

A Philosophical Attempt at Formulating a Working Definition of Information

Sam Berner
March 18, 2001

1. The Importance of Being Informed

It seems to me that the question of why it is important to know what 'constitutes' information is a linguistic misnomer. The reason for this is that the two keywords in the question are "know" and "information", both of them related to one another in a way that makes it very difficult to distinguish. Since "knowledge" is often defined as "information" and vice versa¹ the question above could be reconstructed as "why it is important to know what 'constitutes' knowledge", an attempt at explaining the outcome of a process through the application of the process itself, generating more output as result, and so on and so on, ad infinitum.

'Information' is not the only such definition-defying concept. To talk about a definition of information, let alone a working one, seems equivalent to trying to define such concepts as the First Cause, the Unmovable Mover, or cognition, mind, and universe. As such, any of the available definitions - including those purporting to be value-free or discipline-independent - may work for some, and be totally useless for others. Moreover, just to add to the circularity of the whole issue, the term 'definition' implies 'information'. When we 'define' something, we presumably know what the object of our definition is, we become 'informed' about it. So here we go again, stumbling through the blurry line that divides the process from its output.

Having said that, the 'fact' (another ubiquitous related term) is that there is no such thing as a Universal Definition/Concept/Theory of Information. There are only attempts at explaining it and its attributes: what can be done with it and why we want it so much that we are willing to invest time, energy and money into accessing it. All rational people know it exists, but no one can give it a precise, containable definition that would meet general consensus. To quote Keith Devlin (1991; quoted in Dembski, 1998):

"That there is such a thing as information cannot be disputed... Our very lives depend upon it, upon its gathering, storage, manipulation, transmission, security, and so on. Huge amounts of money change hands in exchange for information. People talk about it all the time. Lives are lost in its pursuit. Vast commercial empires are created in order to manufacture equipment to handle it."

Although some scholars would say otherwise, the world hasn't been a worse place for the lack of a finite definition. It seems to me that all the attempts at attaining one serve no purpose but the gratification of the human need to classify; a need that started - according to different 'information' sources - either on the African plains or in the Garden of Eden.

A "working definition" is precisely that: it is operational. Since no definition of information works for all and sundry, it follows that information has to be defined in the

context within which it is being used, generated and manipulated. The process must be an inherent part of the definition. It is well and good to engage in the lofty pursuit of a "Theory of Everything"; there is nothing wrong with exercising mental muscle. But such definitions, in my opinion, do not have a utilitarian value to information brokers, providers and managers. I propose to show how this is so in section 4 of this paper.

2. Is Diversity a Good Thing?

To clarify how muddled the whole area of "information science" is, let us now look at a number of different, and often conflicting, attempts at defining the term.

The more renowned attempts have come from such disciplines as communications engineering, computer study, cognitive psychology and philosophy, library science, education and, lately, economics and business. There have also been attempts at all-encompassing, discipline-independent definitions; among these we will discuss the work of Flückiger (1996), Losee (1998), and Hofkirchner (1999). An interesting, although difficult to stomach for the non-scientist, article discussing information in the biological field is Dembski (1998). Kando (1994), Marco (1996), Agre (1997), and the Cristal-Ed Discussion (1997) provide attempts at defining information from the library science point of view. Dervin (1996) and Luke (1996) represent the communications and computer science disciplines, while Sveiby (1996) and Kirk (1999) tackle information from the new discipline of Knowledge Management position.

To start with, it seems that most definition-providers either subscribe to, or criticise Shannon's definition of information. Shannon (1969) was a telecommunications engineer, and his definition is technical and narrow. For him, information was the capacity of a communications channel measured in bits, and each bit is a distinction that is meaningful to the parties on each end of the channel. This is a schematic representation of the transmission of a message from an information source to its destination. The "amount of information" or information content can be given in the form of a rarity value (probability): The less frequently an information element appears in a message, the higher is its information content. The interesting part in Shannon's definition, if I understand this correctly, is that the more information there is, the less meaning is contained in the information. Shannon perceived information as entropy. The logical conclusion is that chaos has more information than a structured pattern, but very little meaning to it, and vice versa.

Dembski (1998) defines information as "fundamentally a matter of ruling out possibilities". For him, information presupposes a contingency, not a medium. He rejects the notion that natural laws can be applied to information, since natural laws are deterministic (map x onto y) while for him information is a complex process that deals with a multiplicity of distinct possibilities (mapping more than one variable onto y). He further classifies information into specified and non-specified, patterned and un-patterned. For Dembski, no information can arise out of chance. His final conclusion is that since information is complex and specified and so is life, then life must have originated from information. In my opinion this sounds more like a deistic interpretation of life than a definition of information.

Dervin (1996), a Communications specialist, gives two definitions in his article. The first assumes that information is a description and ordering of reality, with some form of relationship to what it describes. This definition implies that it is something amenable to distribution through channels (either natural or unnatural). The second definition states that information is something that has always been designed by man, either covertly or overtly. Dervin, as an information "designer" subscribes to the second definition. He discusses different philosophical perceptions of information and its interrelationship with reality, observation and power.

Luke (1996) defines information from the standpoint of an "information strategist". He recognizes information as a means to an end: promoting understanding and acquisition of knowledge as well as providing a basis for actions, decisions and planning.

Noriko Kando is an Information Systems specialist. In his article, Kando (1994) acknowledges the dilemma of ambiguity and the fact that information – at least within Library Science – has to be context-defined. He provides examples of various ways in which the term 'information' is defined. He then proceeds to divide these definitions into two groups – objective and subjective, which are interconnected through a third: information as a process. He reaches a conclusion that there is an information continuum between objective information (content) which becomes subjective information through the process. He also suggests that information and knowledge are different stages of the same process. It has to be said, though, that no where in the whole article does Kando actually define the term 'information'; he only proposes ways in which it could be looked at.

The discussion group at Cristal-Ed (1997) provides a definition of information by Rohde (1986). She defines information in the context of information needs studies as "factual data or advice or opinion, a physical object, such as a book or journal, or the channel through which a message is conveyed, for example, oral or written communication." A few of the contributors to this discussion, interestingly, voice an opinion that using the term should be avoided, since it causes nothing but confusion. However, such an approach would be impossible for the lack of any other, more wide-spread term. One opinion stated was that information is what we use it for – again, context seems to rule.

Joining in the chorus of the disappointed is Marco (1996), a librarian, who categorically calls "information science" a false dogma which should be avoided at all cost by librarians. He quotes Aristotle's description of 'valid definition': interconvertible subject and predicate (if all $x = y$, then all $y = x$), and proposes to define information as "organized data in the form of a verbal or graphic text", thus divorcing the process from the content. Marco correctly argues that we cannot give the same name to different things. This complaint applies very well to the various definitions of information.

Agre (1997), on the other hand, comes as a bit of a surprise, in that he proclaims the concept of information "dead". His argument is that the new advances in technology demand a new concept. He sees information as a bridge-building process, closing the gap between artifacts (objective information?) and meanings in people's lives (subjective information?). He then proceeds to give us three definitions of information: as industrial

material that can be processed; as a “fabric of heaven” designed by techno-prophets; and as homogeneous content that can be stored and retrieved by professionals. He disagrees with the notion that information is in any way homogeneous, and as such amenable to systematisation. From that tenet he postulates the need for a new concept of “collective cognition”, where genres replace information. Genres, presumably, will allow information professionals (librarians) to perceive the heterogeneity of information.

Karl-Eric Sveiby is Australia’s Knowledge Management guru. In his article (Sveiby 1994), he discusses the difference between definitions of information given by Wiener (the father of cybernetics) and Shannon. For Wiener, information is the negative quantity of entropy ($I = E-1$), which means that the more information we lose the more chaos we have, the less meaning available. This is opposite to Shannon’s definition which I discussed above, where $i \propto E$, meaning that the more information we have, the more chaos. After criticising both definitions as inadequate, Sveiby argues that information has no value of its own, and that its value comes out in connection with human action, thus partly subscribing to Shannon’s view.

Kirk (1999) discusses information definitions in the field of information management. Her four definitions are based on a policy-making work by Braman (1989), and in short postulates information as a) thing: resource, commodity b) perception of pattern and c) changing force in the society. These definitions are in agreement to my personal working “definition”, since they come from the same field of work and consequently serve the same interests. Information is a creative process producing tacit or explicit knowledge. In other words, information can be an object or a construct, both being very valuable in corporations. Unfortunately, discussion of Knowledge Management is outside the scope of this paper.

Finally, I will summarise the attempts at a unified theory of information that attempts to cover all fields, disciplines and contexts. Hofkirchner (1999) seems to be writing for a readership that does not eke out a living from re-packaging and selling information. He starts by defining information as a property of self-organizing systems, a process that cannot be formalised, mapped or computed. For him, like for Dembski (1998) information is a process by means of which novelty emerges. Hofkirchner then proposes a theory of information that is flexible enough to relate to most various manifestations of information while at the same time saving research from a reductionism, notwithstanding the fact that the majority of computer and telecommunication scientists, not to say librarians and information brokers, would have major reservations with this “unformalised”, unmeasurable” process. If I may be forgiven the colloquialism, mixing Spielberg with good electronics to produce an E.T. does not solve the defining dilemma, and has been practiced for thousands of years under the brand name of philosophical eclecticism.

Even more oblique is Losee (1998). For him, information is the input, the process and the output - nothing new if we changed the names to objective - subjective; thing - construct, etc. What is new is his assertion that one does not need to understand the nature of the process in order to understand the value of the output. He also adopts the view that all processes produce information as output independently of a receiver (the tree in the forest paradigm). Although the validity of this definition could be argued, its utility cannot. Losee, however, is optimistic that his definition will solve research problems across all

disciplines when taken as a base to which qualifiers can then be added. In my opinion, once we start adding qualifiers, we come out with context-based – and totally different – definitions.

Flückiger (1996), too, postulates the need for a unified theory of information, complaining that the Functional-cybernetic definition (Hofkirchner's property of self-organising systems and its derivatives) and the structural-attributive definition (Wiener's information as structure, order, etc) are incompatible. He then opts for a third definition, based on the neurobiological findings of information as a thing, as an individual's private phenomenon. So now, instead of having a unified definition of information, we have as many different definitions as there are human beings, since each one of them will define his "private phenomenon" subjectively, in his mind, with no objective framework out there somewhere against which to test it.

3. "Working" Definition

As an information manager (or, rather, knowledge manager since knowledge in the corporate context is defined as information put to good use), I need a definition of information that reflects the context in which I work. Looking at the day-to-day activities, I have to admit that it would be economically unfeasible to indulge in definitions that do not produce a return on investment. Early on in my career, the notion of "information commodity" used to cause me moral dilemmas. These have subdued in severity as time goes by and I find that I am not alone in marketing information.

When Karl Marx wrote his *Das Kapital*, he considered labour to be a commodity. I am sure that he would strongly approve of defining information/knowledge as commodity, had he been around in this "information age." It makes one wonder if he would have considered "information overload" as "surplus value", though.

Let me explain why some of the definitions discussed above don't "work" in my field. Shannon's definition is unsuitable for policy reasons: it would be difficult to convince a potential client that the more information he/she is provided with, the more he/she is charged – and that they are being charged for buying meaningless chaos. And yet even repackaged, digested and customer-tailored information is not THE meaning; it is the meaning the designer has added/represented to the available information. Dembski's definition, with its ruling out of possibilities and probability value comes close to the notion of informed risk management, but his rejection of formalisation and application of laws would be difficult to apply in an organisation. The three attempts at Universal Solutions simply don't work out of context, for reasons I explained above.

That leaves us with the newer definitions of information as thing, or creative construct, being categorised into tacit and explicit knowledge. Information, for it to be economically viable, has to be valid (or perceived as being valid, i.e. correspond to "reality"), meaningful, promote decision making, and re-cyclable. It should be amenable to measurement, storage, channeling, organisation and retrieval. It should also promote its own positive growth. Information has also to be communicable.

This in short, is what I would call my "working" definition of information. It "works" in my particular context, and I make no claims to any cross-discipline universality. However,

in an age where almost everything is rapidly becoming a commodity, time may still prove that this definition is quite feasible for more than just an information broker.

4. References

Agre, Phil. (1997). The End of Information and the Future of Libraries. [Online]. Available WWW: http://www.libr.org/PL/12-13_Agre.html

Dembski, William A. (1998). Intelligent Design as a Theory of Information. [Online]. Available WWW: http://www.arn.org/docs/dembski/wd_idtheory.htm

Dervin, Brenda. (1996). Chaos, Order, and Sense-making: A Proposed Theory for Information Design. [Online]. Available: <http://edfu.lis.uiuc.edu/allerton/95/s5/dervin.draft.html>

Flückiger, Federico. (1996). Towards a Unified Concept of Information: Presentation of a New Approach. [Online]. Available: <http://mypage.bluewin.ch/federico.flueckiger/Uci/Proceed/FIS96PRC.htm>

Hofkirchner, Wolfgang. (1999). Cognitive Sciences In the Perspective of a Unified Theory of Information. [Online]. Available: http://igw.tuwien.ac.at/igw/menschen/hofkirchner/papers/InfoConcept/CognSciences_UnifiedTheory/9944.htm

Kando, Noriko. (1994). Information Concepts Reexamined. [Online]. Available WWW: <http://www.rd.nacsis.ac.jp/~kando/papers/fid94-information.html>

Kirk, Joyce (1999). Information in Organisations: Directions for Information Management. [Online]. Available WWW: <http://www.shef.ac.uk/~is/publications/infres/paper57.html>

Losee, Robert M. (1998). A Discipline Independent Definition of Information. [Online]. Available WWW: <http://www.ils.unc.edu/~losee/b5/book5.html>

Luke, Michael O. (1996). Some Thoughts About the Relationship Between Information and Understanding. [Online] Available WWW: <http://www.asis.org/midyear-96/luke1.html>

Marco, Guy A. (1996). Two False Dogmas in Information Science. *New Library World* 96(1131): 11-14

Rohde, Nancy Freeman. (1986). "Information Needs". *Advances In Librarianship* (14):49-73

Shannon, Claude E. & Weaver, Warren. (1969) *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, Chicago, London.

Sveiby, Karl-Eric (1994). What is Information. [Online] Available WWW: www.sveiby.com.au/Information.html

What is Information? A Discussion from the CRISTAL-ED listserv. (1997). [Online]. Available: <http://lrs.stcloudstate.edu/cim/courses/im577/infodef.html>