Journal of Computations & Modelling, vol.1, no.2, 2011, 97-126
ISSN: 1792-7625 (print), 1792-8850 (online)
International Scientific Press, 2011

The Notion of "Meaning System" and its use for "Semantic Search"

Lin Liu[1] and Junkang Feng[2]

Abstract

Mingers [16] suggests a notion of 'meaning system' in order to clarify the relationships between data (signs), information and meaning, and their bearings on information systems (IS for short). We observe that there are a few important points that need further investigation, which are centred on the basic notions of information and meaning. In this paper, we summarise the most seemingly influential studies on these two concepts in the field of information systems especially Dretske's *semantic theory of information*. Then we take a close look at the notion of 'meaning system' by drawing on Dretske [8] and Devlin [6] in addition to Mingers [16]. We conclude that the term of 'meaning system' may be seen referring to an overall system whose elementary components are the cognitive systems whereby an human agent generates meaning in responding to stimuli, and it is through such a system an human agent interacts with the world including other humans in which he/she is, which includes crucially access to and

[1] Database and Semantic Web Research Group, University of the West of Scotland, e-mail: lin.liu@uws.ac.uk
[2] Database and Semantic Web Research Group, University of the West of Scotland, e-mail: junkang.feng@uws.ac.uk

Article Info: *Revised* : November 15, 2011. *Published online* : November 30, 2011

subsequently creating information. We explore how this notion may be applied to IS problems by formulating one’s meaning system using an ontology language in order to improve web searching.

Keywords: Information, Meaning, Information content, Propositional content, Meaning system, Ontology, Semantic Web Search

1 Introduction

In the literature, ‘meaning’ is taken as synonymous to *the semantic content* of a *concept* [8, p. 222]. Mingers extends Dretske’s concept of meaning to include some seemingly strong and arbitrary features, “*meaning is generated from information by interpreter, carried by sign through a process of digitalization that abstracts only some of the information available* [16, p. 10]”. Minger’s notion of ‘meaning’ consists of three levels:

1. *Understanding*, the primary or literal meaning of a sign. This corresponds to the semantic content of a sign. This level of meaning is commonly shared by all competent cognitive agents of a community, e.g., what a sentence refers to directly. This is because they invoke the same concepts and the instantiation process can hardly go wrong. Such meaning is embodied in the signal (the sentence in the above example), thus it has *objective* features in that the agent does not contribute anything to it.
2. *Connotation*, secondary meaning. This extends the initial meaning of the sign to include nested consequences known and available to a receiver. This level of meaning is *inter-subjective,* which is captured by a group of people who share the same cultural background and language. But different groups of people may obtain entirely different connotation from a given sign.

3. *Intention,* the third and individual meaning, which is realised by a particular person based on his own personal experiences, feelings and motivations at a particular time. As a result, appropriate action is taken, likes above "It is snowing" example. Therefore, this level of meaning is *subjective*.

Thus, by 'meaning' Mingers [16] also refers to the significances to and the purposes and intention of a cognitive agent that perceives a sign/signal. Putting all these together, Mingers suggests an overarching notion of 'meaning system' within which information system (IS) as technological systems is an integral part.

We observe that the approach embodied by the notion of 'meaning system' is helpful in understanding the nature of IS and in looking at the relationship between data (i.e., a type of signs), information and meaning. However, we also believe that some fundamental concepts that Mingers (ibid.) refers to should be further investigated. For example, above-cited Minger's claim can be boiled down to simply that 'meaning is an integral part of information'. This is questionable, as if it is not continentally true, the meaning of a signal is not part of the information that the signal carries. Moreover, to include 'intention' in 'meaning' is confusing and contradicts Minger's fundamental claim that meaning is part of the information carried by a signal, as intention is subjective and can vary depending on individuals, which leads to an unreasonable conclusion that at least part of the information that a signal carries is decided subjectively by individuals that receives the signal.

Thus we believe that the notion of 'meaning system' should be clarified, further developed and made applicable to IS problems. In this paper, we report our work thus far along this line.

We summarise main viewpoints concerning 'information' and 'meaning' in the literature in Section 2. We give our view on these notions in Section 3. These two sections have to be elaborate as they tackle some fundamental notions for our objective. Then in Section 4, we look at the relationship between IS and meaning system. In Section 5, we show how the notion of meaning system may be applied to the problem of user profiling so that Web search may be more relevant to individual users before concluding the paper in Section 6.

2 Classic Views on the Nature of Information and Meaning

We are living in an ocean of information. Information and representations (signs) of information exist everywhere. Information is generated at every moment of time. A small object (sign) is capable of containing and conveying potentially vast amount of information. Despite of being such an important element to mankind, information seem still an ‘*explicandum term’* [9, p. 351] in academic communities today. People tend to use the word “information” on a daily bases without thinking where its concept lie. Moreover, many believe information is closely related to computing or intelligent life and cannot exist without human cognition. In the past decades, the notion of information was studied by many leading philosophers in different aspects. The Mathematical Theory of Communication proposed by Shannon [21] justifies the statistical attributes of information. In terms of the semantic aspect, Dretske’s [8] Semantic Theory of Information has a fundamental significance to the study of the content of information. Barwise and Seligman [2] developed the Information Flow Channel Theory that enables one to identify information flow between systems with the notion of ‘distributed systems’.

Despite of these well established theories about information, the debates around information have never stopped. Particularly, what is the true nature of information, and is it possible to give a single and universally accepted definition to information? Information has been referred to as processed data, the propositional content of a sign, data plus meaning and many more. Moreover, various natures are being attributed to information including objective, subjective and combinations of both. Therefore, finding a clear, justifiable, and applicable concept of information becomes increasingly vital for academic researchers and society as a whole.

The study of information can be traced back many centuries. According to Harper [15], the notion of "information" is originally invented in 1387 with the definition of "*act of informing*". It was referred to as "*knowledge communicated*" a century later. The development of modern technology has inevitably multiplied the number of definitions for information with varying degrees of complexity. Among them, a common view is that *information is data that has been processed in some way to make it useful for decision makers*, which is revealed by Lewis's [14] survey of 39 IS texts. Information embodies an objective nature according to this assumption, because data is objective and independent to its observer in term of its existence and structure. Mingers [16] argues that *"Information is the propositional content of a sign* [16, p.6] [8, p. 65]". The generation of information is due to reduction in uncertainty of what might have happened.

Bateson suggests that information is *a difference that makes a difference* [3, p. 286], which can be interpreted that it is the difference that generates an event, a sign, a symbol, or an utterance.

Subjectivists Lewis and Checkland believe that information exists within human's cognition. As Lewis argues, *"Different observers will generate different information from the same data since they have differing values, beliefs, and expectations* [14]". Moreover, Checkland formulates this view as *"information equals data plus meaning* [5, p. 303]*".* That is, "by attributing meaning to data, we create information".

It is hardly surprising to experience such fierce controversy over the nature of information. Some philosophers have sensed the powerful, elusive nature of information and brought out an impartial idea – the definition of information depends on different fields of requirements. As Shannon points out "*It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field"* [22, p. 180]. Floridi further emphasises, *“It (information) can be associated with several explanations, depending on the cluster of requirements and desiderata that orientate a theory* [9, p. 351]*”* .

Some philosophers pay their attention on defining other attributes of information. Shannon is the founder of the Mathematical Theory of Communication [21], which focuses on the statistical perspective of information. The basic idea of this theory is that information can be accurately quantified as long as the unlikeliness, i.e., the probability, of the event is known. Philosophers and mathematicians such as Barwise and Seligman [2] and Devlin [7] developed and formulated the Information Flow Channel theory and the Infon theory. Their motivating idea is that information flow is made possible by regularities in distributed systems. *Constraint*s capture what (information) flows, and *channel*s reveal why such flow takes place. For example, a constraint concerning a tree trunk could be *‘Number of rings’* ⇨ *‘Age of tree’.*

Meaning is most commonly used in the field of linguistics, e.g., Semantics, although it plays equally important roles in non-linguistic field like Semiotics. The notion of ‘meaning’ may seem simple, but in reality, the characteristic of meaning is that it is far too ambiguous and hard to define. Furthermore, understanding the relationship correctly between information and meaning is crucial since this decides how IS and meaning system are related.

The study of meaning has the same prolonged history as information. In the past, meaning was referred to as tenor, gist, drift, trend, purport, sense, significance, intention, etc. Grice [12, pp. 377-388] divides the convention of meaning into two categories, *natural meaning* and *non-natural meaning*. The natural meaning is close (if not equivalent) to the ordinary sense of “information”, for example, a blown fuse means the circuit has been overloaded. That it is raining means that the grass is wet. Non-natural meaning is relating to language and semantic studies. In this sense, that it is raining means that rain is dropping down from the sky.

In term of how to define it, Cang and Wang say "*meaning is the link between information and data* [4, p. 2]", which is concerned with communication between people that is completed by the realization of meaning from data to information. In their view, the meaning of information carried by data is just a representation and reflection of the essential integration of objectivity and subjectivity in people's lives. It would appear that their notion of 'meaning' is concerned with what a piece of information means to an individual rather than the literal or conventional meaning of a sign, i.e., what the sign directly refers to.

As a great epistemologist, Dretske has this insight on meaning: *meaning is the semantic content of a concept* [8, p. 222]. It is the propositional content of a structure that exhibits a *higher order of intentionality* [8, p. 176]. Meaning generation consists of conceptualization by digitalizing analogue information and therefore converting it into a semantic digital form.

In Mingers' notion of 'meaning system', as cited earlier, "*meaning is generated from information by an interpreter, carried by sign through a process of digitalization that abstracts only some of the information available* [16, p.10]. According to him, meaning can be divided into three levels, i.e., understanding, annotation and intention. It emphasises human agent's involvement in producing meaning and its implementations to mankind.

Devlin proposes linguistic meaning as a linkage between utterance type and actual situation type. "*The meaning of an assertive sentence Φ is a constraint, an abstract link that connects the type of an utterance of a sentence Φ with the type of the described situation* [7, p. 221]".

3 Our Attempt to Clarify the Notions of Information and Meaning

As aforementioned, due to the elusive and diverse nature of information and meaning, it is extremely hard to find a completely safe way of talking about them, in particularly, an explicit definition covering all appropriate aspects. Our attempt lies on finding a clear conception of 'information' and 'meaning'.

3.1 Sign and its impact on defining information

The nature of information has a significant impact on how to define it. A piece of information can be embodied (represented) and carried by a sign (data are a collection of signs). *The sign signifies something, or rather, it signifies that some event has occurred. It also has implications for the receiver* [16, p. 7]. Anything can be a sign as long as it is 'signifying something, referring to or standing for something other than itself'. A sign is an integration of Representamen (vehicle), Interpretant (sense) and Object (referent) according to Peirce's triadic model [19]. Stamper constructs an organisational semiotic framework, which consists of six levels (properties), namely, *Physical World, Empirics, Syntactics, Semantics, Pragmatics* and *the Social World* [23].

Sign may be seen within an information context. Information can be physically carried by a representamen (i.e., the sign) with some syntactic property as described in Stamper's semiotics framework. The interpretant is implication (significance) of other objects, which can be seen as meaning of the sign. This is at the semantics level of semiotics. For example, a traffic light is a sign. The information that the sign carries is an instruction to traffic. When it turns red in a normal circumstance, for instance, the instruction is 'to stop', which is the meaning of the sign and at the same time one of the pieces of information that it carries. If the traffic light turns red in testing, the meaning of it would still be ‘to stop’, but it does not carry the information of ‘to stop’ as there is no such instruction to traffic in the first place.

Despite the connection between sign and cognitive agents (human beings) in the social world, despite the abilities of cognitive agents in generating information through signs, e.g., traffic signs, the making of the sign is independent of its observer if any, and after a sign has been made, it is an objective commodity that exists independently of its observer if any. Therefore, information as a constituent of a sign (i.e., what a sign can tell us truly) is objective, independent of its carrier (sign) and receiver. It is not created in the mind of the observer of the sign. Once the sign is made, e.g., the utterance of a speaker is out, the information is there no matter who receives it.

How much and what information is available to each individual may vary depending on receiver’s prior knowledge about information source. This is so called ‘*relativization*’ [8, p. 79] of the information content of a signal, which should not being confused with being arbitrary. Lewis’ argument in previous section should be refined as different observers will *receive (not generate)* different information from the same data since they have differing values, beliefs, and expectations.

3.2 Arbitrary and negative information

It may be argued that information can be produced in a human’s mind due to reduction of uncertainty occurring in it. For example, a person stops in front of traffic lights thinking about what he is having for his dinner. The uncertainty is thus reduced since he selects one option, e.g., the fish supper out of other possible choices. However, the “information” generated in people’s minds is not within the domain of information that we discussed above, which is concerned with the states of affairs of the real world, and not something in people’s minds. Moreover, such reduction in uncertainty is not carried by a sign (e.g., traffic lights) but by cognitive states.

Moreover, as an information carrier, a sign has implication to its receivers (Mingers, 1995), which is echoed in Dretske’s nuclear sense of information: A state of affair contains information about X to just that extent to which a suitable placed observer could learn something about X by consulting it (Dretske, 1991, p. 45). Information is capable of yielding knowledge and knowledge requires truth, information requires it too. This truthfulness is a necessary condition for DOS (declarative, objective and semantic) information (Floridi, 2005). Therefore, mis-information or false-information is not information, more precisely it is not in our nuclear sense of information, because they are not true. It could be ‘negative information’(ibid.) (i.e., not information at all) generated due to the *equivocation* or *noise* in a process of information transmission or purely the receiver’s misunderstanding.

The above objective characteristic of information is much clearly taken on-board by the Cambridge dictionary of philosophy, which defines information as: "*an objective (mind independent) entity. It can be generated and carried by messages (words, sentences) or by other products of cognizes (interpreters). Information can be encoded and transmitted, but the information would exist independently of it encoding or transmission*."

Information is also measurable, as long as the probability *P* of an event is known. Let s_a be a state of affairs among a few others of a selection process S, then

Surprisal $I(s_a)$ - the amount of information generated at S can be calculated:

$$I(s_a) = -\log P(s_a) \tag{1}$$

Where $P(s_a)$ is the probability of s_a.

3.3 Information content of a signal

The informational content of a signal can be referred to as the *message it bears* [8, pp. 55, 60], e.g.,'s is a person' from a portrait. In general, the informational content of a signal is only the descriptive or conceptual elements embodied in the predicate expression ("S is F"). *The subject term merely attaches that content* (*F*) *to a particular individual* (*S*) [8, p. 67]. The information content of a signal must be contingently true, which implies that the information content has to meet truth condition in order to qualify, but no all true state of affairs carry information. Information is generated only due to the reduction of uncertainty. For instance, "Earth is flat" is not information because it is not true, neither is "Earth is round" because it is constantly true.

Dretske's definition of *informational content*:

> *A signal r carries the information that s is F = the conditional probability of s's being F, given r (and k), is 1 (but, given k alone, less than 1)*.

To make it simple, If P(s is F | r) = 1, then r has s being F as part of its information content.

The above definition concentrates only on the level of types of events. However, it is *particulars*, i.e., individual things in the real world that carry information [2, p. 27], e.g., "Earth is on quake in location l at time t". Xu and Feng [24, p. 3] modify the above concept using the notion of '*Particular*' presented by Barwise and Seligman

Ri carries the information that there must be some sj existing at time tj and location lj, that is, the state of affairs of s is F at tj and lj, if and only if the conditional probability of s's being F given r is 1 (and less than 1 given k alone).

No single piece of information is entitled to the status of the informational content (i.e., the whole information content) of a signal except that the signal is a cognitive state whose propositional content exhibits the third order of intentionality (7, p.173). The receiver of the signal may be more interested in one piece of information than others, he may succeed in extracting one piece of information without another, but these differences are irrelevant to the information the signal contains [8, p. 72].

Unlike information that its amount can be transmitted in different percentage, the transmission of information content is simply an ‘all or nothing’ affair. It does not make sense to speak of sending out 90 percent of the information that it is raining today.

A signal that has A in its informational content will also include all the nested information of A as part of its informational content. This is the distinguish feature to signal’s *semantic content* which is illustrated in detail later. *S’ semantic content is unique in the way that its informational content is not* [8, p. 178].

3.4 Propositional Content of a signal

After defining information content, we are in the position of explaining a very abstracting term of *propositional content*. The basic element of propositional content is proposition. According to Oxford Dictionary, *a proposition is a sentence expressing judgment or opinion.* In philosophy, a proposition is identified ontologically as an idea, concept, or abstraction whose token instances are patterns of symbols, marks, sounds, or strings of words. Devlin declares *a proposition is a claim about the world to the effect that a certain object is of a certain type* [7, p. 62].

Propositional content of a signal is the proposition embodied by the signal as a result of cognitive process (recognition, identification, and classification). Explicitly speaking, the propositional content is an idea, concept or abstraction expressed by a signal and the creator of the signal must possess some sort of cognitive state to be able to purpose this propositional content to the signal. Therefore, not every information-bearing sign consists of propositional content. For instance, the traffic light has propositional content of traffic instructions, whereas the tree stump in the forest does not contain any propositional content at all.

Cognitive states always, either explicitly or implicitly, have a specific propositional content [8, p. 154]. We know (or believe, or judge, or think) that s is F (identify, classify, or categorize s as F). We have a variety of ways to describe our cognitive states, e.g., Herman realizes that the wine has gone bad.

Propositional content should not be confused with perceptual content of the cognitive state. We perceive (see, hear, smell, taste, and feel) objects and events, which qualifies as perceptual statement. The perception is a process in which sensory information is coded in analogue form in preparation for cognitive utilization (digitalization) [8, p. 154].What determines what we perceive (what object or event) is not what we believe (if anything) about what we perceive, e.g., one can see a pentagon and think it is a square.

3.5 Semantic Content of a signal

The propositional content of a signal is very closely related to the semantic content of the signal given the fact that they are all the results of cognitive activities. As Dretske points out, t*he semantic content is the unique propositional content exhibiting the third order intentionality* [8, p. 173].

Third Order of Intentionality:

(a) It is analytically necessary that Fs be G
(b) S has the content that t is F
(c) S does not have the content that t is G

To illustrate it, one might know t=2 without realizing that t is square root of 4 if his cognitive state exhibits a third order of intentionality.

However, the differences between propositional content and semantic content are apparent. A signal may have more than one propositional content, whereas the semantic content is unique and there is nothing else for a cognitive state.

A signal may carry more than one piece of information and a piece of information may be nested in another. The *outmost information shell* in which all other information is nested either nomically or analytically is said being carried in *digital* form [8, pp. 177-8]. S carries the information that t is F in digital form if and only if that is the most specific piece of information about t that S carries.

To apply the above notions to databases, then data are created by some human originator, and normally have propositional content. However, data in a machine (non-human and therefore there is no cognitive capability) do not have the *third order of intentionality*, therefore data or databases do not have semantic content but only propositional content, which may have many other pieces of information nested in it. These are sources of information for a user of the database to apply a process of digitalisation in order to receive the information that interests them. That is to say, from the semantic content of some intention (which is a cognitive state) of the database originator, which is only one piece of information without anything else nested in it, to the information content of the data that is placed into the database by the operator, the number of pieces of information is greatly increased which would undoubtedly add burdens to database designers in terms of how to convey the explicit information to the users.

3.6 Concept

A concept is a type of internal structure: one whose semantic content, when instantiated, exercises control over system output (7, p. 214). This definition has following consequences. Concepts arise due to the cognitive agent having *selective sensitivity* which allows cognitive agent to build an internal structure known as a *concept* from coding information origin, e.g., classifying swan by observing it. After a concept is formed, it becomes part of internal language consisting of syntax and semantics. When the cognitive agent receives information (input), e.g., seeing a particular swan in a pond, it is digitised and matched against an internalised concept, this concept is then instantiated and the semantic content of a cognitive state arises due to this matching. The semantic content of concept attaches *meaning* (output), such as the swan is a flying bird, and prefers fish and bread, to the incoming information. A concept is determined by its extensions (the instances of the concept). No two concepts are the same unless they have the same extensions.

It should be noticed that a concept can be mistakenly applied through instantiation, the cause of which is that there is no relevant information available, but there is still *meaning* that is given by the concept involved. For example, one may look at a goose and identify it as a swan because he wrongly applies the concept of swan to the goose. This is why the meaning of an instance of a concept is independent of whether this instance carries relevant information. In other word, the *meaning* of an instance of a concept is not necessarily part of the *information* that this instance of concept carries. Moreover, the instantiation of a concept gives rise to a 'belief'. If it is supported (i.e., caused or/and sustained) by relevant information, a belief becomes knowledge.

Once it is established in an agent's mind, a concept has the life of its own, i.e., it retains its semantic content even if the information originally used for the concept to be established is no longer available. However, a concept is not a fixed perception. It is constantly updated on its properties as long as learning process continues. That leads to explanation of why a child's perception is quite different from an adult's.

3.7 Meaning and its relations

We adopt Dretske's semantic (non-natural) approach of meaning to discriminate it from information. *Meaning like other cognitive states, e.g., belief and perception, exhibits a higher order of intentionality than that associated with* *a structure’s informational content* [8, p. 175]. As discussed in the previous section, the meaning of a signal if any is realized by instantiation of an internal concept, which attaches the *semantic content* of the concept to the instance. The semantic content of a concept is initially formed through an information handling process of which some piece of information is being digitalised by a cognitive system and becomes the semantic content of an internal cognitive structure. Any information that is carried in analogue form that is nested in the piece of information digitalised is excluded from the semantic content. For example, the utterance “Sean is a male adult” does not convey meaning of “Sean’ age is equal to 16 years or over” or “Sean is not a female”, although they (as information) are nomically nested in ‘Sean is a male adult’ if it is contingently true.

In terms of digitalisation, different cognitive systems may abstract different pieces of information from those that are carried by a signal depending on its cognitive ability and intention etc, e.g., experience, knowledge and understanding. A broadcast statement “it is snowing” carries a lot of information. It can be interpreted to be cold by an elderly man’s cognitive system and he stays in. Conversely, a boy next door is quite excited to hear it. He is expecting a snow ball fight and rushes out.

A signal might well have meaning without carrying any information. For instance, the utterance “it is raining” has meaning, but carries no information if it is not true (not raining). However, it should be pointed out, although the meaning generated in this example is not from information carried by “it is raining”, but it still comes from other sources, e.g., mis-information or negative information [9, p. 354] mentioned previously, which are not information at all. That is to say, in this example, the concept ‘raining’ is mistakenly instantiated possibly due to mis-information etc.

Unlike information, which can be measured as said earlier, meaning is not measurable. Meaning cannot be measured by the probability of an event. “It snows in July” does not have more quantity of meaning than “It snows in December” even though the former carries a larger account of information than the latter as the probability of the former is lower than that of the latter.

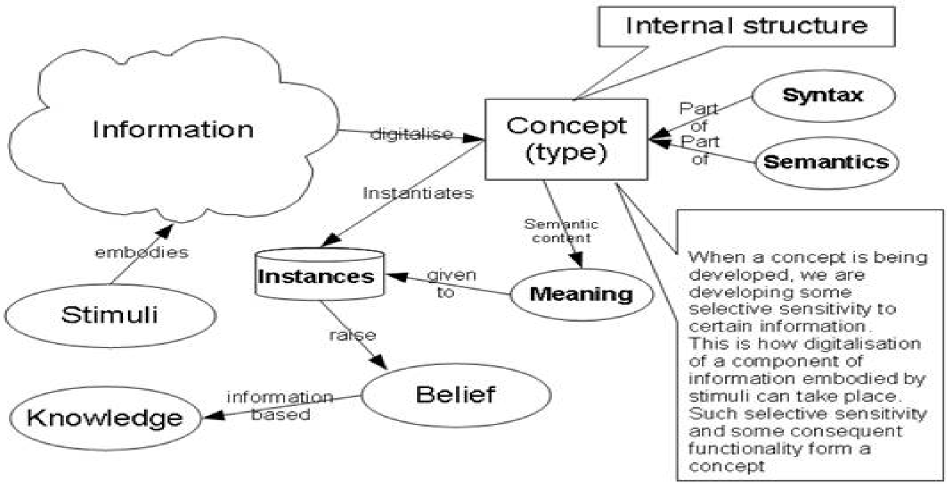

Figure 1: The overall relationship of information-theoretic system between information, concepts, meaning (Courtesy to Douglas Salt, Dr Feng on meeting note August 2010).

4 'Meaning System' Redefined and IS

By 'meaning system' therefore we refer to a humans' epistemological system based on perception and cognition of information from which meaning, is produced through instantiation of a concept. It would seem that such a system has to have a set of concepts as it necessary constituent. To receive information from a signal/sign involves digitalising the information carried by the sign/signal by applying and thus instantiating concepts. Information therefore is untouchable directly to human agents in the sense that humans can only interact with information through their meaning systems and the meaning systems interact with the carriers of information, i.e., signs and signals. Information cannot be received by a human agent until it connects to their meaning systems within which human beings operate. By doing the above we amended Minger’s [16] notion of meaning system. We suggest that the term ‘meaning system’ be used to refer to an overall system whose elementary components are the cognitive systems whereby an human agent generates meaning in responding to stimuli, and it is through such a system an human agent interacts with the world including other humans in which he/she is, which includes crucially access to and subsequently creating information.

We observe therefore that the ‘meaning system’ is indispensible for an information system to work. Information is necessary for knowledge [8] as the latter is information supported belief. In an information system (IS), most information is held in analogue form in the sense that they are not the most specific information that data carry. It is through human’s meaning system a user receives information through digitalisation of it.

Information systems are ultimately designed to serve mankind. Traditional IS implementations are concerned very little with individual requirements; they treat users as a whole group. We observe that information systems should adopt the notion of ‘meaning system’ in developing user-oriented applications, e.g., Web searching, online shopping, digital libraries, and so on, and modern technologies should facilitate it by providing useful mechanisms.

5 Applying the Notion of ‘Meaning System’ to Web Search

We suggest that the notion of ‘Meaning system’ have the potential of a wide spread application across disciplines. We now take a look at *Semantic Web Search* as an example to demonstrate the concept’s significance in the IS field.

5.1 Contemporary Approaches to Web Searching

The mass of content available on the World Wide Web raises critical question over its effective use. The web is largely unstructured with pages authored by many people on a diverse range of topics, making simple browsing too time consuming to be practical.

Traditional search engines, e.g., Google, Yahoo, Sohu, have been designed to work with natural languages, searching web pages by matching explicit queries. However, it is hard for people to articulate what they want, especially if they have limited capability of vocabulary such as keywords. Therefore, the large lists of search results often contain only a handful of useful pages if ever, due to the poor formulated queries. On the other hand, those search engines do not work well with documents encoded in the Semantic Web languages RDF and OWL. They do not understand conventions and structural information such as those involving XML namespace.

The *Semantic Web* technology seems an intelligent way to avoid the obstacles of query searching and to find the Semantic Web Documents on the web. It allows search queries based on web pages marked up with semantic metadata. However, this technology very much depends on the degree of how authors annotate their web pages, and automatic web page annotation is still immature. Furthermore, the vast majority of web pages are constructed without annotation, and it will remain the same in the foreseeable future. Thus, the semantic web is offering limited benefit to the problem of effective searching.

5.2 Ontologies

An ontology as defined in [16] is a formal explicit description of a domain in a human-understandable, but machine-readable format, consisting of *classes*, which are the concepts found in the domain (also called entities). Each class may have one or more parent classes (is-a or inheritance links), formulating thus a specialization/generalization hierarchy; a class has *properties or slots* (also called roles or attributes) describing various features of the modelled class, and restrictions on the slots (also referred to as facets or role descriptions). Each slot, in turn, has *a type* and could have a restricted number of allowed values, which may be of simple types (strings, numbers, Booleans or enumerations) or instances of other classes. Classes may have *instances*, which correspond to individual objects in the domain of discourse; each instance has a concrete value for each slot of the class it belongs to. An ontology together with a set of individual instances of classes constitutes a *domain of knowledge*.

Ontologies have some prominent advantages over other data representation systems such as databases. Firstly it has richly structured metadata sources and web-accessible format. Secondly, ontologies are capable of providing a shared understanding of a domain. Such a shared understanding can be used to overcome terminology ambiguity of search queries. For instance, one application's zip code may be the same as another application's post code. Conversely two applications may use the same term with different meanings. For example, in university A, a course may refer to a degree (like computer science), while in university B it may mean a single subject (e.g., cs101). Such ambiguities can be overcome by mapping the particular terminology to a shared ontology or by defining direct mappings between the ontologies. In either case, it is easy to see that ontologies support semantic interoperability [17].

In addition, by using ontology, web searches can exploit generalization/specialization information. If a query fails to find any relevant documents, the search engine may suggest the user using a more general query. It is even conceivable for the engine to run such queries proactively to reduce the reaction time in case the user adopts a suggestion. Or if too many answers are retrieved, the search engine may suggest to the user some specializations. The most well-known ontology languages are XML, XMLS, RDF, RDFS and OWL. The interested reader is referred to [1].

5.3 User Profiling

The notion of *user profiling* has been introduced in order to record the user context and personalize applications so as to be tailored to the user needs [10]. User profiling is commonly employed nowadays to enhance usability as well as to support personalization, adaptivity and other user-centric features.

Traditionally, the user profile model is constructed on the base of canonical (or homogeneous) user. However, often individual users vary so much that a model of a canonical user model is insufficient. Thus building an individual user profile in the system is highly desirable for modern IS applications, in particular web searching, to filter out the loosely related keyword results.

Ontologies have been proven an effective and efficient tool for modelling user profiles because they may present an overview of domain related to a specific area of interest and be used for browsing and query refinement [10]. Up to this point, Ontologies modelling user profiles are application-specific, with each one having been created specifically for a particular domain. A standard extendable ontology is much needed to model user profiles. More importantly, the reasoning within ontology in current models seems pretty straight forward, and need to be improved.

5.4 Semantic Web Search

Semantic Search seeks to improve search accuracy by understanding searcher intent and the contextual meaning of terms as they appear in the searchable data space, whether on the Web or within a closed system, to generate more relevant results [Wikipedia]. Unlike keyword search or relevancy prediction, Semantic Search uses semantics, which is the meaning of language, to produce highly relevant search results.

The notion of *Semantic Web Search* is referred to various semantic search engines which indexes RDF or OWL data stored on the web and provide an interface to search web pages through crawled data. Semantic Web Search is still in its early stage of development and like Semantic Search, it does not support personalization. Some semantic web search engines that are currently under development are Semantic Web Search Engine (SWSE), Swoogle, Sindice, Yahoo!Microsearch, Falcons and Semantic Web Search.

5.5 Applying ‘meaning system’ to improve Semantic Web Search

In the previous sections, we demonstrated that the notion of *meaning system* is concerned with how an individual receives information through digitalizing incoming information by applying relevant concepts the instantiation of which generates meaning, which is then in turn give rise to belief. And moreover, if the instantiation of concepts is supported by relevant information and in other words, the belief is caused and/or sustained by the relevant information, then the individual gets *to know* something about the real world. As such, the cognitive states of an agent are the centre part of his/her meaning system. We suggest that the ‘meaning system’ of

The merits of ontologies, in particular its ability of digitizing collections and user context [10], have revolutionarily established ontology to be a prime cognitive language in modelling user’s intents. Therefore, it becomes an absolutely crucial tool in building user-oriented IS applications.

In the context of semantic web search, the search system/engine can take reference of user profiles to analyse and reformate queries and thereby to generate the search results that best march the user’s interest. We suggest that the *meaning* *system* of an individual be captured and formulated using ontologies to serve as the main content of a user profile.

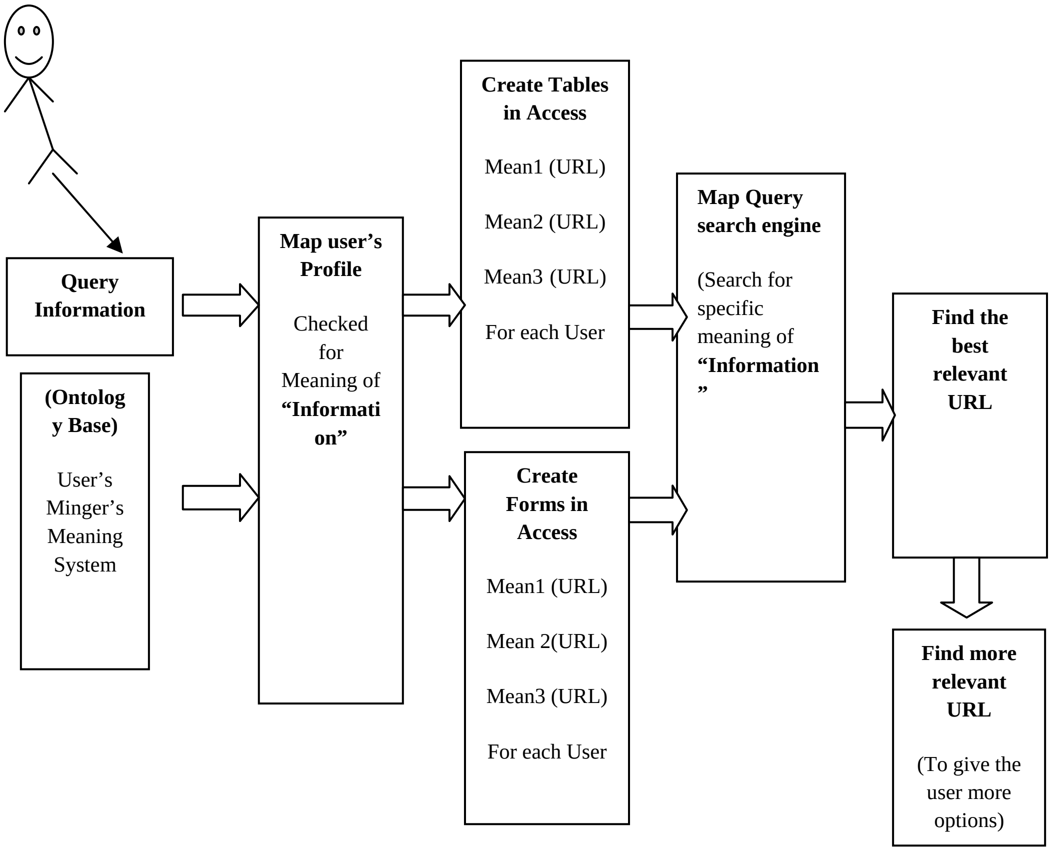

Figure 2: Semantic Web Search System using the notion of 'meaning system' [20]

1. The attempts to be described below are aimed at capturing meaning system using ontologies. The material to be presented here is based on [20]. This approach utilizes user profile to interpret queries. For instance, for "IT", the search engine will produce different URLs for each individual user (or group), e.g., a computing student, a scientist, or an NHS nurse by following their respective meaning systems.

We summarize the procedure and strategy of our system for semantic web search that makes use of the notion of ‘meaning system’ below:

1. After the query has been made by the user, the web search system abstracts the primary meaning for the query (e.g., the term “information”).
2. In order to make the search more specific so that to find the best results, the search system will find out what is meant by the term ”information” for this specific web user by citing his/her profile, which has been stored in the system.

Appropriate search results—URLs are selected and brought forward. Those URLs are directly linked to the primary meaning of the query under this particular user profile. For instance, for the search results on “information”, the relevant URLs will be quite different for IT professionals, philosophers, lawyers, doctors, and so on.

5.6 A Simple User Profile

We use ontology to capture and formulate user profiles. A simple user profile is built here, which may be extended through inheritance, the addition of more classes as well as concept instantiation according to the needs of a specific application. A top-down approach is adapted in conjunction with Gruber’s design criteria [11] (clarity, coherence, extensibility, minimal encoding bias and minimal ontological commitment). The most central class of ontology is “Person” which contains all the user profile characteristics such as “name” and “date of birth”. The rest of classes are Education, Profession, Expertise and Interest, which are used to describe the complex user characteristics.

This user profile ontology below is written in the OWL language, which captures a top class called “Person”. This profile is an integral part of our web search system.

```
<owl:Class rdf:ID=”Person”>
        <owl:DatatypeProperty rdf:ID=”name”>
          <rdfs:domain rdf:resource=” Person”/>
                <rdfs:range rdf:resource=”&xsd;string”/>
        </owl:DatatypeProperty>
        <owl:DatatypeProperty rdf:ID=”dateofbirth”>
           <rdfs:domain rdf:resource=” Person”/>
                <rdfs:range rdf:resource=”&xsd;integer”/>
        </owl:DatatypeProperty>
        <owl:DatatypeProperty rdf:ID=”gender”>
          <rdfs:domain rdf:resource=” Person”/>
                <rdfs:range rdf:resource=”&xsd;string”/>
        </owl:DatatypeProperty>
</owl:Class>

<owl:Class rdf:ID=”Education”>
        <rdfs:subClassOf rdf:resource=” Person”/>
        <owl:DatatypeProperty rdf:ID=”degree”>
          <rdfs:domain rdf:resource=” Education”/>
                <rdfs:range rdf:resource=”&xsd;string”/>
        </owl:DatatypeProperty>
        <owl:DatatypeProperty rdf:ID=”level”>
           <rdfs:domain rdf:resource=” Education”/>
                <rdfs:range rdf:resource=”&xsd;integer”/>
        </owl:DatatypeProperty>
</owl:Class>

<owl:Class rdf:ID=”Profession”>
        <rdfs:subClassOf rdf:resource=” Person”/>
        <owl:DatatypeProperty rdf:ID=”Occupation”>
          <rdfs:domain rdf:resource=” Profession”/>
                <rdfs:range rdf:resource=”&xsd;string”/>
        </owl:DatatypeProperty>
</owl:Class>

 <owl:Class rdf:ID=”Expertise”>
        <rdfs:subClassOf rdf:resource=” Person”/>
        <owl:DatatypeProperty rdf:ID=”skill”>
```

```
        <rdfs:domain rdf:resource=” Expertise”/>
              <rdfs:range rdf:resource=”&xsd;string”/>
      </owl:DatatypeProperty>
      <owl:DatatypeProperty rdf:ID=”depth”>
         <rdfs:domain rdf:resource=” Expertise”/>
              <rdfs:range rdf:resource=”&xsd;string”/>
      </owl:DatatypeProperty>
</owl:Class>

<owl:Class rdf:ID=”Interest”>
      <rdfs:subClassOf rdf:resource=” Person”/>
      <owl:DatatypeProperty rdf:ID=”business”>
        <rdfs:domain rdf:resource=” Interest”/>
              <rdfs:range rdf:resource=”&xsd;string”/>
      </owl:DatatypeProperty>
      <owl:DatatypeProperty rdf:ID=”sports”>
         <rdfs:domain rdf:resource=” Interest”/>
              <rdfs:range rdf:resource=”&xsd;string”/>
      </owl:DatatypeProperty>
      <owl:DatatypeProperty rdf:ID=”others”>
         <rdfs:domain rdf:resource=” Interest”/>
              <rdfs:range rdf:resource=”&xsd;string”/>
      </owl:DatatypeProperty>

</owl:Class>
```

6 Conclusions

In this paper we have taken a look at the notion of ‘meaning system’, and we explored how it might be clarified and extended. We also looked at how it might be applicable through an experimentation of a Semantic Web Search system. To this end, we have provided an analysis of two fundamental but controversial elements: “information” and “meaning”. We subscribe to the viewpoint that information is an objective commodity. It exists independently of the carrier, or receiver, if any, although the quantity and quality of information available to each receiver may vary depending on their background knowledge about information source. But this relativization of information should not be seen as evidence that information *per se* is subjective, and it is only a matter of how the same information source is looked at. We discovered that it is the *cognitive states* (internal structures in human's mind) that are responsible for the creation of *concept*s whose *semantic content* is capable of giving *meaning* to its instances. Thus the creation of meaning involves instantiation of a concept as a response to some stimuli (signals), which an agent's cognitive state will then bear. Furthermore, the instantiation of a concept does not require the availability of the same information that is used to create the concept in the first place, and as a result, meaning may or may not be part of the information that a signal carries. This is why meaning may be objective, inter-subjective or subjective.

We conclude that the term of 'meaning system' may be seen referring to an overall system whose elementary components are the cognitive systems whereby an human agent generates meaning in responding to stimuli, and it is through such a system an human agent interacts with the world including other humans in which he/she is, which includes crucially access to and subsequently creating information.

The notion of 'meaning system' seems useful in designing IS applications. It provides us with a way to convert hard and technology-oriented IS into soft and user-oriented ones. Ontology seems a promising tool for identifying and formulating a meaning system because its main components are concepts, which match the basic structure of ontology well. Our experimentation with semantic web search seems to have given preliminary evidence to support such a hypothesis.

References

[1] G. Antoniou, and F. Harmelen, *A Semantic Web Primer*, The MIT Press, 2 Rev. ED edition, (9 May 2008).

[2] J. Barwise and J. Seligman, *Information Flow: The Logic of Distributed Systems*, Cambridge University Press, 2008.

[3] G. Bateson, *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution and Epistemology*, Chicago University Press, 2000.

[4] P. Cang and S. Wang, The analysis of uncertain knowledge based on meaning of formation, *WSEAS Transactions on Information Science and Applications*, **6**(1), (2009), 136-145.

[5] P. Checkland and J. Scholes, S*oft Systems Methodology in Action*, Wiley Chichester, 1990.

[6] W.B. Croft, H.R. Turtle and D.D. Lewis, The use of phrases and structured queries in information retrieval, *Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval*, Chicago, Illinois, United States: ACM, (1991), 32-45.

[7] K. Devlin, *Logic and Information*, Cambridge University Press, 1995.

[8] F.I. Dretske, *Knowledge and the Flow of Information*, Cambridge University Press, 1999.

[9] L. Floridi, Is semantic information meaningful data?, *Philosophy and Phenomenological Research*, **70**, (2005), 351-370.

[10] M. Golemati, A. Katifori, C. Vassilakis, G. Lepouras and C. Halatsis, Creating an ontology for the user profile: Method and applications, *Proceedings of the First RCIS Conference*, (2007), 407-412.

[11] T.R. Gruber, *Toward principles for the design of ontologies used for knowledge sharing*. Originally in N. Guarino and R. Poli. (Eds), International Workshop on Formal Ontology, Padove, Italy. Revised August 1993. Published in *International Journal of Human-Computer Studies*, **43**(5-6), (Nov/Dec. 1995), 907-928.

[12] H.P. Grice, *The philosophical Review*, **66**(3), (July, 1957).

[13] N. Guarino and P. Giaretta, Ontologies and knowledge bases: Towards a terminological clarification. In Mars, N., editor, *Towards Very Large KnowledgeBases: Knowledge Building and Knowledge Sharing*, p. 25-32, Amsterdam, NL. IOS Press, 1995.

[14] P.J. Lewis, Linking soft systems methodology with data-focused information systems development, *Information Systems Journal*, **3**, (1993), 169-186.

[15] K. Lyytinen, H. Klein and R. Hirschheim, The effectiveness of office information systems: a social action perspective, *Information Systems Journal*, **1**, (1991), 41-60.

[16] J.C. Mingers, Information and meaning: foundations for an intersubjective account, *Information Systems Journal*, **5**, (1995), 285-306.

[17] N.F. Noy and D.L. McGuiness, *Ontology Development 101: A Guide to Creating Your First Ontology*, Stanford Knowledge Systems Laboratory technical Report KSL-01-05, March 2001.

[18] L. Obrst, *Ontologies for semantically interoperable systems,* Conference on Information and Knowledge Management, Proceedings of the twelfth International conference on Information and Knowledge management.

[19] Charles Sanders Peirce, *Logic as semiotic: The theory of signs*. In Buchler, J. (ed.), The Philosophical Writings of Peirce, Dover, New York, pp.98-119, 1897-55.

[20] R. Reda, Improving Web Search by Drawing on Semantic Information Theories, *MSc Dissertation* of the University of West Scotland, 2009.

[21] C. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, 1949.

[22] C.E. Shannon, *Claude Elwood Shannon: collected papers*, IEEE Press, 1993.

[23] R. Stamper, The semiotic framework for information systems research, *Information systems research: Contemporary approaches and emergent traditions*, (1997), 515–527.

[24] Xu, Kaibo and J. Feng, *Defining the notion of ‘Information Content’ and reasoning about it in a database*, University of the West of Scotland, Jounel, 2009.