What is Information?
An Information Veteran Looks Back*

Robert Fugmann

Robert Fugmann was until 1982 Head of Scientific Department for Documentation at Hoechst AG. He is known "for development of the GREMAS system (Genealogical Retrieval of Magnetic tape Storage), the first truly sophisticated computerized retrieval system, based on a faceted hierarchical fragment code for each part of a chemical molecule, and for development of the TOSAR system (TOpological representation of Synthetic and Analytical system Relations) for the retrieval of reactions and other concepts, including establishment of indexing concepts for nonstructural information and creation of theoretical basis of information systems", for which he was given the Skolnik award by the American Chemistry Society. He was also a foundation member and chairman of the Fachgruppe Chemie-Information (SIG Chemical Information) in the GdCh, the German Chemical Society, Visiting Professor at the School of Library and Information Science, Indiana University in Bloomington, Indiana, and part-time lecturer at the University for Applied Science in Darmstadt and Hannover as well as at the former University of Library Science in Stuttgart. In addition to the Skolnik Award, he was the recipient of the best paper award of the Journal of the American Society for Information Science 1985, and the Ranganathan Award of the Committee Classification Research of the FID

Robert Fugmann was one of the founders of ISKO alongside Ingetraut Dahlberg, in 1989, having previously been vice president of its predecessor, Gesellschaft für Klassifikation, under Dahlberg's presidency. In 1993 he wrote the first of the ISKO series of knowledge organization theoretical texts, S*ubject Analysis and Indexing: Theoretical Foundation and Practical Advice*, and in 1996 articulated the aims of ISKO in the Society's journal *Knowledge Organization*.

Fugmann, Robert. 2022. "What is Information? An Information Veteran Looks Back." *Knowledge Organization* 49(1): 3-5. 6 references. DOI:10.5771/0943-7444-2022-1-3.

1.0 Introduction

"Information"[1] is a concept that is of fundamental importance for everyone in the world. Depending on the field in which the term "information" is used, it may have various meanings and definitions, embedded in a network of links to related concepts, sometimes with a confusing variety of meanings (which are discussed below). The semiotic triangle (Figure 1) described by C.K. Ogden and I.A. Richards (1923, p. 11) provides a good way of studying the ideas involved:

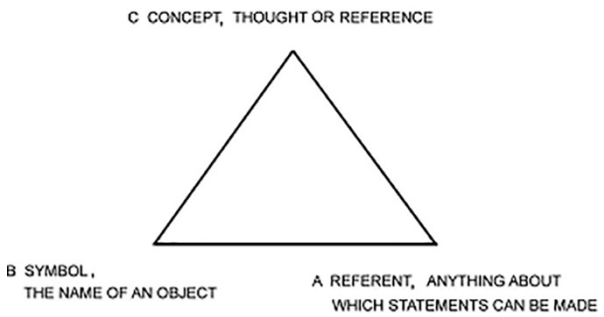

Figure 1. The semiotic triangle.

2.0 Referent (A in Figure 1)

The "referent," in the bottom right corner of the semiotic triangle, is anything about which a statement can be made. Someone who has created or discovered an object and wants to communicate it to others has both a right and a need to name the object. He will find himself in opposition to the *Zeitgeist,* the spirit of the times, or even facing hostility from it. The *Zeitgeist* is instinctively opposed to any innovation, as a matter of principle.

3.0 The naming of an object (B in Figure 1)

We distinguish lexical naming from non-lexical naming of objects, both with typical strengths and weaknesses. Lexical naming proceeds through the assignment of a word to the object of interest. It always requires agreement on such a word among a group of people. It is often conclusively arrived at only with considerable delay.

The non-lexical type of naming is the original and primary form of naming an object. It is freely phrased in the language of the originator or discoverer of an object and it is immediately available. It often takes the form of a definition.

However, there is an infinite number of different ways of describing and defining an object, each of which could have been chosen for the description of the object. Hence, an object of interest may be contained in a store of non-lexical names in a multitude of variations. If complete retrieval of the literature about an object is the goal, each of these infinite possibilities would have to be addressed in an alternative search, since each of the possibilities could have been chosen in the naming of the object. Since this is impossible, any complete search in a store of non-lexical object names is bound to remain incomplete. In addition, it is laborious to compile large numbers of search alternatives.

For a long time, the literature about a new infectious disease existed exclusively in the non-lexical mode. When this became unacceptable, the name "AIDS" (acquired immune deficiency syndrome) was coined, and a working group was formed in the USA with the task of working through the medical literature of the previous ten years to look for relevant papers. Whenever one was found, the new word was assigned to it. The lexicalization of a substantial part of the relevant literature was executed in this way.

The vocabulary of natural language is subject to wild proliferation, mainly driven by people who are not language experts. And misunderstanding of colloquial expressions is widespread.

When misunderstanding is not acceptable (as in the field of medicine, for example) artificial languages are used, in which only unambiguous and precisely expressive words are used for the objects of interest. Special committees are established to maintain such vocabularies, e.g., for Latin in the Vatican, or for the artificial language Esperanto (although it lacks wider acceptance).

4.0 Concept (C in Figure 1)

The concept, as a unit of thought, comprises the essential features of an object ("thought or reference" in the upper corner of the semiotic triangle.).

Various groups of people associate different ideas with one and the same object, depending on what is regarded as being essential in each case. In the case of benzene, for example: for the fire department, the essential aspect is that it burns with a severely smoky flame, forms explosive vapors together with air, and that a benzene fire cannot be extinguished using water, but instead using foam. For a physician, the essential aspect is that benzene is a blood toxin that takes effect via inhalation. For a chemist, the essential aspect is that benzene is a six-member ring molecule of carbon atoms that forms the basis for a large group of chemical compounds.

It would be absurd and presumptuous if the fire department claimed the exclusive validity of its own concept of benzene for all fields of knowledge, without restriction. However, this mistaken attitude is common in the way in which the term "information" is used by telecommunications engineers and computer scientists (see section F below). Widespread confusion is caused when "concept" and "word" are used synonymously.

5.0 Message

Message is an attempt by one living being to add to or correct another living being's store of knowledge.

6.0 Information to be taken note of

Information is a message that is of interest to the recipient. This is the viewpoint of people in information science and of the person in the street. That and similar definitions are widely found in the literature. The essential aspect is the involvement of the receiver of a message and the subjective decision in assessing the message's value. This definition takes into account the highly subjective nature of "information." Each person has different interests, moves in a different environment, and thus regards different things as being interesting and essential.

A great deal of confusion arises from the fact that a distinction is often not made between message and information.[2]

This definition of information also illustrates the impossibility of determining what information a message contains in a programmable way with certainty and without including unnecessary ballast.

7.0 Information to be conveyed

Telecommunications engineers are exclusively interested in problem-free, rapid, and cost-saving ways of *conveying messages.* Starting from this point of view, telecommunications engineers represent a strong contrast to the group mentioned under section 5.0 above, the representatives of information science and practical information technology.

Evidence for this view is provided by the telecommunications engineers Claude E. Shannon and Warren Weaver in their *Mathematical Theory of Communication.* This theory explicitly avoids taking into account the involvement in a "communication" of the receiver of a message, even though without such a receiver, no "communication" at all can actually take place, and even though the word "communication" appears on their title page.

The principle applying in this theory is: "The measure for information is the bit" (binary digit). In this view, two books of the same kind must contain twice as much of this type of "information" as a single book. Seen in this way, the sentence "Three times three is nine" contains more "information" than the statement "3 × 3 = 9," because it contains more bits.

If an object happens to drop onto the space bar of my keyboard during the writing of this text, and if hundreds of additional word spaces with the corresponding number of bits are consequently created, then from that point of view, a correspondingly large increase in "information" has occurred. But in fact what has actually taken place is only an annoying mishap that is unquestionably of no use to anyone. Correcting the resulting error would represent a loss of Shannon–Weaver "information." Any form of computer science that is capable of teaching such things is obviously deeply astray in a crisis of definition.

If you were to search the entire history of science for an approach similar to that of Shannon and Weaver, you would not have any success. It would be like a chemist who regarded the system of chemical elements with its eight periods as representing the "mathematical" basis of chemistry and refused to include the large field of interconnected carbon atoms in his considerations, as seen for example in proteins, fats, and carbohydrates.

Shannon–Weaver's "information" measure of a text using bits does not deserve to be associated with the concept of information. At best it represents only a futile attempt, by hook or by crook, to make something measurable out of a text.[3] You might just as well use a book's weight in kilograms as a yardstick and regard that as a measure for this type of "information."

Subsequent generations of telecommunications engineers continued to cling to this disregard of the involvement of the receiver of a message in the communication process. The attribute "mathematical" that the theory had adopted probably shielded it from criticism.[4]

It is absurd and presumptuous for computer scientists to project the validity of their own specialized concept of "information" onto every other field in which the concept plays a role. Many other attempts to define the concept of information have also been published. In some of them it is quite unclear which groups of people such a definition could be useful to.

8.0 Living beings

The widely accepted definition of "living being" (outside the fields of medicine, biology, or genetics) includes an independent metabolism, an ability to reproduce, and an ability to process stimuli as the essential characteristics of a living being. Biologists may need to rethink this, as living beings known as tardigrades have been identified that are capable of reproduction, but do not require their own metabolism, at least for a time. Even if they are exposed to the cold and radiation of outer space without protection (as happened on the ISS space station), they do not suffer any lasting damage.

Notes

* A German version of this text, "Was ist Information?", is forthcoming in *Information – Wissenschaft & Praxis* 73 (2022) 153–156.

1. It has been suggested that "information" should be regarded as a third fundamental category alongside energy and matter; see also Peter Rechenberg, *Was ist Informatik?* (Munich: Hanser, 2000, p. 302).
2. The loudspeaker message: "The train to Basel will be departing today not from platform 5, but from platform 7" represents information only for people who want to use the train, but not for anybody who is planning stay overnight in the station hotel.
3. Werner Gitt also criticizes this approach in his book *Information* (Bielefeld: CLV, 7th ed. 2020; esp. p. 390 and following) and seeks and proposes alternative solutions.
4. The attribute "mathematical" is widely associated with meaning something grand and unassailable (see also Peter Rechenberg, *Was ist Informatik,* Munich: Hanser, 2000, p. 275).

References

Fugmann, Robert. 1993. *Subject Analysis and Indexing: Theoretical Foundation and Practical Advice*. Frankfurt: Indeks Verlag.

Fugmann, Robert. 1996. "The Goals of ISKO". *Knowledge Organization* 23, no. 3: 173-7.

Gitt, Werner. 2020. *Information* 7th ed. Bielefeld: CLV.

Ogden, C.K. and I.A. Richards. 1923. *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. New York: Harcourt Brace.

Rechenberg, Peter. 2000. *Was ist Informatik,* Munich: Hanser.

Shannon, Claude E. and Warren Weaver. 1949. *The Mathematical Theory of Communication.* Urbana, IL: University of Illinois.