A NEW APPROACH TO INFORMATION SCIENCE: FRAMING A NATURALISTIC PERSPECTIVE

Jorge Navarro, Raquel del Moral & Pedro C. Marijuán

Abstract: *In recent decades, an increasing variety of research fields are converging into information science conceptualizations. They are accompanied by astounding new uses of knowledge and even more astounding social transformations that revolve around information technologies. Whether a robust information science will finally emerge may not only depend on successful discussions about the philosophy of information and the social impact of the new technologies. The most important adjustment to make is about framing a "new way of thinking" or new perspective about information science itself: an inner philosophy interconnecting research practices in fundamental areas of the new science. Like in the historical birth of other major sciences, the empirical, comparative understanding of informational phenomena and informational entities should take place first. Further, a naturalistic perspective, biologically inspired, would help in determining what model systems should be adopted for advancing the comparative study of informational entities. Rather than attempting ill-fated definitions [is information definable at all?] or closely remaining within the confines of information theory, a naturalistic, empirically oriented strategy offers room for advancement. In the end, crafting a great scientific domain around information science –rather than around computing – should be a common goal for the scholars and researchers both from natural sciences and engineering and from social sciences involved in these new studies.*

Keywords: *Information Flow, Cell Production, Communication, Information Science,*

ACM Classification Keywords: *Theory*

The informational transformation of contemporary science

The emergence of a renewed information science along the stunning *informational uprising* of contemporary societies would hardly be a surprise. The historical parallel with the emergence of thermodynamics in the generations closer to the industrial and social experimentation with steam-engines provides a sensible argument. As a matter of fact, most of scientific growth during the last two decades has been produced in information-related fields: bioinformatics and *omic* revolution, neuroinformatics, robotics, artificial intelligence, quantum information, new computing tools, models and simulations, virtual economy, virtual learning, and so on. The stunning transformation of many branches of science and technology fueled by computers, including physics itself, is leading to reconsider the natural structures and processes of information, not only in the quantum-cosmological realms or in the biomolecular, or in the human-to-machine interaction, but particularly in the human-to-human interrelationship and in the bonding structures of societies as the great drivers of global communication networks.

At the time being, the efforts to build a coherent discipline around the manifold scientific manifestations of information are only partially succeeding, but they are gaining more and more momentum and should be contemplated with high hopes and a lofty ambition. Rather than attempting the construction of another average or standard discipline, information science is about the making out of one of the “great scientific domains” of contemporary knowledge [Rosenbloom, 2013]. It is information science –rather than computer science and technology as argued by Rosenbloom– the strategic domain that has the mandate to perform the greatest cognitive interrelationships that our social system of knowledge demands: the natural sciences with the social sciences and humanities; information with meaning, knowledge, and intelligence; the individual creation of knowledge with its social organization and management; the limitations of human communication with social complexity and political structures; the anthropogenic economic flows with the geogenic planetary flows.

Providing the next generation with a more coherent and workable system of knowledge would be a valuable legacy from our generation indeed—by having launched in the chaotic sea of disciplines an *information hub* of fresh design. In this regard, the Foundations of Information Science [FIS] initiative was launched almost two decades ago with the explicit goal of developing a ‘vertical’ or ‘transdisciplinary’ science connecting the different threads and scales of informational processes [Conrad, 1996; Marijuán, 1996]. The five FIS conferences held until now represent a valuable experience to guide in the transformations that have to be made within information science itself. Actually, the FIS meeting in the 2014 ITHEA GIT Conference represents a celebration of FIS 20th anniversary.

Can information be defined? Advancing a new way of thinking

Rather than discussing on the hundreds of information definitions proposed [Lenski, 2010], let us definitely discard that. Information cannot have a universal definition because it corresponds to the open-ended interactions that a subject can engage with whatsoever elements or objects in its environment. Information pertains to an unmediated coupling between objects-subjects, and involves their contact. We demand either that a direct physical contact be established between the pair object-subject, or that a “channel” be arranged to bring to the adjacency of the subject those objects or their very effects.

Advocating the undefinability of information does not mean the ineffability of the term; rather what it means from a naturalistic perspective is that the plurality of subjects (can’t we talk about information in cells, in organisms, in individuals, or in social bodies as legitimate subjects?), plus the vastness of possible coupled objects, and the multiplicity of coupling modes between subjects and objects do not permit any universalistic definition. Achieving a universal information definition and even more a unified theory of information represents a similar attempt to the “universal non-linear theory” that John von Neumann was quipping as “the non-elephants theory”.

Changing the main analytical focus from the “name” information as an object in itself to the “adjective” informational, what entities might be attributed the quality or characteristic of being informational? The response is concise: those capable of both *communication* and *self-production*. The first trait is well known in many different fields related to information science—apparently. From our perspective, however, the communication theme demands full clarification. In the same way that the energetic-metabolic aspect of biological self-production was better understood once the *energy flow* was properly recognized and characterized [Morowitz, 1968], the *information flow* of communication demands its own general conceptualization, beyond current conceptions restricted to specialized uses.

How the informational entity organizes its openness towards the environment in the communication exchanges? The capture of the surrounding signaling flows is propitiated by *ad hoc* designs that tend to cover increased regions of space throughout special sensing structures, incorporating fractal distributions for instance. In general, the communication trait has been evolved adaptively: exploiting the maximum surrounding flows with the highest possible diversity of objects/events, using the lowest capture energy, and applying the faster processing rhythms compatible with the own structures. This can be observed in cellular signaling systems, in nervous systems, and in the communication systems of human organizations and in societies at large. The external communication flow, or better, the *information flow,* antedates and permeates the evolution of whatsoever structures of complexity.

In the other essential trait of the informational entity, *self-production*, streams or flows of objects, destined to fuel the self-production processes, are introjected or channeled inwards to be consumed—also appearing a parallel flow of extrojected material outputs. The ingoing and outgoing flows constitute the metabolism of the self-producing informational entity. As said, the *energy flow* has been well-characterized for the biosphere and ecosystems [Morowitz, 1986]; technically it is also a crucial ingredient of industrial engineering, urban planning, and economic management. From this energetic point of view, the informational entity appears as an open system, out from equilibrium, endowed with the required energy-entropy-matter flows, and often involving the generation of well-ordered dissipative structures. But these flows do not spontaneously run to encounter the entity; rather they have be searched, channeled, and internalized by following optimized behavioral and search strategies [Holland, 2012].

Although the information flow and the energy flow may partially overlap in their constitutive components, they are treated in a highly different way: "reading" the environment becomes utterly different from, and prior to, "eating" it. The living cell is the clearest instance [Marijuán, 2010]. In general, the high-energy, highly valuable self-production flows will be anticipated, detected, and captured by the faster and cheaper communication flows tended with the surrounding environment. A frequent commonality of forms occurs too, manifested in supporting structures that often display fractal forms derived from the necessity to cover a region of space and to transport the affordances of both the material "stuff" and the communicational "fluff" to a center [Bejan and Peder, 2012].

Around *the information flow*, a new conceptual framework might be established. At its core, information science should deal with informational entities that exist "in the flow", that both communicate and self-produce. By intertwining the two operational realms a seamless existential unity is generated. As Lanham [2006] has put for our societies, it is the almost immaterial *fluff* of communication which provides guidance to the constitutive material *stuff*. And vice versa, it is social life which provides the functions and goals of communication systems—always in service of social self-production. Hence an "economy of attention" emerges: how the networks of self-production processes look at the world for their survival.

This informational approach can be generalized upwards. Actually every new organization realm develops some specific information flow; at the frontiers or junctures between compartments the previous, more basic information flow is maintained. This successive composition of information layers seems valid for organisms, individuals and societies. However, a homogeneous description is out of hand—all these informational entities belong to conceptual disciplines worlds apart [Marijuán, 2010]. A parsimonious strategy, herein advocated, may consist in attempting a relatively independent description for each one of the scales, to be followed then by some tentative abstractions/conclusions interrelating them “vertically” [Conrad, 1996; Marijuán, 1996].

This new perspective is both naturalistic and empirically oriented. Its focus is in the autonomous existing informational entities, taking as their distinctive characteristic the intertwining of communication and self-production flows. Natural informational entities to study as model systems are cells, organisms [nervous systems], individuals, and societies. In spite of the disciplinary differences and conceptual obstacles, a vertical understanding of the communication and self-production flows seems feasible. The realization of multiple empirical studies on informational aspects of these natural entities will help to advance the common understanding.

The primordial informational entity: the living cell

Prokaryotic cells will be taken as a practical case (almost the single one in this brief paper) to discuss the fundamentals of the new perspective on information science. These cells contain perhaps one of the most interesting panoramas for informational analysis: where at least nearly every intervening phenomena can be molecularly described. In the description that follows the language has to switch necessarily towards the molecular-biological.

How different looks the animate from the inanimate! Just counting the number of different molecular species teeming up at the interior of a prokaryotic cell (the multiple classes of peptides, enzymes, receptors, phospholipids, RNAs, DNA, metals, nutrients, ions, etc.) the figure is staggering: in the order of 10,000 species, unthinkable of any regular physical system that magnitude - a mere cubic micron. The special properties of water, the polymerization strategies, the organization in “architectures”, the coded correspondence between triplets and amino acids, the folding process, the catalytic properties of enzymes, the semi-permeable membrane, etc. should be invoked as molecular builders of the cellular order, providers of “processing power” to this basic informational entity.

From our informational perspective, it is interesting that the vocabulary developed by molecular biologists was crafted, almost from the very beginning, around the *communication* metaphor: molecular recognition, genetic and epigenetic codes, transcription, translation, processors, messengers, signalling systems, effectors, transducers, second messengers, regulators, interferences, complexes, networks, modules, etc. It is quite revealing that the founding fathers of molecular biology so heavily relied on the communication metaphor for naming the molecular events they were uncovering. Advertently or inadvertently, they were fundamentally right.

Briefly analyzing the two essential informational traits, *self-production* is basically performed by a “network society” of specialized enzyme and protein agents that are coded onto the DNA “sequential architecture” and are continuously exchanging information about their specific activities thanks to the especial solvent properties of the water matrix. In response to signals of the environment, to be described later, enzymes and proteins are synthesized (and also degraded) out from the sequential information of DNA and RNA, which are themselves incessantly subject to evolutionary combinatory games [Marijuán, 2002, 2010]. The whole productive processes culminate in the regularity of a specific cell-cycle that is open to the environment.

Nevertheless, by itself the transcription network expressing the DNA genes is closed, “blind”. In other words, the coupling between the sequential genetic architecture and the diluted architecture of enzymes and proteins needs the injection of further adaptive capability to respond to environmental demands. This is done by means of signaling guidance, so to partially deploy the genetic circuits in response to relevant happenstances of the environment and also of the cellular interior. Most of the *topological governance* of the transcription regulatory network, the decision of what parts should be activated or what particular circuits should be inhibited, is achieved by means of the cellular signaling system [Navarro, 2010].

In the implementation of *communication*, a variety of signaling systems may be found in prokaryotic cells, ranging from simple transcription-sensory regulators [a single protein comprising two domains], to those systems of multiple components and interconnected pathways that regulate key stages of the cell cycle, such as latency, pathogenesis, replication, and dispersion. A basic taxonomy of bacterial signaling systems was proposed somewhere else [Marijuán *et al*., 2010], which was centered on “the 1-2-3 scheme.” In a specific bacterium, for instance *M. tuberculosis* or *E. coli*, the number of different signaling pathways is close to one hundred, the majority belonging to the “one component systems” class, with around one or two dozen members of the “two component systems” class, and a few members of miscellaneous classes (“three component systems” and others). Every one of these signaling pathways may be present in a range of one or two orders of magnitude—some dozen to some hundred molecules. Conversely, in the eukaryotic signaling system, hundreds of different pathways and thousands of dedicated molecular agents may participate [receptors, ion channels, transducers, amplification cascades, second messengers, intermediate effectors, final effectors]; and they can be arranged differently in each tissue [Marijuán *et al*., 2013]. The flows of information crossing throughout the different receptors and further transduced along the signaling pathways are informing in prokaryotic cells about the presence of metabolic items in the environment, while in eukaryotic cells these very flows mostly relate to direct cell-to-cell communication about developmental and physiological matters. In both cases the information flows become systematically transformed into variations of the self-producing structure.

Let us emphasize the difference between signals and metabolites. Essential molecular components of the environment are continuously scanned by the signaling system: nutrients, ions, metals, peptides, amino acids, toxics, signaling hormones, etc. Once important molecular presences are detected, the system activates gene expression programs or directly induces changes in the cytoplasm and membrane. But signals themselves are left untouched: the molecules participating as “messengers” are merely recognized as signals and do not suffer further transformations as happens with metabolites along metabolic pathways.

Thereafter, *meaning* may be defined throughout *molecular mining*: as the [signal] induced changes in components and connectivity of the constitutive enzyme-protein populations and the associate metabolites and substrates. The *relevance* or *value* of the signal can subsequently be considered and gauged —this in general corresponds to second messengers and the cell cycle “checkpoints”. Completion of the cell cycle always appears as the fundamental reference. The phenomenon of *knowledge* may be appended too, once the generative codes of the protein agents implementing successful responses have been evolutionarily selected, refined, and cohered within the life cycle. The *recombination* strategy performed upon the DNA stretches emerges as an abstract problem-solving tool of far reaching evolutionary consequences, basically understood as “domain recombination” [Marijuán and del Moral, 2007; del Moral *et al.*, 2011].

To conclude this brief description of the simplest informational entity [at the time being the only one we can describe molecularly with some reasonable approximation], the methodological problems found are not insurmountable. Like in other more mature disciplines, pragmatic solutions may also be found for the informational analysis of living entities: those safe guidelines under which an efficient reduction of complexity may be achieved.

The generation of new domains of complexity: social knowledge

The informational approach we have drafted for the cell can somehow be generalized upwards. The information flow – the communication sphere – exists around the successive compartments of cells and organisms, accompanied always by a parallel flow of nutrients and maintenance and repair components. Actually every new organization realm develops some specific information flow; at the frontiers or junctures between compartments the previous, more basic information flow is maintained—as happens for instance with “old” molecular diffusion in between “modern” synaptic contacts. This successive composition of information layers seems valid for organisms, individuals and societies too.

Is there any “natural” information flow we can point at in human societies? Anthropologically, the most distinguishing trait of our species is the use of language for communication among individuals. Evolutionarily, these communication activities, basically face-to-face conversation, have taken a genuine evolutionary category along the “social brain hypothesis” recently framed [Dunbar, 2004]. It means that the most plausible evolutionary hypothesis on the hypertrophic growth of human brain relates to the adaptation to bigger social groups and to the use of language as a fundamental vehicle of communication and as new form of social “grooming” [Allman, 1999; Silk, 2007]. Oral exchanges become the way to share group experiences, techniques, knowledge, culture, etc., far beyond any previous “cultural repertoires” of other anthropoids. When the open-ended combinatory structure of language is finally pictured into written “permanent” form [the parallel with the “permanent” DNA of the cell is unavoidable], a whole new world for the circulation of information and knowledge is generated in human societies.

The historical fact is that the development of social complexity has become irreversibly linked to a chain of inventions for communication and knowledge generation. There is a succession of fundamental inventions that dramatically alter the “infostructure” of societies: numbers, writing, alphabet, codices, universities, printing press, books, steam engines, means of communication, computers, etc. [Hobart and Schiffman, 1998]. Another crucial aspect related to human cognition is the *recombination of knowledge*. It is a phenomenon which has passed almost unnoticed in traditional philosophy of science notwithstanding its massive presence in contemporary scientific-technological societies [Scott, 1998; Arthur, 2009]. The term is not exclusively related to the social or to the biomolecular fields (DNA domain recombination). It is quite relevant that contemporary neuroscience has also recognized the importance of the recombinatory dynamics in the “neuronal workspace” of individuals [Dehaene, 2009]. The social creation of knowledge, and obviously the growth of science, derives from knowledge recombination processes taking place in the cerebral workspace of individuals. The strict conditions put by the scientific method would represent the ways and means to directly interconnect standardized individual perceptions and actions beyond time, space, and cultural limitations, allowing the social decomposability of problems and the occurrence of knowledge recombination dynamics at a global and intergenerational scale [Marijuán *et al*., 2012].

A new vantage point about the whole social dynamics of scientific knowledge is needed in today's "information societies"—our sciences have become a Babel Tower of more than 6,000 scientific and technological disciplines, where classical information [library] science is helpless to provide any interesting guidance to society [Marijuán *et al*., 2012]. Information science should promote a new vision to help make sense of the historical expansion of science, and of the requirements of human knowledge in action.

Information science, properly developed and linked with computer science and mathematics, should constitute one of the Great Domains of contemporary science. The informational would go together with the physical, the biological, and the social: constituting the four great domains of science. Rather than attempting the construction of another average or standard discipline, information science is about the making out of one of the "great scientific domains" of contemporary knowledge.

Bibliography

[Allman, 1999] J.M. Allman Evolving Brains. Scientific American Library. New York, NY, USA. 1999.

[Arthur, 2009] B.W. Arthur. The Nature of Technology: What It Is and How it Evolves. New York: The Free Press. 2009.

[Bejan and Peder, 2012] A. Bejan and J. Peder. Design in nature. Doubleday. New York. 2012.

[Conrad, 1996] M. Conrad. Cross-scale information processing in evolution, development and intelligence. BioSystems. 38 97-109. 1996.

[Dehaene, 2009] S. Dehaene. Reading in the brain. Penguin. New York. 2009.

[del Moral et al., 2013] R. del Moral, J. Navarro, P.C. Marijuán. New times and new challenges for information science. Proceedings of the 5th Foundations of Information Science. Information 2014, 5, 101-119. 2013.

[Dunbar, 2004] R. Dunbar. The Human Story: A New History of Mankind's Evolution. Faber & Faber Ltd. London. 2004.

[Hobart and Schiffman, 1998] M.E. Hobart and Z.S. Schiffman. Information Ages. The John Hopkins University. Baltimore, MD, USA. 1998.

[Holland, 2012] J.H. Holland. Signals and Boundaries: Building Blocks for Complex Adaptive Systems. The MIT Press. Cambridge, MA, USA. 2012.

[Lanham, 2006] R.A. Lanham. The Economics of Attention. The University Chicago Press. Chicago, IL, USA. 2006.

[Lenski, 2010] W. Lenski. Information: A Conceptual Investigation. Information. 1 [2] 74-118. 2010.

[Marijuan and del Moral, 2007] P.C. Marijuán and del R. Moral. The informational architectures of biological complexity, In: Computation, Information, Cognition –The Nexus and The Liminal. Dodig-Crnkovic G. and Stuart S. [eds.]. Cambridge University Press, Cambridge. 2007.

[Marijuán et al., 2010] P.C. Marijuán, J. Navarro, R. del Moral. On prokaryotic intelligence: strategies for sensing the environment. BioSystems. 99 94-103. 2010.

[Marijuán et al., 2012] P.C. Marijuán, R. del Moral J. Navarro. Scientomics: An emergent perspective in knowledge organization. Knowledge Organization. 39 [3] 153-164. 2012, [Marijuán et al., 2013] .C. Marijuán, R. del Moral 2012, J. Navarro. On eukaryotic intelligence: Signaling system's guidance in the evolution of multicellular organization. Biosystems. Vol. 114, Issue 1: 8–24. 2013.

[Marijuán, 1996] P.C. Marijuán. First conference on foundations of information science: From computers and quantum physics, to cells, nervous systems, and societies. BioSystems. 38 87-96. 1996.

[Marijuán, 2002] P.C. Marijuán. Bioinformation: untangling the networks of life. BioSystems. 64 11-118. 2002.

[Marijuán, 2010] P.C. Marijuán. Knowledge recombination: on the informational adaptability of cells, nervous systems, and societies. International Journal “Information Theories and Applications. Vol. 8, No. 1, pp. 3-15. 2010.

[Morowitz, 1968] H. J. Morowitz Energy flow in biology: biological organization as a problem in thermal physics. Academic press New York, London 1968.

[Navarro, 2010] J. Navarro. Transcriptional Regulatory Network of M. tuberculosis: Functional and Signaling Aspects. Master Thesis. Universidad de Zaragoza, Zaragoza, Spain. 2010.

[Rosenbloom, 2013] P.S. Rosenbloom. On Computing. The Fourth Great Scientific Domain. The MIT Press. Cambridge [MA]. 2013.

[Scott, 1998] J.C. Scott. Seeing Like a State. Yale University Press. New Haven, CT, USA. 1998.

[Silk, 2007] J.B. Silk. Social Components of Fitness in Primate Groups. Science. 317: 1347-51. 2007.

Authors' Information

Jorge Navarro López *- (jnavarro.iacs@aragon.es), Graduate in Chemical Engineering, University of Zaragoza. Master in Physics and Physical Technologies, University of Zaragoza. In 2008 he joined the Bioinformation Group of the Aragon Health Sciences Institute (IACS) as Research Assistant. His main research interest is on Systems Biology, but also cellular signalling systems, cellular intelligence and the nature of biological information.*

Raquel del Moral Bergós *- (rdelmoral.iacs@aragon.es), Graduate in Biology, Complutense University of Madrid. Master in Molecular and Cellular Biology, University of Zaragoza. In 2008 she joined the Bioinformation Group of the Aragon Health Sciences Institute (IACS) as Research Assistant. Her main research interest is on Neuroscience, but she is also keen on the nature of social information, and the parallel with the biology.*

Pedro C. Marijuán *- (pcmarijuan.iacs@aragon.es), Senior Researcher, is the leader of the Bioinformation & Systems Biology Group at the Aragon Health Sciences Institute (IACS). Engineer and Doctor in Cognitive Neuroscience (PhD Thesis on "Natural Intelligence", University of Barcelona, 1989). During more than 20 years he has advanced research on the nature of biological (cellular) information, communication, knowledge, and intelligence. He has covered both the intracellular molecular mechanisms and the systemic, behavioural (brain), and social realms. He was co-founder with Michael Conrad of FIS (Foundations of Information Science).*