Outline of a theoretical framework of Architecture of Information: a School of Brasilia proposal

Mamede Lima–Marques

1 Introduction

The motivation for this paper came about with the perception that, from the beginning of 2000, there has been a need for theoretical fundaments for the Architecture of Information (AI) [Haverty, 2002, Dillon, 2002, Robins, 2002, Dale, 2002]. The challenge to be faced is the possibility of proposing a theoretical framework of Architecture of Information with the aim of lessening the conceptual gap currently presented.

The situation becomes more complex when it is noticed that all sciences and all institutions in our society deal with information. The human being is immersed in information. As a consequence a fundamental question arises: what is the organization of information underlying human understanding and interest?

The notion of information has become a crucial topic in several emerging scientific discipline [Doucette *et al.*, 2007] and it shall lead a clear epistemology for the Science of Information. In line with various authors [Bates, 2005, Hofkirchner, 1999], the Science of Information is considered a wide field of human knowledge and so is its object of study: information. As a consequence, our vision aligns with this approach in the sense of creating a "Science of Information" (SI) [Doucette *et al.*, 2007], and thus an "Architecture of Information".

Therefore it aligns itself with the initiative of 3^{rd} International Conference on the Foundations of Information Science, Paris, July 2005, where a new expanded field was proposed: 'Science of Information'. This not to be confused with the older term 'Information Science', which sometimes is understood as advanced "library science", rather it is to take into consideration a newer and larger perspective encompassing many academic disciplines and new fields of interest.

2 On the basis

The search for adequate epistemological elements for engineering a solid foundation for the scientific explanations within SI and AI is crucial. Positioning oneself as regards matters of the core elements – data, information and knowledge – is very complex not only due to the high level of polysemy their usage comprises but also due to what is described by [Floridi, 2004] and revised by [Crnkovic and Hofkirchner, 2011] as unsolved problems. Such positioning is, however, fundamental for building coherent scientific theories and developing advanced applications.

A position about each of these fundamental elements is presented in what follows. These positions, far from solving unsolved problems, define epistemological frameworks from which is hoped to advance in the discussions.

2.1 On knowledge

The nature of knowledge is closely connected to the idea of AI. Phenomenology is suggested as the theoretical framework for understanding the phenomenon of knowledge. Therefore, the ideas of its most influential thinkers will be addressed, what will serve as basis for the discussions presented in this paper.

Phenomenology as a philosophy postulates that "all doing is, in essence, significant" began with the ideas of German philosopher Franz Clemens Brentano (1838-1917). Brentano defined two classes of phenomena: physical and mental. According to him, research about physical phenomena could be conducted through traditional positivist methods insofar as these phenomena become direct objects of sense perception; on the other hand, the positivist method would not the applied for mental phenomena due to the primary characteristic of such phenomena: 'intentionality', defined hereinafter [Hirschheim, 1985].

Edmund Husserl (1859-1938), philosopher of Israeli ascent born in Moravia (region in the Check Republic) and a follower of Brentano, is considered to be the founder of the phenomenological movement. His Phenomenology consists in a philosophical method proposing the description of lived experience from consciousness, the manifestations of which are purged from its real or empirical characteristics and considered against the plan of essential generality.

Phenomenon, from Husserl's point of view [Husserl, 1961, Husserl, 1963], does not mean "the simple appearance that opposes to the truth of a being or number", as it is in Plato and Kant, it is apparition rather than the appearance; it is the appearance of the object accessible through consciousness; the full manifestation of sense. And Philosophy needs find clarification of this sense.

A core point in Husserl's phenomenological concept is the 'intentional' character of consciousness, according to which 'consciousness is always consciousness about something'. Therefore, 'intentionality' consists of consciousness tending to an object and giving it a meaning. Kant's Phenomenology describes consciousness and experience, but abstains from considerations regarding its intentional content [Smith, 2011].

For Husserl, Phenomenology is fundamentally interested in the structure of various forms of experience: perception, thought, memory, imagination, emotion, will and volition to bodily awareness, embodied action, and social activity, including linguistic activity. The structure of these forms of experience comprises various intentionalities. This view, influenced by Brentano, determines the direction of experience for the objects in the world, meaning that viewing an object as data, as imaginary or as past is possible [Smith, 2011].

The phenomenon of knowledge, for Husserl, presents itself in its fundamental aspects. As reported by [Hessen, 1978], in knowledge the 'subject' and the 'object' face one another. Knowledge appears as a relation between these two elements, which remain eternally separate from each other. The subject-object dualism pertains to the essence of knowledge. The relation between these two elements is balanced – as a correlation, not equivalence – meaning that subject is subject and object is object. Both are only for one another. The function of the subject is to apprehend the object in terms of their properties, and function of the object is be apprehended by the subject. The subject is altered according to knowledge. In the subject arises an 'image' of the object, i.e. a set of object properties.

Therefore, knowledge is an image, a set of properties of the object apprehended by the subject. Knowledge is different from subject and from object. Knowledge appears as a third element that through correlation connects with those two elements thus forming a trinity.

2.2 On the Data

In the context of the theory of knowledge, supported by the phenomenology of Husserl, knowledge plays an important role in the sense of convergence of the intentionality of the subject, ready to deliver a unique experience: the set of object properties.

Intentionality refers to the notion that consciousness is always consciousness of something. Consciousness occurs as the simultaneity of a conscious act and its object. Intentionality is often summed up as 'aboutness'. Whether this something that consciousness is about is in direct perception or in fantasy is not consequential to the concept of intentionality itself; whatever consciousness is directed at, i.e. what consciousness is consciousness of. Therefore the object of consciousness doesn't have to be a physical object apprehended in perception. It can just as well be abstractness or an 'ideal object'. These 'structures' of consciousness, i.e., perception, memory, fantasy, are called intentionality.

Based on these conditions, taking intentionality as central pivot of the phenomenon, the nature of data can be its genesis related to the moment in which the apprehension occurs. It is proposed therefore that data is the state of the object properties to the instant immediately prior to his apprehension by the subject.

Unlike content found in literature that relates data and information, a direct relationship of data and knowledge is here. Both can be understood as different dimensions of the intentionality of the subject.

2.3 On the ontological status of information

The concept of information is diverse in its meanings, from daily to technical use [Crnkovic and Hofkirchner, 2011]. Generally, the concept of information is closely connected to the notions of restriction, communication, control, data, form, instruction, knowledge, mental stimuli, pattern, perception, representation, record, among others.

The Greek concept of *form* is represented by various words related particularly with *view*: the view or appearance of something. The ancient words are μορφή (*morphē*), εἶδος (*eidos*) and ιδέα (*idia*), "the type, the idea, the form". "*Eidos*" was used by Plato – and, later on, by Aristotle – to indicate the ideal identity or essence of something (Theory of Forms). This word may also be associated with the concept of thought, proposition or even of concept. The words φαινόμενα (*phainomena*), "appearance" and φαινό (*phainō*) "of glow and light" still carry similar meaning.

The *Theory of Forms*, developed in Phaedo one of Plato's dialogues, realizes an understanding of various concepts and proceeds in his theory. For Plato the idea of concept, according to this new doctrine, is immutable, timeless, intellectually comprehensible and capable of a precise definition in a pure reasoning, because it is a real thing and it exists as an independent thing, an entity. Immortality of the soul is evidenced through our capacity of apprehending the concept of eternal, the object Plato called *form*.

In Physics, 'information' has a well-established meaning examples of which include quantum phenomena, and even the possibility of violating the second Law of thermodynamics by Maxwell's demon.

In information theory, entropy is a measure of the uncertainty associated with a random variable. Thus, the term usually refers to the Shannon entropy, which quantifies the expected value of the information contained in a message, usually in units such as bits. In this context, a 'message' means a specific realization of the random variable.

In statistical thermodynamics, Boltzmann's equation regards to probability related to the entropy S of an ideal gas to the quantity W (*Wahrscheinlichkeit*), which is the number of microstates corresponding to a given macrostate, or a configuration, or, yet, *complexions* (from Latin *complex* and the Greek suffix for diminutive *ions*):

$$(1) \quad S = k \; log \; W$$

In statistical mechanics, a microstate is a particular microscopic configuration of a thermodynamic system that the system may occupy with certain probability, in the course of its thermal fluctuations. As a counterpart, the macrostate of a system refers to its macroscopic properties, such as its temperature and pressure.

Macrostate is characterized by a distribution of probabilities of possible states in a particular statistical set of all microstates. This distribution describes the probability of finding the system in a determined microstate. In the thermodynamic limit, the microstate visited by a macroscopic system during its fluctuation possesses the same macroscopic properties.

The idea of information arises, then, from equation 1 and then we have:

$$(2) \quad S = \; k \; ln2 \; \; log_2 \; W$$

where:

$$\begin{aligned} k &= \text{Boltzman's constant} = 1,3806505 \cdot 10^{-23} \cdot J/K^{-1} \\ log_2 W &= \text{information} \\ kln2 = 0.69k &= \text{minimum of information} \end{aligned}$$

The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used the resulting units may be called binary digits, or more briefly bits [Shannon, 1948].

Any experiment by which an information about a physical system is obtained corresponds in average to an increase of entropy in the system or in its surroundings. This average increase is always larger than (or equal to) the amount of information obtained [Brillouin, 1953].

Used in the more specific sense of information theory, information is a quantity that can be measured in bits [Lloyd, 2008].

The idea of information measures is a more general definition of information than either Shannon information or algorithmic information content. Information measures allow for the identification of effective complexity, a measure of the amount of information required to describe a system's regularities or rule-governed behavior.

Due to the real configuration of atoms and molecules of a gas in a specific space being unknown, *entropy* is associated to the *information we do not have*, in such way that when information is obtained, entropy is reduced. Entropy measures ignorance [Gell-Mann and Lloyd, 1996]. Therefore, the nature of *information* is related to *change in entropy.* In microscopic states, *information* is *complexions* of a gas, of a system or of an object. Further study of these microscopic states allows inference of some properties of the information.

Despite the reluctance of Claude Shannon (1916–2001) to extend the scope of his Mathematical Theory of Communication to other areas of knowledge such as physics, it is possible to conceive an integration, or at least a correlation between the entropy as defined in Shannon'n theory, and Heisenberg's Uncertainty Principle, through the concepts of uncertainty inherent to these two theories [de Carvalho Pineda, 2006].

Matter in subatomic levels, as evidenced experimentally, is presented in a superposition state, may take simultaneously more than one physical microstate within a set of possibilities. These quantum states of subatomic particles, as provided by the Uncertainty Principle formulated in 1927 by Werner Heisenberg (1901–1976), had an uncertainty inherent in probabilistic models, and are only persisted in the observer at the moment of his apprehension (in this case, it is the scientific instruments used in the experiments that allow observation of subatomic particles behavior, such as extension of human senses of perception).

In a quantum sense, a position from a perspective of information as to how the representation of primary properties of matter would occur becomes necessary. Following this line of thought, it is possible to define data as a snapshot of this information, produced by the process of decoherence at the time of his apprehension and when occurs the decaying of superposition state to a single state persisted. Information on a quantum level would thus have an inherent uncertainty and be correlated to the set of possibilities of different quantum states, which could be assumed by subatomic particles.

Based on an analogy between the messages of Shannon's Mathematical Theory of Communication and a perspective of quantum information, as described above, it is possible to arrive at this correlation. Assuming that there is a finite set of possible quantum states, each of these quantum states with a certain probability of occurrence, it is clear that the uncertainty inherent in the perspective of quantum information will be reduced by the occurrence of the decaying of superposition state to a single state at the time of his apprehension. The amount of uncertainty reduced by the process of decoherence is related to the probability of occurrence of the quantum state apprehended at the moment of observation, in the same way that the receipt of a most probable message reduces the uncertainty less than the receipt of a message less probable.

This approach seems to adhere to both the theory of Shannon's and the concept of perspective Quantum information: "Information is a reduction of uncertainty given when you get an answer to a question".

In this approach, information represents the primary properties of the object independently of the subject and therefore strictly ontological, while this information would be data persisted at the exact moment of his apprehension by the subject, a snapshot. For this approach, there would be a fundamental distinction between data and information, making it necessary to agree to a proper terminology for this the model. From the arguments presented one can conclude that information is 'thing', i.e. information belongs to ontological level, and that data is the condition of the object properties on the instant immediately prior to its apprehension by the subject.

3 Elements of Architecture of Information

3.1 On space

Spencer Brown, in his book *Laws of Form*, [Brown, 1969], introduces the idea of form as a 'distinction' in a space, proposes a logical system and overcomes a few boundaries between mathematics and philosophy. The idea of distinction and the idea of indication, and that we cannot make an indication without drawing a distinction, are taken as a given. We take, therefore, the form of distinction for the form. By definition, distinction is perfect continence. A distinction is drawn by arranging a boundary with separate sides so that a point on one side cannot reach the other side without crossing the boundary. For example, in a plane space a circle draws a distinction. So, it is possible to postulate that there is no space without distinction. Once a distinction is drawn, the spaces, states, or contents on each side of the boundary, being distinct, can be indicated.

3.2 On state

Distinguished space has a state. Time is related to state. Distinguished space has content. Content is composed by things. Things have properties. Therefore, we assume that 'space of information' is the set of distinguished information in a distinguished space.

DEFINITION 1. A state $\mathcal{E}$ is an unique configuration of information in an interval of time Δ_t, denoted as $\mathcal{E}_{\Delta_t}$.

DEFINITION 2. A dynamic $\mathcal{D}$, of the spaces of information e is defined as:

$$\mathcal{D} = \{\mathcal{E}_{\Delta_{t_1}}, \mathcal{E}_{\Delta_{t_2}}, \mathcal{E}_{\Delta_{t_3}}, \cdots, \mathcal{E}_{\Delta_{t_n}}\}$$

3.3 On Architecture of Information

It is shown that the concept of AI can be applied in any information space. Examples of information spaces can be characterized generally as any set of things. We may think of a DNA structure as an information space, hence an AI, or, the solar system, or my office desk with its objects over, or even any object as particulars information spaces, i.e. it is possible consider all these meanings from the point of view of an AI.

The concern in organizing and structuring knowledge accompanies human history for centuries. The phenomenon of information explosion took even greater from the World Wide Web, and caused a growing concern with the systematization and access to knowledge. The concept of AI is to be inserted in this context, despite its origin dating from ancient times.

The term 'information architecture', as recorded in the literature, was first used by architect Richard Saul Wurman in 1976, who described it as "science and art of creating instructions for organized spaces". Wurman viewed the problem of searching, organizing and presenting information as analogous to the problems of the architecture of buildings, which will provide for the needs of its residents, because the architect needs to identify these needs, organize them into a coherent pattern that determines its nature and their interactions, and design a building that satisfies them.

The publications: Information Anxiety, [Wurman, 1989] and Information Anxiety 2, [Wurman, 2001] show an overview of the fundamental principles that motivated the author in his previous work, highlighting how dramatic is the explosion of information.

In view of Wurman, the assemblage, the organization and presentation of information served the purpose of the tasks characteristic of Architecture. The Information Architecture would be an expansion of the profession of architecture, but applied to spaces of information. And information structures interactions influence the world in the same way that the building structures encourage or limit social interactions. In 1976, Wurman organized the National Conference of the American Institute of Architects (AIA) and chose "The Architecture of Information" as the theme of the conference, coincidentally 100 years after the first meeting of the American Library Association. Today we have worked with a much broader idea of AI, particularly in the proposal of the School of Brasilia.

The first concept of AI arises as a result of above subsection, where for a given information space we consider the configuration of information, or the information being.

DEFINITION 3. AI is the states configuration of the constituent elements of the thing itself and its properties, characterized by space-temporality of distinguished information.

It is shown that this concept of AI can be applied in any information space. AI is inherent to any information space, in any domain. As a consequence of the definitions, one can say that there is no space of information without AI.

From an extension of the first concept of AI, we get the second concept, characterized by the need of changing of state. Better saying, considering a time interval Δ_t for an information space, an initial state $\mathcal{E}_{\Delta_{t_1}}$ of a configuration of properties of their constituent elements (snapshot information), a changing to future state $\mathcal{E}_{\Delta_{t_{n+1}}}$ is performed by a transformation. Evidently the intermediate states are characterized by a dynamic $\mathcal{D}$ belonging to this context.

DEFINITION 4. A *transformation* is a set of events, applied to a particular state, in order to provoke changes to future states.

This perspective of AI can be applied in any situation. For example, one can show that there is no information system without AI. In fact, when designing an organization's information, and information systems are patterns of organization of information, you can not do so without regard to AI, or AI is inherent to any information system.

4 Architecture of Information as Social Science

The third concept is related to the perspective of applying a *transformation*. This perspective is related to the performance of a *subject*. As a result, we are in the field of Applied Social Sciences, or on the application of transformations performed by an individual or by a *subject*.

Philosophy of Language originated the *Theory of Speech Acts*, at the beginning of the sixties having been appropriated by Pragmatics later on. John Langshaw Austin (1911-1960), philosopher of the Oxford Analitic School, followed by John Searle and others, understood language as a form of action: "all speech is an action". The various types of human action realized through language were reflected upon: *the speech acts*. The Theory of Speech Acts was published posthumously, in 1962, in the book *How to do Things with Words* [Austin, 1962, page23]. For Austin, speech is not only a way of passing on information, but also – and foremost – it is a way of acting on the interlocutor and on the world.

Until then, linguists and philosophers, in general, thought that the claims serve only to describe a state of things, and thus were true or false. Austin calls into question this view of descriptive language, showing that certain statements do not serve to describe anything, but to take action.

The *School of Brasilia* proposes an extension of Austin's Speech Act in the sense that "all doing is an act" and "all act is a transformation".

DEFINITION 5. *Transformation acts* are sets of events, applied to a particular state, by subject, in order to provoke changes to future states.

In the literature there are dozens of books whose titles we see the term 'information architecture'. With the theoretical framework presented, examples of AI in the literature become particular cases or examples, of the general concept. Some of the titles available are: *Enterprise Information Architecture, Information Assurance Architecture, e-Gov Information Architecture, Federal Enterprise Architecture, Strategic Information Architecture, Supply Chain Information Architecture, Web Information Architecture, Information Security Architecture*, etc.

Enterprise Information Architecture is the most recognized by industry. In the proposal of this approach, we can consider an enterprise as an information space. If we take a snapshot at some time interval, we obtain the configuration information of the enterprise. This is the initial situation. We may analyze this situation from the perspectives of the current paradigms for organizing information in Enterprises. From a future perspective of a possible desired state of the current situation, it is possible to design a future state. Acts of transformation will occur to make it possible to achieve the desired future state. This same approach can be applied to any space, where it is necessary to consider any aspect of the "information life cycle", from a perspective of SI, and consequently AI.

BIBLIOGRAPHY

[Austin, 1962] John L. Austin. *How to Do Things With Words.* Oxford University Press, Oxford, 1962.

[Bates, 2005] M.J. Bates. Information and knowledge: an evolutionary framework for information science. *Information research*, 10(4):10–4, 2005.

[Brillouin, 1953] L. Brillouin. The negentropy principle of information. *Journal of Applied Physics*, 24(9), April 1953.

[Brown, 1969] George Specer Brown. *Laws of form.* George Allen & Unwin London, 1969.

[Crnkovic and Hofkirchner, 2011] G.D. Crnkovic and W. Hofkirchner. Floridi's "open problems in philosophy of information", ten years later. *Information*, 2(2):327–359, 2011.

[Dale, 2002] Adrian Dale. Letter 12: Information architecture: the next professional battleground? *Journal of Information Science*, 28(6):523–525, 2002.

[de Carvalho Pineda, 2006] José Octávio de Carvalho Pineda. A entropia segundo Claude Shannon: desenvolvimento do conceito fundamental da teoria da informa cão. Master's thesis, PUC-SP, São Paulo, 2006.

[Dillon, 2002] A. Dillon. Information architecture in jasist: just where did we come from? *Journal of the American Society for Information Science and Technology*, 53(10):821–823, 2002.

[Doucette *et al.*, 2007] D. Doucette, R. Bichler, W. Hofkirchner, and C. Raffl. Toward a new science of information. *Data Science Journal*, 6(0):198–205, 2007.

[Floridi, 2004] Luciano Floridi. Open problems in the philosophy of information. *Metaphilosophy*, Volume 35(Number 4):pp. 554–582, July 2004.

[Gell-Mann and Lloyd, 1996] M. Gell-Mann and S. Lloyd. Information measures, effective complexity, and total information. *Complexity*, 2(1):44–52, 1996.

[Haverty, 2002] Marsha Haverty. Information architecture without internal theory: an inductive design process. *Journal of the American Society for Information Science and Technology*, 53(10):839–845, 2002.

[Hessen, 1978] Johannes Hessen. *Teoria do conhecimento.* A. Amado, Coimbra, 7 edition, 1978.

[Hirschheim, 1985] R. Hirschheim. Information systems epistemology: An historical perspective. *Research methods in information systems*, pages 13–35, 1985.

[Hofkirchner, 1999] W. Hofkirchner. *The quest for a unified theory of information: proceedings of the Second International Conference on the Foundations of Information Science.* Routledge, 1999.

[Husserl, 1961] Edmund Husserl. *Recherches logiques.* Presses universitaires de France, France, 1961.

[Husserl, 1963] Edmund Husserl. *Idées directrices pour une phénoménologie.* Gallimard, France, 1963.

[Lloyd, 2008] Seth Lloyd. Quantum information science. USA, 2008.

[Robins, 2002] David Robins. Information archtecture in library and information sicence curricula. *Bulletin of the American Society for Information Science and Technology*, 28(2):20–22, 2002.

[Shannon, 1948] Claude Elwood Shannon. *A Mathematical Theory of Communication*, volume 27. Bell System Technical Journal, Octuber 1948.

[Smith, 2011] David Woodruff Smith. Phenomenology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* The Metaphysics Research Lab, Stanford University, fall 2011 edition, 2011.

[Wurman, 1989] Richard Saul Wurman. *Information anxiety.* Doubleday, New York, 1989.

[Wurman, 2001] Richard Saul Wurman. *Information Anxiety 2.* Hayden/Que, Indianapolis, USA, 2001.

Mamede Lima-Marques
Centre for Research on Architecture of Information – FACE/UnB and
Faculty of Information Science

University of Braslia
Brasilia, Brazil
e-mail: mamede@unb.br

Mamede Lima-Marques is full professor of Information Science at University of Brasilia. He received his Doctorat in Informatique (1992) from Université Paul Sabatier, Toulouse III, France. He has been working on Technology and Architecture of Information, especially in epistemology, science and praxis levels.

@incollection{lima-marques:2011,
Author = {Mamede Lima{--}Marques},
Booktitle = {Logic without Frontiers: Festschrift for Walter Alexandre Carnielli on the
occasion of his 60th Birthday},
Title = {Outline of a theoretical framework of Architecture of Information: a School
of Brasilia proposal},
Editor = {Jean{-}Yves B\'{e}ziau and Marcelo Esteban Coniglio},
Month = {December},
Publisher = {College Publications},
Series = {Tribute Series},
Volume = {17},
Address = {London},
Year = {2011}}