# Defining the "Information" Part of "Information System"

Author(s)
Affiliation(s)
*Email address(es)*

**Abstract.** In order to construct a quantitative model of an information system, an operational definition of "information" is needed. This operational definition needs to mold the concept into a quantitative entity. There are many differend definitions of information already, but they have been formulated with other purposes. In this article the concept of operation is operationally defined with the long-term purpose of building simulation model of an information system where the human actors are considered part of the system. The concept is discussed from the viewpoint of earlier definitions, related concepts, things that happen to it and its various aspects. The curriculum of information is reformulated into quantitative terms, in order to finally arrive in a new definition of information within the current context: Information is an explicitly constructed message, always ultimately constructed by a sentient being. Along with describing the properties assigned to this entity, it is believed that it would be possible to construct a quantitative model of a complete information system.

## Background

In order to measure and compare information systems in an efficient manner, a quantitative model of the information flows within the information system is needed. In order to make a quantitative model of an information system we need an operational definition of information. A quantitative model of an information system, with an operational definition of information, could be used to build a simulation of an information system. The operational definition presented in this article is a part of the way towards building such a simulation.

The concept "information" needs to be properly defined and put into its context. The definition need to be concise and operational (see below), since its purpose is to make information possible to represent quantitatively. The use of the operational definition is, in the extension, to make it possible to quantitatively compare the efficiency of information systems.

In the following I will concentrate on "information" and view "information system" as an extension to it. Further, as a catch 22, I will for now use the term "information" in the discussion without defining it, because the actual definition is the whole purpose of the article. I am aware that this may feel confusing at first.

## An operational definition?

The difference between a definition in the common sense and an operational definition is that the operational definition needs to be concise and exact enough to precisely fit into the purpose for which the operational definition was made. It is not only a description of the thing as such, but also a way to mold a concept into something which possible to use in practice.

An operational definition can here be viewed as a design: We construct and mold concept in order to use it as a tool in the following work. A design is a specification of something new or changed, not a summary of something that exists. Thus it does not disturb me in any horrible manner that the definition does not altogether agree with what earlier authorities within the area has said, although it is still important that the definition relates to those authorities in some way.

## The problem with "information"

Now, there are several available definitions of "information". Why are these not usable within the current context? There are a number of problems with them, and with the everyday use of the concept.

First we must make a hard distinction between data and information. Data is already measurable, and has been since it was operationalized by Shannon (Shannon C, 1949). However, and unfortunately, this definition was made within something called "information science", which has led to much confusion. Shannon never measured information. He measured data. Information implies abstract concepts distributed via the bits and bytes in the data transfer, it is not equal to the data transfer. Shannon cannot be used to measure information transfers other than indirectly.

However, due to the initial labeling, data and information are commonly confused. If you speak about an "information system", the association most people get is a heap of cabling, switches and routers, computer stuff, which is used to transfer data. The efficiency of such a system is commonly measured with terms

like megabit per second, but it does not really say anything about whether the actual information was efficiently transferred.

When constructing an information system, measuring data flows is not sufficient. We must be able to know in advance whether the information flows are likely to be efficient. To know this we must be able to construct a quantitative model of information, not data. To make this quantification we must know exactly what it is we want to represent and the first step down that road is to make an operational definition of the term "information".

### This article

The following article is intended to specify such an operational definition. It should be noted, that the definition is operational within the current context: The simulation of information systems.

It should not be viewed as a generic definition of the term, and it is possible that it does not entirely overlap the commonly used sense of the word.

The approach and method for conducting the operational design definition, is to a) study earlier definitions, b) discuss related concepts, c) discuss aspects of the concept as such, in order to d) conclude a summarized operational definition.

## Information system

Before beginning to define the concept "information", we need to know within what scope the definition is to be applied. As mentioned, the target area is "information systems", but this term needs to be further clarified. In common speech it can mean everything from cabling and routers to an organizational abstraction of how communication channels are organized.

### What is an Information System?

In the words of this article, an information system is all that within an organization contributing to the distribution of information. The information system consists of nodes capable of holding information, channels able to distribute information and actors acting and re-acting upon the information. Thus the actors - the humans within the organization - are *parts* of the information system rather than *users of* the information system. With this definition, there is no such thing as an information system without people.

Since the actors are parts of the system, and since "the system" is our arbitrary grouping of the activities and components we study, the information system can be characterized as a human activity system in the sense of Checkland and Scholes (Checkland P & Scholes J, 1997). The system is an aggregate function of the organization, its people and its technological artifacts. The higher-order object, the

human activity system, the "information system" does not exist as a separate tangible object. It is rather an abstraction, a holon, to which it is convenient to attach properties and behaviors. When further mentioned, the term "information system" shall denote this abstraction as described above.

The short-term purpose of the information system is to distribute the right piece of information to the right actor in the right time. The long-term purpose of the information system is to support the viability of the organization through providing a sound base for decision, information merging and organizational development.

To make certain that no confusion arises upon this subject: let me state explicitly that an information system here is not the same thing as a software system. A software system, such as a customer database, may certainly be a part of the information system, but the information system is an aggregate phenomenon on the organizational level. There is no such thing as an information system without humans. However, it is very possible to have an information system without a single computer and even without electricity.

To summarize, an information system is all that, abstract or not, within an organization that contributes to and shapes the distribution of information: Humans, routines, policies, information nodes, spatial design, location of the coffee machine and the size of the tables in the lunch room.

# Information

The concept "information" is somewhat tricky to define, as it is usually not a tangible object. It is rather an abstract phenomenon which arises between tangible objects.

Thus, rather than trying to immediately define the "information" concept, I will in the following try to define the context of the concept: What relates to it, what its aspects are and what happens to it. When this is done, a collected definition will be presented.

## Survey of literature

The first thing we need to be aware of is that "information" is not one thing. It is an abstraction which differs significantly across different disciplines.

The next few sections will present a brief discussion about the concept from the viewpoint of literature within information science and infology. The fields can be said to be somewhat arbitrarily separated, as the current field can be said to be a conjunction of several.

First we need to recognize that information science is not really about information as such, it is about the communication of information. This has been pointed out in various summaries, such as by Holmström and Koli (Holmström J &

Koli T, 2002). Indeed, one of the most important classical bases for information science is called "A mathematical theory of communication" (Shannon C, 1949). Thus, it might not be entirely fruitful to discuss the information science definition or view of information as such.

However, whether defined or not, a view of information is implicit within information science. In my personal understanding, data and information are seen as equal. For example, one definition of information is "Information is any input into the system that initiates a change of state" (Vickery B & Vickery A, 1987). The authors of this definition do, however, seem to be aware of the problem with the information definition, as they later in the text state that defining information as such would not be fruitful.

Implicitly we could understand information as a data transfer which caused an effect at the receiver. This should largely be a paraphrase of the above stated definition.

Several authors working within the information science paradigm, or close to it, approach the problem in a similar way. Cole, for example, defines information as "that which modifies knowledge structure" (Cole C, 1997a), although he himself sees this definition as not entirely clear and that the distinction between information-as-a-process and information-as-a-thing must be made clear (Cole C, 1997b).

This method of managing the information concept seems to be common within the information science literature. Information is not defined directly, rather it is specified what happens to it, what it causes and how it is treated.

One version of this is trying to define information as a process rather than as an entity. Losee, for example, tries his hands on a discipline-independent definition, and defines information as the following "Information may be understood as the value attached or instantiated to a characteristic or variable returned by a function or produced by a process" (Losee R, 1998).

The border between "information science" and "infology" might not be altogether obvious. In this text I shall treat the latter as the "soft" variant which concentrates on meaning, interpretation and understanding, while the former is the hard part concentrating on measurement, quantity and communication channels.

The classifications of which category authors fall into may be considered somewhat arbitrary. For example, Meadow and Yuan (Meadow C & Yuan W, 1997) discusses the distinction between data and information, from a base of mostly information science literature, but end up into definitions that I feel are infological. They end up in viewing information as partly the function of the recipient; that the interpretation process is a part of the information as such.

One of the fundaments of infology is the infological equation "I = i(D,S,t)", that information is the data (S) and the preknowledge (S) interpreted (i) during the time (t) (Langefors B, 1995). The important consequences of this are that information is never objective. Information only exists within a subject who has

pre-knowledge, and who has performed an interpretation process of a certain piece of data. The objective components of this equation are data and time. Pre-knowledge, interpretation and information are subjective and largely uncontrollable. We can also see a hard distinction between information and data.

This view could be said to largely coincide with, in Buckland's wording, "information-as-knowledge". Buckland states that this is one of three major approaches for viewing information, the other two being information-as-process and information-as-thing (Buckland M, 1991).

Stonier (Stonier T, 1996) discusses information as an entity or concept which is inter-convertible with entropy. In essence, the view is that information equals energy, which performs a work to change the state of a system, which in its turn is an organization of existing entropy. To drive it even further, Stonier claims that "order is considered to be a manifestation of information" (Stonier T, 1996). A closer analysis of Stonier also reveals that he treats "information" in largely the same way that infology treats "data", and what infology would call "information" he calls "message".

## Narrowing down the field

Clearly, "information" can mean many different things, some of them incompatible. To delimit the aim of this work, I shall begin with a root definition (Checkland P & Scholes J, 1997). The system in focus (the system we will study through the use of the simulation, not the simulation itself) is "an actor-owned and manned system to distribute and treat information, through the use of technological and social information channels, in keeping with the overall organizational goals, in order to support the viability of the organization through keeping all its parts up to date with necessary information." (This is a paraphrase of a root definition by Checkland & Scholes 1997).

## Related concepts

In the discussion about information, there are a number of other concepts which can be considered close or related, but which themselves are not the same thing as information. In order to delimit what information in itself is, these other concepts need to be made explicit too.

The following discussion will be concentrated on how the related concepts can be formulated in relation to an operational definition of information. Thus, the approach is to operationalize these concepts too.

### Data

As previously mentioned, one key concept is that of data. It is important since it is often confused with information. However, while data is something tangible,

information is not, at least not yet. "Information" implies meaning, while "data" only implies the *possibility* of meaning (Meadow C, 1996). In the following I shall view "data" as synonymous with the carrier in an information channel (see below).

## Concepts related to human understanding

One major thing which differentiates information from data is the concept that information contains some form of meaning. Meaning is implicitly something humanly attributed: It exists only in the relation between data and a conscient, interpreting entity. For all practical reasons we can for now equal this entity with a human being. While it is conceivable to manage some kind of AI or communicating with animals, this is not the common use of an information system today, and it is therefore for now disregarded.

From Langefors (Langefors B, 1995) we acquire two key concepts here: interpretation and pre-understanding. Pre-understanding can be expanded into knowledge and experience. For completeness we can also include the terms wisdom and understanding here.

In the current context, the problem with all these concepts is that they are highly relative, or at least subjective. They are a function of a human being, and not representable in any easy manner. This is a problem with the current context, not with Langefors' definition: It is correctly observed that information is relative and subjective. This does not help us right now though. It is for us to find a work-around for this.

Thus, to work around the problem, the concepts need to be operationalized. If they can be formulated in a quantifiable manner, then quantifying the thing they relate too will be easier too.

Thus, let us begin with "knowledge" and "experience". These can be readily operationalized if we presuppose that there is a quantitative and objective definition of information. Essentially they are synonymous: knowledge and experience is the amount of information that an individual has successfully managed to receive. Knowledge and experience within a certain topic is the amount of information received concerning that topic.

Of course, we have to take into account the situation where a person is exposed to information, but does not manage to receive it. Or to use more general terms, where someone got the information but did not understand it.

This is where the term "interpretation" comes into focus. To make things easier, I will presuppose that all information in an information system is, in the end, theoretically possible to understand. It might require other information in order to be understandable, and it might take longer or shorter time to understand, but theoretically all individuals in the information system can, with some effort, understand all information available.

With this delimitation in mind, "interpretation" becomes the process in which information is integrated with the already acquired knowledge and experience. The

result of an interpretation process is either that the amount of known information is expanded, or that the information in focus is rejected because there was not enough knowledge and experience available to be able to accept it. With this delimitation, the term "understanding" can be left out of the equation, since all information is understood once it has passed the interpretation process.

The existing knowledge and wisdom forms a frame within which new information is interpreted. This frame can be said to be synonymous with "pre-understanding". Thus, pre-understanding is operationalized to denote the availability of knowledge and experience. Or, in consequence, pre-understanding is the already acquired relevant information.

This leaves us with the term "wisdom". Essentially, it is possible to continue the discussion without taking this term into account, but for completeness, let's define it. And operational definition of "wisdom" in this context is that it equals the speed and success factor of the interpretation process. A person is wise if he has much knowledge and speedily can acquire new information.

Essentially, the above leaves us with the opportunity to make a collected reformulation of Langefors' infological equation. This is an interesting opportunity which, as we will see, will end up being useful.

## Concepts related to communication

As the term information system implies communication, it is relevant to think that communication has an impact on the definition of information in this context. Descriptions of communication usually involve the actors and the path between the actors. Leaving the actors themselves aside for now (they will be discussed below in the section on human behavior), we can list a number of important concepts related to the path between two actors in a communication.

First we have the path itself, the channel. This is the complete path, with all transformations, from actor A (sender) to actor B (receiver). It does not include the actors themselves, and does thus not include the interpretation process mentioned above. The "channel" starts when information is transmitted and ends right before it is received.

The "channel" consists of at least one medium, at least one carrier and at least one syntax. While information can be conceived as a somewhat abstract concept, it does not exist in an abstract void. It relies on a physical environment to be transmitted, and during the transmission it has a form and a physical representation. The environment through which the information is transmitted is the medium, and through that medium the information is transported using a carrier. For example, if two people are shouting to each other, the medium is the air and the carrier is the sound waves. If sending an email, the medium is computer cabling and the carrier the electronic signals making up the mail.

The thing that separates information from line noise (apart from the meaning and the content which we disregard for now), is the "syntax". The syntax is the

structure of the transmission, the protocol agreed upon by both sides of the channel. Without the syntax, the transmission can not be properly received at the end of the channel.

To summarize, "communication" is the transmission of information in a structured way through a medium by a carrier. The term "communication" does here not include the processes required to prepare the information for transmission at the start of the channel, nor does it include the interpretation process at the end of the channel. These delimitations are chosen in order make the "communication" as such objective. The subjective activities related to the communication are discussed later.

## Concepts related to decision-making and human behavior

As mentioned before, with the definition of information system used in this article, there is no such thing as an information system without humans. As an implied consequence we could also question whether there is such a thing as information without humans.

As a practical postulate, I will state here that there is not. All information presupposes at least one thinking and interpreting being somewhere. However, and as a distinction against other definitions, the information does not absolutely require two ends of a communication channel.

Having accepted that the definition of information requires at least one human, we need to specify how. We might need to take into account things such as information behavior, how people behave in relation to information.

The first aspect to point out here is how the phenomenon of information arises. As a continuation of the previously stated practical postulate, I will claim that information is exclusively caused by human beings, although possibly through their behavior and interaction with their environment.

Completely new information can arise through observation. However, usually new information is dependent on existing information. Partially new information will arise through the processing of other information, in human activities such as a guess, through inference or through deduction.

The experiences of a human form how that human processes further information. The human has formed schemata and scripts for information processing.

Finally, there is an output in behavior (of which non-action is an example). The human has made a decision, usually after having come to a conclusion.

These concepts, related to construction, processing and output of information, need to be made operational to make an operational definition of information possible.

In a model of an information system, the human entity is a part which needs to be modeled. However, the detailedness of model of this entity only need to fulfill the requirements for the information system model, it does not have to be a

detailed model of human behavior. As noted by Schmidt (Schmidt B, 2000) a model does not equal a replica. Here, only a basic representation is required.

So, as for the representation of the production of information. As we cannot model the detailed contents of information, we must represent it through a reduction, such as an amount of categorized information. Since the information is roughly represented, its arisal should likely be equally roughly represented. Thus, in a simulation, completely new information can arise on a statistical basis in the human entities. An observation, a guess or an inference is thus the event when new information arises in a human entity based on a statistical model. To discuss exactly how this statistical model should look is beyond the scope of this discussion.

For the processing of information we have the terms "schemata" and "script", which are essentially terms fetched from cognitive psychology. The operationalization of both these terms are the allowed combinations of already acquired and incoming information. The complete reduction is a list of which categories of information is allowed to be combined with which other categories to produce new categories of information. These combination and categories are implementation-dependent.

Finally, decision and conclusion can be operationalized to denote output (or non-output) of information. A visible decision or a conclusion is when two chunks of information have been combined and result in the human entity transmitting the resulting information. An invisible ditto is then the combination has happened, but the human does not transmit the result, and simply keep it to himself for later.

For an implementation of a simulation of human information behavior there of course remains much to be said. A model would need to include operationalizations of how humans work, look for information and communicate. However, these models are beyond the scope of this discussion.

## Other important concepts

There are also a number of random other concepts which relate closely to the concept of information. Especially of note in this discussion are terms such as fact and context for content and interpretation of information, and document and message for its transmission.

To begin with "fact", it is essentially left without meaning in an operationalization. Information is information, when it comes to quantitatively representing information flows. Whether the information is true or not is not and does not have to be measurable. Thus, concepts such as disinformation, too, lose their meanings. All information according to the coming operationalized definition is informative whether it is true or not.

In a simulation, some things will be too fuzzy to take into account. One of those things is the context in which information is received. In reality, the context has a heavy influence on interpretation and evaluation of received information.

However, since we do not take into account the truthfulness of information, and since it is infeasible to cover the detailed contents of an information transmission, the notion of context is also dropped: the prerequisites for the influence of context of the interpretation process does not exist, and thus the context cannot be taken into account.

Tangible concepts such as "message" and "document" become all the more interesting though. I shall here view these terms as more or less synonymous. A document is merely a formalized message. All documents are thus messages, although not all messages are documents. The notion of message fulfills many of the criteria of quantification we want, and as we will see below, it can be useful for the final operational definition of information.

### Excluded concepts

This discussion has necessarily excluded several concepts. The obvious synonyms have been left out, but also things that might be considered key concepts by some.

Psychologists of various brands might be put off by the fact that I have left out concepts such as action, reaction, collective consciousness and subconcsciousness. The first two were left outside, since I felt they were already covered by the section on human behavior. The last two were left outside since I did not think them relevant for the discussion: The terms are fuzzy from the start, and would not contribute to the operationalization of the information concept.

Infologists, information scientists, computer scientists and AI researchers might react against the exclusion of terms such as data type, scale, denotation, proof, capta and elementary message. Scale and data type are left outside because they are too lowlevel to be relevant for the discussion. Denotation and proof are left outside, because they concern the contents of information rather than the information itself. Capta and elementary message are theoretical constructs concerning the interaction between data and interpretation and are too abstract to be considered important.

## Related activities

Having discussed what other entities relate to information, we can now discuss what happens to our elusive concept itself. Some of these activities should not be very controversially defined, but some may seem a bit strange at first. As before, the goal with the definitions here is to make the information concept as explicit and quantifiable as possible.

### Transmission and storage

Transmission is the entire chain from (but not including) the mind of individual A, via encoding to a format suitable for sending, through a communication channel on a carrier, to decoding by individual B to (but not including) the mind of individual

B. On the route, the information might also be stored for longer or shorted durations.

The communication channel and the carrier have been discussed earlier, and they are thus left outside of the discussion here.

Decoding is synonymous with interpretation. It is the process in which received information is integrated with the known information. This can be equaled with time. Encoding could be said to be this process but backwards, it is the process of moving known information to a carrier on a communication channel. This, too, can be equaled with time, although not necessarily the same time as for the decoding.

Information can be stored for longer or shorter durations. A short duration is when a technological artifact holds the information during the process of transmitting it along a communication channel. A longer duration is what is usually understood by the term storage: It can be put in an archive. This, too, is done in some form of technological artifact (which may, of course, be an analogous traditional archive).

## Production and destruction

When building a model of information flows, my experience has been that the trickiest part to specify is how and when information arises. Once information exists, making it flow through a model is usually not all that difficult.

To address this we must clearly specify the production and destruction of information. Where does it arise, and where does it disappear?

In the following, I will apart from the term "production" also use aggregation as a term related to production. There is a small but important difference between the terms.

As mentioned in the discussion about human information behavior, new information can arise on a statistical basis in a human being. This is one source of information. The new information can be completely new (produced) in which case it has spontaneously arised (representatively as a result of observation, guess or inference), or it can be a combination of earlier acquired information (an aggregation). These are the only internal sources of information within an organization.

However, information systems does not exist hanging in a void. They have input and output, as all open systems. Thus, an additional source of information is external input. This represents information coming into the organization from the outside world, through for example phone calls or emails.

Concerning the destruction of information, it is never explicitly destroyed. Information exists as long as it has a carrier. If a human has received information, then the information is known as long as the human remains in the information system. Information also remains in the system if it is carried by another carrier, such as technological artifact or an archive. From these, the information can be explicitly expunged on the basis of dating.

## Aspects of information

We have now almost surrounded the information concept by defining things close to it, and what happens with it. Before stating clearly what information actually is, I will address some of its aspects.

The purpose with this is mainly to specify what can be specified in the collected definition, and why. The major outcome of this section is a list of things that cannot be considered in order for there to be an operational definition.

### Aspects concerning content

The thing that makes information difficult to nail down is its content, and the terms relating to it, such as relevance, integrity, value, reliability and authenticity.

A model of an information system needs to specify information content somehow. A flow model need to take into account which information should end up where. This causes a severe conflict with the need to make the same model simple and possible to overview.

On one extreme of the simplicity-relevance axis here we have the complete copy: We know the actual contents of the information, and we specify that content for each information transmission. This makes the model huge and cumbersome and, say, a computerized implementation would be difficult to make. Further, knowing all the information in the information system in advance is likely not possible.

On the other extreme, we have the complete reduction of information into a meaningless token. We know nothing about the contents of the information, and the flow model could look exactly the same as if we shuffled coins or water around. This makes the model simple, but hardly relevant. Some notion of content is necessary in order to build a model of an information system, since the paths of information distribution is highly dependent on the contents of the information.

Now, a compromise would be ending up somewhere around the middle of the axis. We could specify a representation or classification of the information without specifying the contents fully.

### Aspects concerning quantification

The whole purpose of the information definition is to make information quantifiable. Thus, terms such as quantity, mass and size need to be specified.

In the end the information entity must just that: an entity. It should therefore be possible to count, and therefore be possible to come in a quantity. In common language, information is not countable. You can have "a piece of information", but not (at least not grammatically correct) "two informations". However, with information as an entity, you can have two informations. "One information" is one instance of the entity we call information. We must, however, recognize that not all information is equal in size. The operational definition of size is here equaled

with time for interpretation and/or for encoding. A large information takes time to accept. A small information is quickly accepted. Thus, the measurement of the size of an information is specified in average number of seconds it takes for a recipient to accept it.

To make a difference between interpretation and transmission, I will also specify the "mass" of an information. This is here defined as the time in seconds it takes to transmit it via a communication channel. In implication, it can be viewed as the amount of data required for the carrier in the communication channel.

## Operationalization and Definition

Now with the entire curriculum described, it is finally possible to deal with the concept as such. First a short discussion about the reformulation of the infological equation, then the definition as such.

### Reducing and operationalizing the infological equation

As mentioned before, one classical base for defining information is Langefors' Infological Equation (Langefors B, 1995), which says that "I = i (D, S, t)", that information is the result of the interpretation process "i" operating on data "D" and pre-knowledge "S" over time "t".

As seen earlier, this definition of information is not suitable as an operational definition, since it explicitly states that information is always subjective, and thus not in practice quantifiable. However, with some reformulation, it is my view that it could be molded into a usable form.

In a simulation, we would need to pre-suppose a model S for the agents, thus making S implicit in our definition. In consequence with the previously stated definition of pre-knowledge, it is simply the collection of previously known information. Since it is a function of information, it cannot be a pre-requisite of information. Thus we should move S outside the equation for now: it is a question of implementation later, and does not affect the actual definition of information.

The interpretation process "i" is globally defined as the time it takes to accept information. This is according to our previously stated definition that presumes that all information in an information system is ultimately understandable by all parts of the information system. The interpretation process is thus something that happens to information, not a part of it. As with "S", "i" should thus be removed from the equation for now, it is a question of later implementation. Further since we already equaled "i" with time, we can also remove "t" from the equation.

That leaves us with the data transmission, which in this case is a meaningful chunk of "something". I am reluctant to call it "data" in this case, since it is the theoretical contents of a carrier transported on an information channel. A better term for D here would here be "message" M. This leaves us with a reduced "I = M", information equals "message", a self-contained entity containing meaning,

meta-data and other relevant properties. The term "message" implies intention, and thus as a further delimitation, a message is always explicitly constructed by a sentient being. The construction might have been done through using a technological tool, but ultimately, a "message" is always constructed by sentience.

The reduced parts, time (interpretation process) and pre-knowledge remain hanging. These are re-introduced as time and behavior further down.

### The new definition

This finally makes it possible to define the now tangible entity "information": ***Information is an explicitly constructed message, always ultimately constructed by a sentient being.***

This is the base definition which enables an understanding of the thing as such. However, to fully operationalize the concept, we further need to specify its properties and qualities, and specify what it is and is not. This will be done in the following sections.

## Properties of the newly-defined entity

While the entity as such is thus defined, we will not be helped until its properties are described. While these properties do not necessarily take part of the operational definition, they are part of the operationalization.

These properties do not necessarily have to be general across implementations where the operational definition is used. The use of them depends on what a model of an information system is intended to be used for. This, said, the following are properties which will likely be relevant for most implementations.

### Automatically assigned properties

The automatically assigned properties of "an information", are those things that describes it curriculum without touching its contents. Examples useful for tracing the paths of information through an information system are things like creation date, author, and intended recipient. As noted, since information according to the definition always arises ultimately as caused by an actor, these properties should be possible to automatically assign in a model.

### Behavior

Most aspects of information are, as seen above, possible to reduce to time. Thus most of the behavior of information is reducible to time. Thus we can specify the size and mass of information as the times it takes to interpret it and to send it respectively. Apart from this we also need to specify specific requirements. These are things such as format requirements, such as whether this current information is not possible or feasible to transform to certain channel formats. Further we have

interpretation requirements, which operationally is a list of other information which is necessary to have acquired before this current information can be integrated.

Category, subject and instances of information

As mentioned, we cannot fully model the contents of information: It would void the purpose of the model. However, we can represent the information contents through marking information with category and subject. These are arbitrary labels, which are case-scenario specific. We can then likely find a few large categories (such as "marketing", "sales" and "environment") within each we can find a number of subjects. These subjects do not necessarily have to be named. Each instance of information has a category and a subject, but several instances can have the same category and subject while still being unique.

The category and subject properties are the base for the routing of information: The actors produce information of a certain classification, and look for information of a certain classification.

# Summary and conclusions

Information is an explicitly constructed message, always ultimately constructed by a sentient being. This is the basis for the operational definition which is the result of this discussion. With the discussion around the definition as a complement, it should be possible to use this definition for building a model of an information system, in which the information flows are quantifiable.

Implicitly, the discussion above has formulated both an ontology and an epistemology surrounding the concept of "information". The ontology has, consciously, been focused on a reduction of reality in order to mold the concept into something tangible. The epistemology has been focused on quantification and measurability, the re-formulation of qualitative concepts into distinct quantitative entities.

To iterate, the definition is not to be considered a generally applicable truth. It is a construction, a design, of a tool. It could look differently, and for another context it probably *must* look different.

The definition and the curriculum is a theoretical base for a model of an information system, it is not the model as such. An actual functional model needs to implement further operationalizations which are case-specific.

# References

Buckland, M (1991): *Information as Thing*. Journal of the american society for information science vol 42 no 5

Checkland, P. Scholes, J (1997): *Soft Systems Methodology in Action*. Wiley, USA

Cole, C (1997a): *Calculating the information content of an information process for a domain expert using Shannon's mathematical theory of communication: A preliminary analysis.* Information processing & management vol 33 no 6

Cole, C (1997b): *Information as a process: the difference between corroborating evidence and "information" in humanistic research domains*. Information processing & management vol 33 no 1

Holmström, J. Koli, T (2002): *Making the Concept of Information Operational*. Department of Information Technology and Media, Mid Sweden University, Sweden. Master thesis in informatics, http://www.itm.mh.se/~vivasp/term8/Exarbeten/02D-008.pdf

Langefors, B (1995): *Essays on Infology*. Studentlitteratur, Sweden

Losee, R (1998): *A discipline independent definition of information*. Journal of the american society for information science vol 48 no 3

Meadow, C (1996): *A Proposed Method of Measuring the Utility of Individual Information Retrieval Tools.* Canadian Journal of Information and Library Science. Vol 21 no 1.

Meadow, C. Yuan, W (1997): *Measuring the impact of information: Defining the concepts.* Information processing & management vol 33 no 6

Schmidt, B (2000): *The Modelling of Human Behaviour*. SCS-Europe, Belgium

Shannon, C (1949): *A mathematical theory of communication*. University of Illinois Press. USA.

Stonier, T (1996): *Information as a basic property of the universe*. BioSystems vol 38. Elsevier Science Ltd

Vickery, B. Vickery, A (1994): *Information science in theory and practise*. Bowker-saur, UK