

# WHAT IS INFORMATION ANYWAY?

**Peter Van Garderen**

*<https://vangarderen.net>  
ORCID 0000-0002-4611-0315*

**Many definitions for archival materials depend on an understanding of the concept of information but do not explain it. This working paper attempts to provide a definition. It concludes that an information object is an entity that contains the content of a message and has the required structure and context to allow that message to be decoded and understood.**

I published a working paper to establish an access-based definition for the concept of archival materials. The purpose is to better define the domain of archives management for software development. [1]

Like most definitions in the archives and records management literature, it leans heavily on the concept of information. However, this term is seldom defined further. It is usually expected that the author and reader share the same understanding of what is assumed to be a universal concept. Therefore, it is necessary to ask "what is information?"

A literal, etymological definition of information is to give form to something. In the modern usage of the word this means to give form to a message by moulding it into a shape or pattern that can be communicated. Since information is such a universal concept there are many other valid definitions for information originating from a variety of disciplines that make heavy use of the concept, e.g. physics, genetics, neurology, cybernetics, computer science, economics, communications, knowledge management, media studies, library science, archival science, etc..

## ***Measurement versus meaning***

Most of the existing definitions for information can be grouped roughly into

quantitative and qualitative categories. The qualitative definitions are focused on the criteria which add meaning to the message that is communicated as information. The quantitative definitions originate from physics. They are focused on measuring the quantity of information units or the strength of its transmission. This approach reduces information to a binary set of symbols or signals (not necessarily electronic). The quantity or frequency of the total number of possible messages that these symbols can make are calculated against the backdrop of random or intentional noise in the transmission channel (i.e. the signal to noise ratio). This is the foundation of classical information theory which was first introduced by Claude Shannon in 1948 and continues to be used in the communications industry to invent and improve the transmission of information over copper wires, through the air, through fibre-optic cable, etc..

More recently, quantum physics has introduced the concept of qubits, essentially quantum bits of information, as being the fundamental building block of the knowable world instead of matter, atoms or the quantum field. [2] Of course, information is all around us, all the time. Indeed, as Marcia Bates notes, "the only thing in the universe that does not contain information is total entropy." [3] As human beings we are exposed to a daily barrage of information which are sent to our senses as photons of light, scents, sound waves, vibrations, pressure, flavours, etc.. Our eyes, ears, noses, tongues, and skin send these messages to the brain for filtering, processing and recording as memories of experiences, facts and ideas. [4]

Classical information theory and physics is interested specifically in the symbols and

signals of these messages and not their intended or interpreted meaning. This is in contrast to the qualitative definitions of information found in the humanistic information sciences (including archival science) that study the processing and use of information. These are very much concerned with the meaning and understanding of the message that the information communicates.

### ***Meeting somewhere in the middle***

In his recent book, *Information: The New Language of Science*, the physicist Hans Christian von Baeyer has called for bringing the quantitative and qualitative definitions of information closer together. [5] Interestingly, it was the visionary archivist Hugh Taylor who anticipated something similar ten years ago when he wrote:

As archivists, we are coming to understand more fully the meaning of documentary relationships in all their richness. If we take the universe and its Creator, or the scientific concept of the 'Big Bang' theory as the beginning of cosmic evolution, then cosmogenesis, the creation moment, becomes the ultimate context of all matter as it moves down through the galaxies, nebula, planets, and stars to life in all its forms on our own planet; all creation is connected in various ways in a marvelous spatial balance. Out of the formation of new entities has emerged information resulting in communication and memory. [6]

As an archivist, Taylor starts with documentary context and works his way back, all the way back, in this case, to the ultimate context of cosmogenesis. As a physicist, von Baeyer starts with the atom and works his way up:

All empirical evidence in science is collected through the mediation of the senses. We learn about atoms by peering through scanning tunneling microscopes, by translating their random motion, revealed to the touch as warmth, into thermometer readings, by converting invisible nuclear events into the audible clicks of Geiger counters...[Information] is the strange, compressible stuff that flows out of a tangible object, be it an atom, a DNA molecule, a book or a piano, and, after a complex series of transformations involving the senses, lodges in the conscious brain. Information mediates between the material and the abstract, between the real and the ideal...Knowledge of

the world is information; and since information is naturally quantized into bits, the world also appears quantized. If it didn't we wouldn't be able to understand it. [7]

Somewhere at the intersection of Taylor and Von Baeyer's grand statements, the same concept emerges; information is at the core of all human existence and at the core of our personal and collective perception of that reality. We use information to give shape, understanding and meaning to our experiences and it functions to communicate and recall those experiences. In very elemental terms this is why we preserve information and retrieve it using tools such as archives access systems. As Eric Ketelaar notes:

Archiving - all the activities from creation and management to use of records and archives - has always been directed towards transmitting human activity and experience through time and, secondly, through space... Archives are time machines enabling man to carry his thoughts, experiences and achievements through time. [8]

### ***The importance of context and knowledge***

To be sure, in order for information to be useful as memory aids or, if you will, as a time machine, it must be more than the physicists' stripped down signals or qubits. The qualitative definitions of information require that the recipient of the signals and patterns is able to decode the message and understand what is being communicated. [9] This requires that the recipient of the message has adequate contextual and background knowledge to decode the signals, symbols or patterns and then to understand the message that these represent. Without the necessary context and knowledge, the message is simply a string of data. [10] That is why knowledge management experts typically refer to information as "data that makes a difference." [11]

In order to decode the data in a message, the recipient might require knowledge of a spoken or written language and the sounds or symbols (e.g. alphabet) that it uses. [12] It could also require specialist knowledge of a specific code and its signals, such as the Morse code used by telegraph operators or the international flag code that is used by sailors. It could require specific technical equipment; such as a radio that converts the

high frequency electromagnetic waves sent by radio stations into the low frequency sounds that the human ear can detect.

Once the message is decoded, a certain level of knowledge is required to process, interpret and make sense of it. Only then is the data within the message considered to be information. Prusak and Davenport define knowledge as:

a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences...It originates and is applied in the mind of knowers. [13]

The Open Archival Information System (OAIS) standard uses the concept of a 'knowledge base' to determine if a person or system has the requisite contextual and background information to understand received information. [14] If a message recipient does not already have the requisite knowledge to understand the information then it needs to be included as part of the communication process (e.g. a codebook, a dictionary, a manual, an explanatory note, etc.). Otherwise, the message will never be more than an insignificant string of data. Of course, one of the primary roles of the archival descriptions that are stored and searched in archives access systems is to provide adequate contextual information (e.g. dates of creation) and background knowledge (e.g. biographical sketch of the creators) so that archival materials can be understood and interpreted.

### ***Let's objectify it***

The key characteristic of archival materials is that they preserve information for future use. This means that the message which transmits or communicates the information must be recorded so that at some point in the future it may be retrieved and re-communicated or re-experienced. This requires that the message transmission is captured and converted into an object that can be carried forward through space and time. An information object is an entity that contains the content of a message and has the required structure and context to allow that message to be decoded and understood.

## REFERENCES

- [1] Van Garderen, Peter. "Archival materials: a practical definition" (January 22, 2007) <http://doi.org/10.5281/zenodo.4998947>
- [2] The qubit was first introduced by Anton Zeilinger in 1999 [see "A Foundational Principle for Quantum Mechanics." *Foundations of Physics* (29:4) (Kluwer, 1999)]. For a very readable introduction to this concept see Von Baeyer, Hans Christian. "[In the Beginning Was the Bit](#)" *New Scientist* (2278: February 17, 2001) [last accessed on January 29, 2007].
- [3] Bates, Marcia "[Fundamental Forms of Information](#)" *Journal of the American Society for Information and Technology* [in press] (2005) [last accessed on January 29, 2007].
- [4] Any memory that persists longer than thirty seconds is generally thought to be stored into our long-term memory through a process called long-term potentiation which uses proteins to increase the number and/or size of synaptic connections between neurons in the brain. Although alternate theories exist on how memories are actually stored (e.g. altered genes that encode memories onto proteins using a process similar to DNA blueprints), most neuroscientists agree that most of a person's lifetime experiences are recorded permanently in the brain. [Furlow, Bryant. "[You Must Remember This](#)" *New Scientist* (2308: 15 September 2001) [last accessed on January 29, 2007] ].
- [5] Von Baeyer, Hans Christian. [Information: The New Language of Science](#) (Phoenix, 2003), p.27. Marcia Bates has made a similar call to her colleagues in the humanistic information sciences, see Bates, Marcia "[Information and knowledge: an evolutionary framework for information science](#)" *Information Research* (10:5) (Wilson, 2005) [last accessed on January 29, 2007].
- [6] Taylor, Hugh. "The Archivist, the Letter, and the Spirit" *Archivaria* (43) (Association of Canadian Archivists, 1997), pp.5-6.
- [7] Von Baeyer, Hans Christian. *Information: The New Language of Science* (Phoenix, 2003). pp.15, 17, and 229.
- [8] Ketelaar, Eric. "The Archive as Time Machine" *Proceedings of the DLM-Forum*

2002. (INSAR European Archives News, 2002), pp.578 and 580.

- [9] Marcia Bates distinguishes between quantitative and qualitative definitions of information as 'information 1' and 'information 2' where the former is defined as the "pattern of organization of matter and energy" and the latter as "some pattern of organization and matter and energy given meaning by a living being (or its constituent parts)." Bates, Marcia "[Fundamental Forms of Information](#)" Journal of the American Society for Information and Technology [in press] (2005) [last accessed on January 29, 2007].
- [10] "Data is a set of discrete, objective facts about events...there is no inherent meaning in data" Davenport, Thomas and Prusak, Laurence. Working Knowledge: How Organizations Manage What They Know. (Harvard Business School Press, 2000), pp.2-3.
- [11] Davenport and Prusak, Working Knowledge, p.3
- [12] The Open Archival Information System (OAIS) standard uses the concept of 'representation information' to define, generically, the mechanisms that are used in any given archival context to decode data into useful information. [International Organization for Standardization. ISO 14721 -- Open Archival Information System - Reference Model (2003), p 2-3]
- [13] Davenport and Prusak, Working Knowledge, p.5.
- [14] Open Archival Information System (OAIS), p. 2-3.