

Article

New Times and New Challenges for Information Science: From Cellular Systems to Human Societies

Raquel del Moral *, Jorge Navarro and Pedro C. Marijuán

Bioinformation and Systems Biology Group, Aragon Institute of Health Science (IACS), Avda San Juan Bosco, 13, Zaragoza 50009, Spain; E-Mail: jnavarro.iacs@aragon.es (J.N.); pcmarijuan.iacs@aragon.es (P.C.M.)

* Author to whom correspondence should be addressed; E-Mail: rdelmoral.iacs@aragon.es; Tel.: +34-976-71-44-76; Fax: +34-976-71-55-54.

Received: 11 December 2013; in revised form: 24 January 2014 / Accepted: 26 January 2014/

Published: 14 February 2014

Abstract: The extraordinary scientific-technical, economic, and social transformations related to the widespread use of computers and to the whole information and communication technologies have not been accompanied by the development of a scientific “informational” perspective helping make a coherent sense of the spectacular changes occurring. Like in other industrial revolutions of the past, technical praxis antedates the emergence of theoretical disciplines. Apart from the difficulties in handling new empirical domains and in framing new ways of thinking, the case of information science implies the difficult re-evaluation of important bodies of knowledge already well accommodated in specific disciplines. Herein, we will discuss how a new understanding of the “natural information flows” as they prototypically occur in living beings—even in the simplest cells—could provide a sound basis for reappraising fundamental problems of the new science. The role of a renewed information science, multidisciplinary conceived and empirically grounded, widely transcends the limited “library” and knowledge-repositories mission into which classical information science was cajoled during past decades. Paraphrasing the Spanish philosopher J. Ortega y Gasset, the overhaul of information science itself becomes “the challenge of our time”.

Keywords: information science; natural information flow; communication; self-production; cellular model system; knowledge recombination; scientomics

1. Introduction: An Exciting New Epoch for Information Science

Looking 20 years backwards, just the time when the FIS initiative (Foundations of Information Science) begun, how different the panorama for information studies was. In 1992, Michael Conrad and the senior author of this paper undertook the adventure of convoking a number of researchers and scholars from many different disciplines and countries who exchanged views on the viability of the foundational project and soon gathered in an international FIS conference, in Madrid 1994 [1,2]. It was the first of the FIS series, to be followed by Vienna 1996, Paris 2005, Beijing 2010 and Moscow 2013. A discussion list, created after the Vienna conference, has been quite active all these years [3].

What was the status of information science at that time, 20 years ago? Arguably, it was not very relevant. In the early 60s, as a result of the two conferences held at the Georgia Institute of Technology, an information science of sorts was launched, with the particular mandate of searching for unified human-human and human-machine communication; but the resultant attempt was scarcely influential. In a decade or so, most of its multidisciplinary luster was lost in favor of the nascent computer science and artificial intelligence communities [1,2]. Actually, in the late 80s and early 90s, a number of multidisciplinary adventures were launched around computer science (parallel processing), artificial intelligence, artificial life, theoretical biology, biocomputing, bioenergetics, chaos theory, complexity theories, theoretical physics, *etc.*

Many of these enterprises are still alive and well, but the general background has sensibly changed in favor of the information studies. Developments closely related to the informational approach are increasingly present in the scientific avant-garde of today: quantum information science, bioinformatics, systems biology, synthetic biology, neural connectome, brain atlas projects, network science, information society, plus the World Wide Web's explosion and all those curious terms related to new ways of communication and social cognition—blogs, Wikipedia, LinkedIn, Facebook, YouTube, Twitter, Whatsapp, Instagram, crowdfunding, “the cloud”.

In our time, the enterprise of a renewed information science looks easier and closer, and even more necessary, as witnessed by the increasing number of journals, books, monographs, scholarly publications, research conferences, and so on that are devoted to information. Notwithstanding this more favorable scientific and cultural climate, information science is still a field in disarray. The traditional approach has not rallied and joined forces with the new foundational attempt yet, and as B. Hjørland discusses [4], it is curious that parallel series of conferences are taking place respectively convoked by FIS and CoLIS (Conceptions on Library and Information Science) without any overlap in the participant authors. In the communication engineering field, the ITHEA organization (Information Theory and Applications) has also held regular conference series in the quest for a General Information Theory. Within computer sciences themselves different specialized branches and research institutions are also involved in “information science” development and in the potential unification under the umbrella of computing [5]. In short, most of the conceptual and institutional obstacles already found in the early 90s still persist: flimsy disciplinary conceptual structures, scarce scholarly recognition, disinterest by neighboring disciplines, “occupied” territories, definition obsession (the cottage industry devoted to *what is information!*), excessive bias towards philosophical discussions, lack of research projects, *etc.*

At stake is whether a departure towards more empirical and applied research themes would overcome those obstacles directly—and would gradually achieve recognition by staying closer to important social and scientific problems where the informational point of view may perform a useful problem-solving work. There are plenty of research fields where information science can interact with other major disciplines: from molecular recognition and cellular signaling systems to biosystems and integrative physiology; from the neurodynamics of behavior to emotions and consciousness correlates; from social media to individual bonds and complex social organizations; from a better explanation of the combinatory dynamics of knowledge to quantum and cosmological interpretations. Some of these themes will be briefly discussed in this paper.

An important factor that impels the development of a renewed—and critically oriented—information science is precisely the stunning transformative success of Information and Communication Technologies (ICTs). Their accelerated expansion during last two decades has produced speedy changes in the way individuals communicate, learn, store their knowledge, externalize their memories, travel, socialize, organize their work, and dispose their homes. It is quite curious a social phenomenon: the enormous intensification of “artificial” information flows in most aspects of individual and collective life has forced a myriad of relational and cultural *adaptive* changes. A paradoxical outcome is the increase of social atomization—the extent to which a stressful sense of impoverished socialization, of loneliness and depression, and even of diminished mental health is accompanying the “information society” advancement. Although many other social, cultural, and economic factors could be involved, that atomization effect seems to be conspicuously related to the pervasive presence of ICTs. Looking more carefully, the discussion on this conjecture should also depend on a deeper understanding of the *informational* characteristics of human individuals, the limitations they confront when managing the different information flows around them. Clarifying those informational/communicational limitations is one of the essential contributions that a renewed information science should provide. We will approach some of these matters in Section 2 and in Section 4.

Changing the ways citizens communicate, and modifying the nature and intensity of the information flows around them, impinges on numerous aspects of social life. In actuality, the most revolutionary inventions in history have been those related to the channeling of social information: tablets, scrolls, and alphabet (Ancient science and Greek science); codices, copyists’ schools, universities (Medieval Awakening); printing press, bookshops, and libraries (scientific revolution); new energy and production technologies, communication and education systems (industrial revolution); engineering, electronics, new materials, and giant corporations (scientific-technological 2nd revolution); media, computers, and internet (information revolution). See the corresponding references in [6–11]. However, we are barely at the beginning of the social history of information, and not too far from the pioneering ideas of W.J. Ong and H.M. McLuhan.

The increase of social connectivity underlies another distinctive characteristic of contemporary societies: the global expressions of social discontent. The possibilities of direct, massive communication among throngs of citizens have fostered unheard forms of “revolt of the masses”—echoing J. Ortega y Gasset [12]: they are bottom up self-organized but with immediate contagion at the national and international scales. In the same way than a global interconnection of national economies, of financial systems, and of health risks have dramatically expanded and amplified

the scope of world crises, the systems of political representation are nowadays challenged by delegitimation crises and sudden storms of discontent. Authoritarian regimes as well as classical, well-consolidated democratic systems are suffering other unanticipated consequences of ICTs' brute connectivity: either in new forms of "direct action" by isolated-networked individuals and self-organized groups, or in new instances of hidden global controls. It may look ironic, but the hyper-connected world shows diminished resilience to internal and external threats [13]. In parallel with the paradoxical effects of hyper-connection in the impoverished socialization of the individual, it is the spectrum of a "new world disorder" that we see on the rise, rather than the perspective of an emerging *ecumene*.

However, discussing the possible deleterious effects or unintended consequences of the ongoing information revolution is not the main goal of this paper. Rather, *our basic proposal is the development of a new conception on information, biologically inspired, so that a new understanding might be gained on some unapproachable social themes of informational nature*, such as the mentioned conjecture that the excess of "artificial" information flows could be interfering with the "natural" information flows and the bonding structures of social life. As we will propose herein, a new understanding of the "natural information flows" as they prototypically occur in living beings—even in the simplest cells—could provide a sound basis for discussing the most general problems of the new science.

There is an essential point about the biological intertwining between communication flows, almost immaterial, and stuff flows related to self-production processes. Around the distinction and conjunction between communication and self-production flows revolve some of the most important and general characteristics of informational entities: cells, organisms, brains, societies. It is an approach that can be easily conducted to a variety of empirical, comparative studies and applied research themes. By translating the idea of the information flows to the complex societies of today, a very different panorama may be obtained about the informational limitations of the individual, the complexity growth of social organizations, and the social use of knowledge. We will deal with these matters in the sections that follow. In order to initiate properly the discussion we must first draft the new conceptions proposed about information.

2. Some New Conceptions on Information

Every social crisis is also a knowledge crisis—the fact is that none of the existing disciplines provides a coherent understanding of the information flows in a complex organization or society. As already argued, it may well be the case that with the acceleration of artificial information flows around individuals, a progressive interference occurs with regard to the natural information flows and the bonding structures of social life, causing unknown new "forces" to stress and fracture basic social organizations. No matter the terms used to approach these phenomena, only a vague scientific understanding may be gained for the time being.

Information Science should be the fundamental theoretical-empirical discipline studying the way informational entities "exist" and how they organize their permanence amidst endless flows and exchanges with the environment. The narrow conceptions behind most approaches to information are predisposed to neglect that informational entities do not only communicate, they must self-construct as well. Information and complexity studies have traditionally fallen into a dichotomy: emphasis in formal communicative aspects along the Shannonian or semiotic cultures (sources, channels,

messages, codes, signals, meaning, *etc.*), *versus* emphasis in physical and engineering aspects of self-organization, dissipative systems, energy-matter-entropy flows, transportation, control, self-production. However, all bio-social entities around us are endlessly caught in both communicational and productive flows, always interrelated.

Let us underscore that important fact: all biological systems, and all societies, are organized by putting together networks of communication and networks of self-production processes, always mixing them, interconnecting them. If living beings scan their environment and communicate with each other, it is in order to fulfill their needs and advance their life cycles. “Reading” the environment becomes prior to “eating” it (*i.e.*, whatever is ingested from the outside has been previously detected and inspected by the sensory systems), and it is obvious that the respective energetic and communicational inputs are treated quite differently. Herein, we will explain in detail how both kinds of openness are organized—intertwined—in some of the simplest living organisms: prokaryotic cells. Further, we will argue that the meticulous integration of both kinds of processes is precisely the only way to arrange a complex organization, either molecular, or cellular-organismic, or social. Communicative flows have to infiltrate the material flows of self-production processes and guide them, and also the vice versa.

Whether the concept of *flow* should accompany information, either referred to the communication processes, almost immaterial, or to the material stuff of self-production processes, becomes another aspect of the debate [14]. The time scale considered represents an essential factor. Regarding communication, the classical Shannonian notion was factually embracing the flow idea, along the *source* and *channel* metaphoric terms; discrete messages are belonging to a continuum of communication, and there is no obstacle to consider them into a communication flow in a larger time scale, which may be compartmented into discrete elements and sequences at shorter scales. On the self-production side, the informational entity becomes an open system, out from equilibrium, with necessary energy-entropy-matter flows. The self-production process is based in those flows of stuff—ions, nutrients, affordances of the ecological niches, goods and services, *etc.*—which must be looked for, and anticipated, usually following optimized search strategies [14,15]. They are easily detected and controlled by anticipatory information from the faster and cheaper communication flows. Orchestrating the mutual congruence between these two realms is critical for the viability of informational entities.

There appears a frequent commonality of forms between communication flows and material flows—manifested in the structures that support them. Often both kinds of systems are displaying fractal forms, derived from the necessity to cover a region of space and to transport the affordances of both the material and the communicational to a center [14]. Given their interpenetration, communication and self-production flows appear as spatially catching each other. In general, it is the “infostructure” that dominates the “infrastructure”: fast and cheap information provides guidance and control over slower and more expensive matter. In philosophical terms, it is like the difference between mind and matter, brain and body, software and hardware, controller and controlled, *etc.* Whatever the point of view, the whole theme is fiendishly complex; adequately cohering and integrating the communication and self-production dimensions is still in its infancy. An important line of thought that connects with some of the ideas herein discussed may be found in T.W. Deacon’s work “Incomplete Nature” [16]; also in A. Bejan’s engineering work on flow systems [14]; this approach might be connected with M. Burgin’s theoretical unification too [17].

Very briefly, we have to deal with the conception of information itself. The famous *what is information* question has been debated endless times, and hundreds of information definitions have been produced just in the last decades, most of them anthropocentric [18]. See for instance the proceedings of the successive FIS and CoLIS conferences, or the recent compilation in the special issue of *Information* [19]. Herein, we are restricting ourselves to the quest of a bio-inspired notion of information that could also be extended partially to human and social realms. A non-trivial problem for anthropocentric approaches to information is that whatever kind of phenomena the human observer is attuned to, or is expecting to receive, it might be considered as a legitimate form of information and then susceptible to being generalized and defined conceptually. Actually, every discipline becomes an artificial ordering imposed on information and knowledge, following rigorous formal and experimental procedures, and can legitimately establish or define its own version of what is information.

Scientifically, there cannot be a universal definition of information for the same reason that nowadays there are hundreds of *different* scientific disciplines [20–22]. Every separate realm of knowledge has legitimately distinguished a separate form of information. Only ideal observers could link all of them—never the real observers. However, consensual notions might be established with more or less success, width of application, and rigor in order to cover and interconnect different fields or domains of experience.

Although a good working notion might be crafted, universality is precluded. There will always be some other fields in which even the best established notion will not work appropriately; all the more when explicit conditions on what should count as *subject*, *object*, and *coupling mode* in the information phenomena have to enter into the game too—and may be changed on the spot along highly volatile communicative interactions.

That a phenomenon does not yield to rigorous conceptual definition does not mean its scientific lack of value or that it cannot be appropriately quantified. That is precisely what happens with two of the most fundamental categories of physics: space and time. We could paraphrase St. Augustine's famous question about time: *What then is information? If no one asks of me, I know; if I wish to explain to him who asks, I know not.*

Thereafter, if we advocate undefinability, it 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. It does not follow that a consensus notion could not be established. The achievement of a "winning" notion getting the highest marks in most fields is quite possible and highly recommended (actually we are going to propose an interesting candidate right in the next paragraph): implicitly, that is what most of the proponents of information definitions have in mind.

The tentative notion of information as "*distinction on the adjacent*" has been advocated by the present authors [2,23]. The *distinction* term refers to the capabilities of the subject or informational entity engaged in the communication process: what "receiving the information" means in some restricted communication logics. It is about how the informational entity may create preliminary streams of relationships associated to the impinging signaling flows. A scheme of distinctional logic based on multidimensional partitions was discussed somewhere else [24].

The *adjacent* term refers to the physical contact to be achieved, and the need of counting with sensory elements or with excitatory surfaces to be impinged upon by the incoming communication signals. Increasing the adjacency, extending the territory covered by the communication processes is a formidable drive of biological evolution: cellular pili, flagella, cilia, arborizations of axons and dendrites, the neuronal multiplicity of sensors and receptors, specialized maps, sensing modalities, *etc.* Not to speak about the cultural artifacts, means of communication, and scientific-technological apparatuses that transcend the immediate adjacency of subjects in the complex information flows of contemporary societies. By transcending the limits of immediate space-time adjacency (and memory), subjects may perform a myriad of further distinctions and cognitive operations.

Actually, it is a triad of entangled concepts—information, knowledge, intelligence—upon which the nucleus of a future information science has to be properly assembled. This view is in agreement with the basic outline presented by Y.X. Zhong [25]; but there is some disagreement in the sense that nowadays these concepts cannot be elucidated from any single discipline, either philosophical, or computational, or scientific. As already said, they form part of the very nucleus of information science itself, and the procedures to abstract them should essentially include *the comparative study among canonical informational entities*: cells, nervous systems, social organizations. This “comparative” point of view, although absent in the mainstream literature, has already been initiated by the authors in some previous papers [20,26].

Continuing with that comparative approach, the main novelty of this paper is that we connect the idea of flow with both metabolic and signaling exchanges of the living cell, finding and highlighting the essential processing difference between them. Thus, in the following section (the third), we will approach with some detail the way the living cell organizes its “real” self-production processes, and how it communicates with the external. Although the former processes are relatively well known (either as metabolism or as gene expression), the latter, performed by the specialized “cellular signalling system”, are widely disregarded and even ignored by molecular-biological researchers themselves.

We think that the utterly different way in which both kinds of input-flows are treated within living cells has a universal informational significance. Thereby, we will discuss how the organization of processes by the cellular system in the interconnection between communication and self-production provides a new bio-inspired notion on the triad information-meaning-knowledge. Subsequently, in Section 4, we will make a bold leap: comparing the cellular handling of knowledge, through protein (DNA) domain recombination, with the way modern societies handle the multi-disciplinary processes of knowledge recombination. A new term, scientomics, will be proposed.

3. The Informational Organization of the Cell: Communication and Self-Production

3.1. The Cell as an Informational Model System

From an information science point of view, the living cell is an astonishing system: it incorporates the highest trove of informational phenomena that one can think of at the molecular scale. It is a micro-world teeming up with millions of specific molecular recognition events, genetic codes, transcription and translation processes, molecular machines and self-assembling complexes, signaling systems, messengers, transducers, second messengers, regulators, effectors, connectivity networks, interferences, *etc.*

Conspicuously, the information metaphor has become the natural way of talking about biomolecular phenomena, almost from the very beginning of molecular biology, and even more along the current *bioinformatic* and “*omic*”—genomics, proteomics, transcriptomics, metabolomics—whole revolution [27].

Today, the great advantage fuelling the expansion of the bio-information paradigm is that cellular information processes may be defined almost to completion at the molecular scale—at least in the case of the simplest cells. Contemporary biological research is factually answering some of the most poignant questions associated with traditionally ill-defined, anthropocentric concepts such as information, knowledge, and intelligence, and is providing new tools to overcome the classical limitations of information theory and other formal procedures applied to the organization of life. The way living cells self-produce, communicate, and collectively organize multi-cell systems becomes a paradigmatic case of informational relationships involving an adaptation to the environment, which is knowledge dependent.

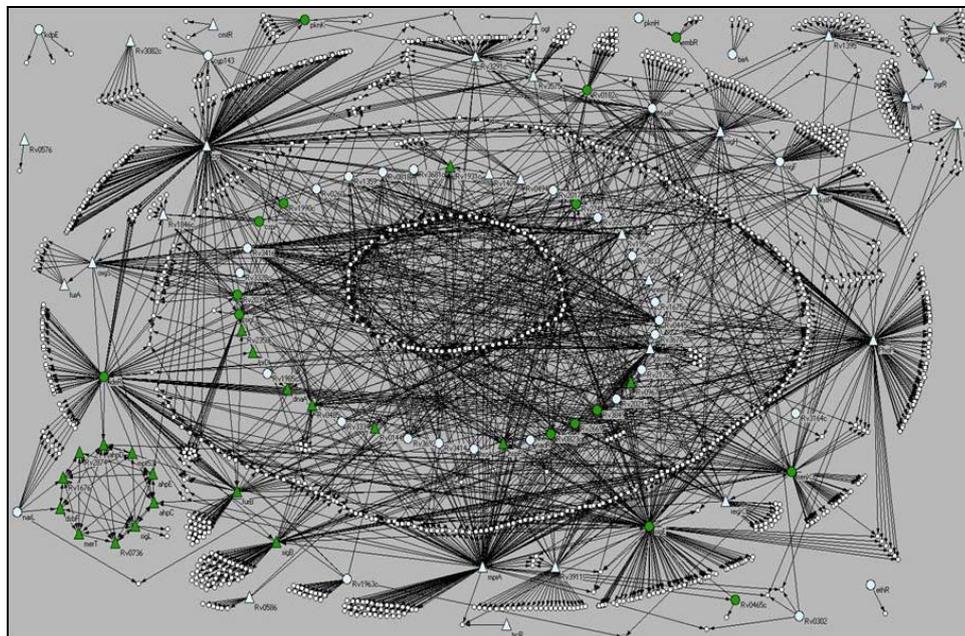
3.2. Cellular Self-Production

Essentially, the cellular game is about an informational problem-solving dynamics applied to self-production of the own structures, implying both synthesis and degradation. It is performed both in physiological time and in evolutionary time by a “network society” of specialized enzyme and protein agents, continuously exchanging information about their specific activities thanks to the especial solvent properties of the water matrix. Enzymes and proteins are “nanomolecular processors” endowed with a peculiar modular structure, based in *protein domains*; they are synthesized (and also degraded) out from the sequential information of the DNA and RNA “data bases”, which in their turn are incessantly subject to evolutionary combinatorial games [28].

Traditionally most studies have focused on the expression of individual genes and not on the overall network and systemic instances of control. Currently, however, massive transcriptional regulatory networks are built for different prokaryotic microorganisms and eukaryotic cellular functions and specialized cell-types. As an instance of such networks, the authors’ research team has cooperated in compiling a large-scale *M. tuberculosis* transcriptional regulatory network, which has been built upon a previously published TR network [29], the largest to date, with the further addition of different kinds of resources pertaining to publicly available sources: DNA microarrays, operons, orthology approaches, and synthetic biology experiments [30,31] (see Figure 1).

The 1400 network nodes represented all correspond to specific genes of *M. tuberculosis* and their protein products, while the 2304 links correspond to gene expression regulatory interactions by 94 transcription factors. The network shows a clear organization in structural levels that correspond with the complex functions and life-cycle stages of this highly sophisticated pathogen. Overall, the genome of the bacillus contains more than 4000 genes and close to 190 transcription factors. In general, an increased number of transcription factors per genome translates into greater genetic network connectivity, which is correlated with increased complexity of the microorganism structures and life cycle [32].

Figure 1. The Transcriptional-Regulatory (ETR) Network of *M. tuberculosis*. Nodes represent *Mt*'s genes, and links represent their regulatory interactions. Approximately 35% of the genome is covered by this network (Modified from [30]).



3.3. Cellular Communication

By itself, the transcription network is “blind”. It needs the injection of further adaptive capability to respond to environmental demands. This is done by means of signaling guidance, so to express the genetic circuits in response to relevant happenstances of the environment or from within the cell. 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*. It is really the organ of cellular communication.

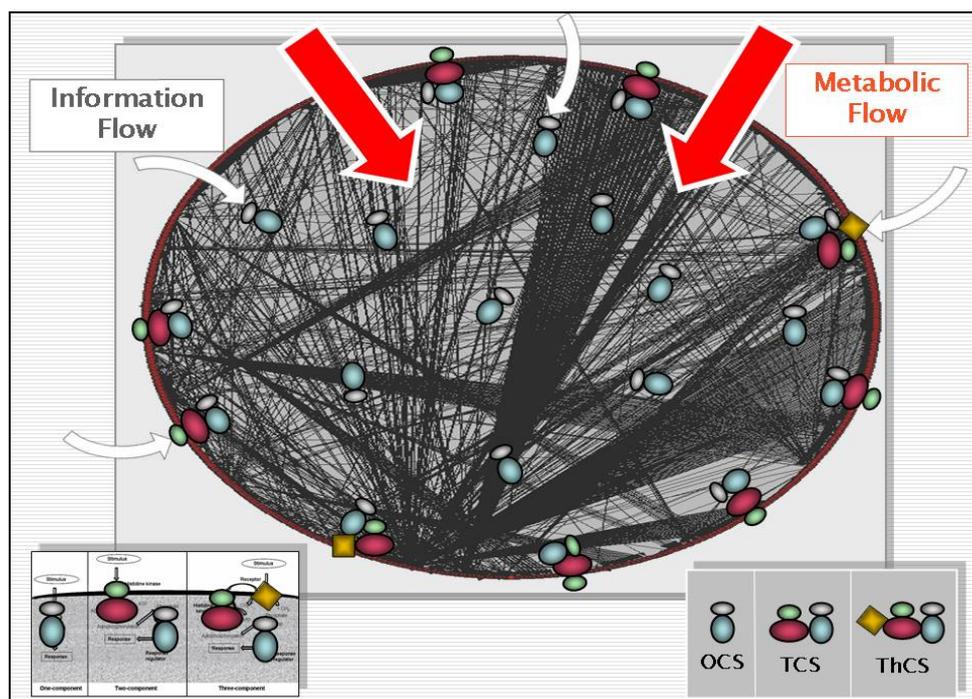
In prokaryotes, a variety of molecular systems are involved in signaling duties, ranging from simple transcription-sensory regulators (a single protein comprising two domains), such as the well-known *embR*, *alkAorfurB*, to those systems of multiple components and interconnected pathways that regulate key stages of the cell cycle, such as latency, pathogenesis, replication, and dispersion. Essential molecular components of the environment are continuously scanned by the signaling system: nutrients, ions, metals, peptides, amino acids, *etc.* Once important molecular presences are detected, the system activates gene expression programs or directly induces changes at the membrane or in the cytoplasm. These continuous communication flows are different from the metabolic flows. See in Figure 2 how the molecules participating are recognized as signals and not processed as metabolites.

The case of lactose may illustrate what we mean. François Jacob and Jacques Monod were the first scientists to discover the basic mechanisms involved in lactose metabolism and regulation, what is called the “lac operon” [33]. Metabolically, the bacterium first choice is glucose, but if the concentration of that substance is low and molecules of lactose are detected by an inner receptor (LacI), then the whole lactose operon is expressed, with the consequence that two specialized enzymes to degrade lactose enter into action as well as a third partner—a permease—to import it more easily

across the plasma membrane. Further mechanisms make sure that these genes are turned off if lactose is not present in the environment or if glucose concentration increases. The essential point is that lactose is treated first as a signal, specifically received by a “one component system” (LacI), and then, after expression of the lac operon, it is treated metabolically by the other three enzymes: transported through the membrane by the permease, and processed by the beta-galactosidase and the transacetylase. So, lactose is “detected” first as a signal, and thereafter it is “eaten” as a metabolite. Similar happenstances may occur for several dozen substances (nutrients, metals, signaling peptides, *etc.*) Figure 2 highlights the difference in kind between faint communication flows, mediated by specific receptor binding, and bulk metabolic flows, which are actively introduced by permeases, transporters, pumps, *etc.*

A basic taxonomy of bacterial signaling systems was proposed somewhere else [34], which was centered on “the 1-2-3 scheme” (see Figure 2). In the concrete case of *M. tuberculosis*, the number of different signaling pathways is close to one hundred, the majority belonging to the “one component systems” class, with one dozen “two component systems” class, and around another dozen of miscellaneous classes. Every signaling pathway may be present (as for the number of acting molecules) within a range in between one or two orders of magnitude each, so from a few molecules to a few hundred.

Figure 2. Metabolic flows and signaling or communication flows are shown respectively as thick (red) arrows and as thinner (white) ones. The representation highlights the difference in kind between faint communication flows, mediated by receptor binding, and bulk metabolic flows, which are actively introduced by permeases, transporters, and pumps. The figure also shows the three characteristic signaling pathways developed by prokaryotes: One Component Systems (OCS), Two Component Systems (TCS), and Three Component Systems (ThCS). These three different options imply very different information processing capabilities and metabolic costs (see Marijuán *et al.* [34]).



In eukaryotes, the signaling system comprises many hundreds of different classes of dedicated molecular agents (receptors, ion channels, transducers, amplification cascades, second messengers, intermediate effectors, final effectors) that can be arranged differently in each tissue. Particularly throughout the very fast changes in second messenger concentrations, an integrated perspective (measurement) of the different internal and external influences—all the communication flows at play—is obtained within the cell, and is subsequently passed towards intermediate chains and the final effectors.

The flows of information crossing throughout the receptors, most of them informing on the presence of metabolic items in the environment, are systematically transformed in variations of the self-producing structure of the cell [26,27]. Thus, the living cell may systematically respond to the signals flowing from the environment, and produce the “meaning” they imply, by letting the signals themselves interfere with the ongoing molecular dynamics of the cellular self-production flows.

Therefore, *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 would correspond to second messengers and the 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 [35].

Most of that complexity growth has been built by tinkering upon multi-domain enzymes and proteins, so that primary functional codes or addresses and secondary addresses regarding the circumstances of the primary functions have been put together (though often in separate domains) onto the same DNA memory bank. By means of bioinformatic tools, one can track down how the different combinations of protein domain families have been progressively formed within genes, generating new protein domains and new gene families in a sort of “Big Bang” of protein evolution, from the early forms of life to the more modern genera. Very old domains can be visualized as they have interpenetrated and recombined with recent domains within more complex proteins, following prokaryote horizontal gene transfer as well as genetic recombinations of all kinds in both prokaryotes and eukaryotes, systemically putting into action more efficient genomes with improved sets of protein domain functionalities.

The existing protein domains coded into the genomic DNA and their combinatory processes may be seen as the *stock of knowledge* of each species, and globally, as the knowledge repository of the biosphere as a whole. Genomes are continuously in the making, self-adapting and trying new knowledge solutions for each individual species in a continuous existential answer to the selective demands posed by every particular niche.

As we have already said, the swapping or recombination between protein domains becomes the main adaptive tool of individual genomes. In this light, *knowledge recombination* becomes the essential cognizing strategy of the global biosphere, a molecular based *noosphere* actually. Then, a poignant question is: To what extent could human societies partake of this recombination strategy at their own societal level? In Sections 4 and 5, we are going to discuss abstract similarities in the way living cells and complex societies deal with the phenomenon of knowledge—the “society of enzymes” *versus* the “society of individuals”. We will make a parallel between the cellular recombination of protein domains and the social dynamics of multidisciplinary recombination. Then, the *scientomics*

term will be proposed not only as a new conceptual construct corresponding to that parallel, but also as a feasible project germane to recent realizations in *culturomics*.

4. The Social Recombination of Knowledge

Approaching science itself as a composite informational construction and particularly as *knowledge recombination* looks feasible [36]. We can quote from W.B. Arthur [37], in his recent approach to the nature of technological change, which is so close to the dynamics of change in science itself: “Technologies therefore share common ancestries and combine, morph, and combine again to create further technologies. Technology evolves much as a coral reef builds itself from activities of small organisms—it creates itself from itself; all technologies are descended from earlier technologies...”

The natural division of work within scientific communities seems to reflect the presence of knowledge recombination processes: the need of specialized disciplines and the reliance on paradigms, the fracture and emergence of new fields, the systematic increase in the number of disciplines during recent centuries, the clusters and citation networking structures within scientific publications, *etc.*

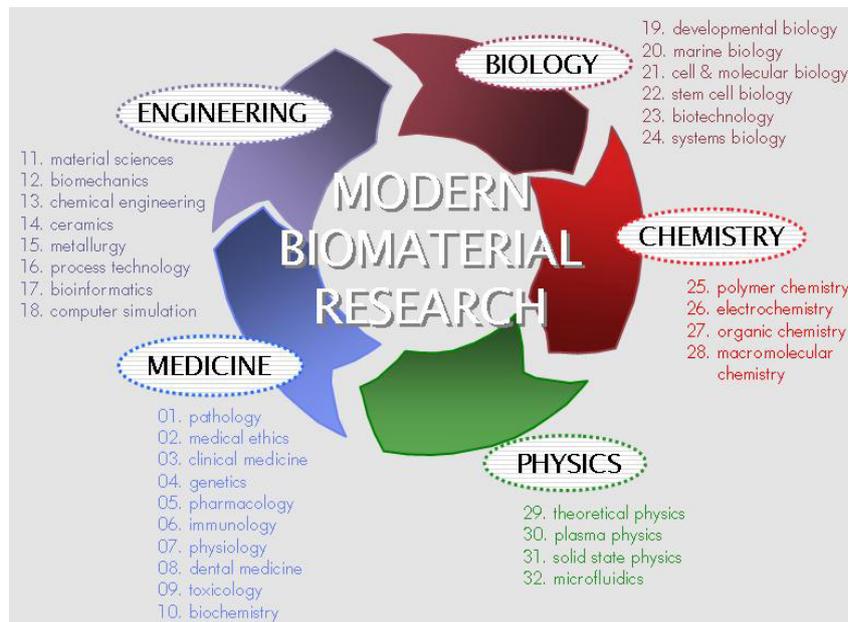
Disciplines, rather than being isolated fields, are continuously mixing and rearranging their contents, *recombining* them, for the sake of the problems they have to solve, and factually giving birth to successive generations of inter-disciplines (e.g., information-physics, physical chemistry, biophysics, biochemistry, bioenergetics, bioengineering, socio-physics, sociobiology, psycho-sociology, neuro-psychiatry, socio-information, *etc.*).

See in Figure 3 how research in a very advanced field—biomaterial research—is contemplated by one of its leading practitioners [38]. The crowding of subdisciplines and specialties is remarkable: up to 32 different ones are listed. It could remind the domain accretion of some large protein of late eukaryotic evolution, as the figure itself suggests by representing specialties in a common circle of domains. Like in the evolutionary process, it makes sense that the most advanced scientific explorations incorporate larger troves of disciplines and specializations. That is particularly true in biomedical research, which has become one of the central and most complex scientific hubs of today.

As Figure 3 suggests, all major research fields have to be surrounded by a “cloud” of disciplines in order to convey the necessary scientific-technologic knowledge. We propose the term “*domain of knowledge*” to the particular collegiums of disciplines surrounding every major research field and potentially contributing to its knowledge recombination processes. It is clear, as in the case of Figure 3, that only some specialties or subdisciplines of each major science are actively involved in the exchange processes. Even at the level of a concrete subdiscipline, the real granularity of the exchanges concerns “modules of specialization” that incorporate theoretical and practical knowledge. Research fields are but niches of opportunity that attract expertise of different disciplines and organize new domains of knowledge; if the research is successful and expectations are fulfilled, new disciplines of inter-multi-disciplinary nature will arise subtended by a new, *ad hoc* research community [36].

It has been estimated that after the industrial revolution, the number of scientists and research fields has approximately doubled with each passing generation [39]. At the height of the 1990s, it has been estimated that more than 8000 research topics or fields were supported by approximately 4000 disciplines [21]. To the extent in which those estimates are cogent, the number of disciplines could have increased to 5000–6000 nowadays, supporting around 10,000 research fields.

Figure 3. Disciplines involved in modern biomaterial research. The representation is based on the description made by bioengineer J. Kirkpatrick [39] (Modified from [20]).



Scientometric studies have already provided rigorous and useful “science maps”, during the last three decades; they were based on citation structures and have grown enormously—keeping pace with Moore’s Law—up to impressive dimensions and multidisciplinary coverage. Representations derived from some million papers covering almost all research fields have been obtained recently [40]. Notwithstanding that computational prowess, generative hypotheses on the overall science and society relationships underlying that unstoppable scientific growth continue to be in short supply.

5. The “Scientomic” Approach

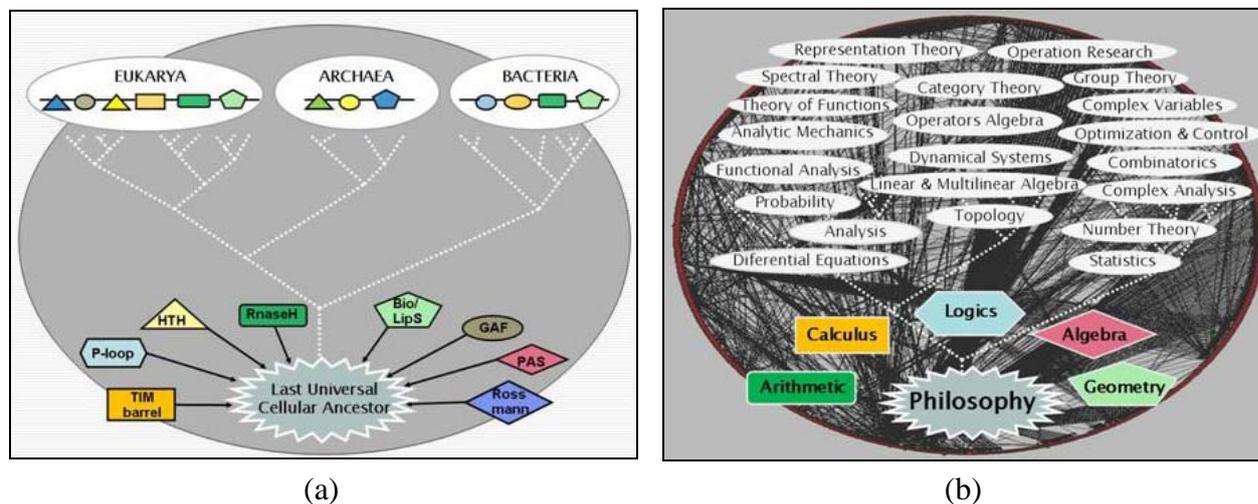
From the point of view of information science, in the same way that philosophies of science, history of science, and psychology and sociology of science have already been developed, we would also need a genuine informational approach to sciences’ generative processes: *scientomics* [36].

In practical terms, at the time being there might be sufficient scope to compare the biological evolution of DNA codes of protein domains and the social-historical evolution of scientific disciplinary contents. Do cognitive “modules” exist within disciplines that travel to other disciplines and generate new fields there? If so, could the combinatory processes in both realms be interrelated? See Figure 4, which illustrates the parallel between the Big Bang of protein domains and the explosive growth of the sciences.

Culturomics might have already paved part of the way. Borrowing the main concepts and techniques from evolutionary biology, J.B. Michel and E.L. Aiden were able to track the growth, change, and decline of the most meaningful *published words* during the last centuries [41]. The new term they have coined, *culturomics*, means the application of “genomic techniques” of high-throughput data collection and analysis to the study of human culture, as sampled in a vast mapping of words from a corpus of digitized books, containing about 4% of all printed books ever published. Further sources might be incorporated to the culturomic stock: newspapers, manuscripts, maps, artwork, *etc.* Analysis

of this corpus enables a new qualitative and quantitative investigation of cultural trends, social and political influences, fashions, and all sort of cultural phenomena.

Figure 4. Parallel between recombination events in the evolution of the protein universe and in the evolution of the sciences. (a) The “big bang” of protein universe; (b) Subdisciplines of mathematics (Modified from [20]).



Thus, the knowledge recombination hypothesis applied to the historical evolution of science might be considered in scientomic terms, as an evolutionary quest on the combinatory activity of disciplinary modules or domains of theoretical-practical knowledge travelling to other disciplines and changing there the local textures of knowledge, altering the regional maps of science, and the whole complexion of the world of knowledge at large. In other words, influential modules such as Euclidian geometry, Newtonian mechanics, differential equations, genetics, and so on (and a multitude of other minor modules), would have generated the history of sciences, not only “developmentally” inside their own fields, but even more “combinatorially”, propelling the multidisciplinary evolution and cross-fertilization among scientific disciplines.

In terms of education science, something similar would happen, for an abridged recapitulation resembling Haeckel’s law seems to be taking place in the ontogenetic development of an individual’s knowledge, which somehow recapitulates the fundamentals of the social acquisition of knowledge along history.

Scientomics, as we are suggesting, could be an important future task for the consolidation of information science, as well as a multidisciplinary research-project running in parallel to current achievements of culturomics in the cultural realm, though pointing to some more ambitious epistemic goals. Indeed, the creation of a proficient “*scientomics*” new field would help to make sense of the historical processes of science, and of human knowledge in action.

From different disciplinary sources, pioneering authors have already recognized the multi-disciplinary implications of *knowledge recombination* [36]. At the philosophical and scientific scale, W. Ostwald’s “Kombinatorik” [42] was notoriously applied not only to nature but also to knowledge organization and creativity processes [43]. In the social realm, J.C. Scott [10] has discussed how the limitations of human expertise are forcing cognizing individuals to “play” recombination

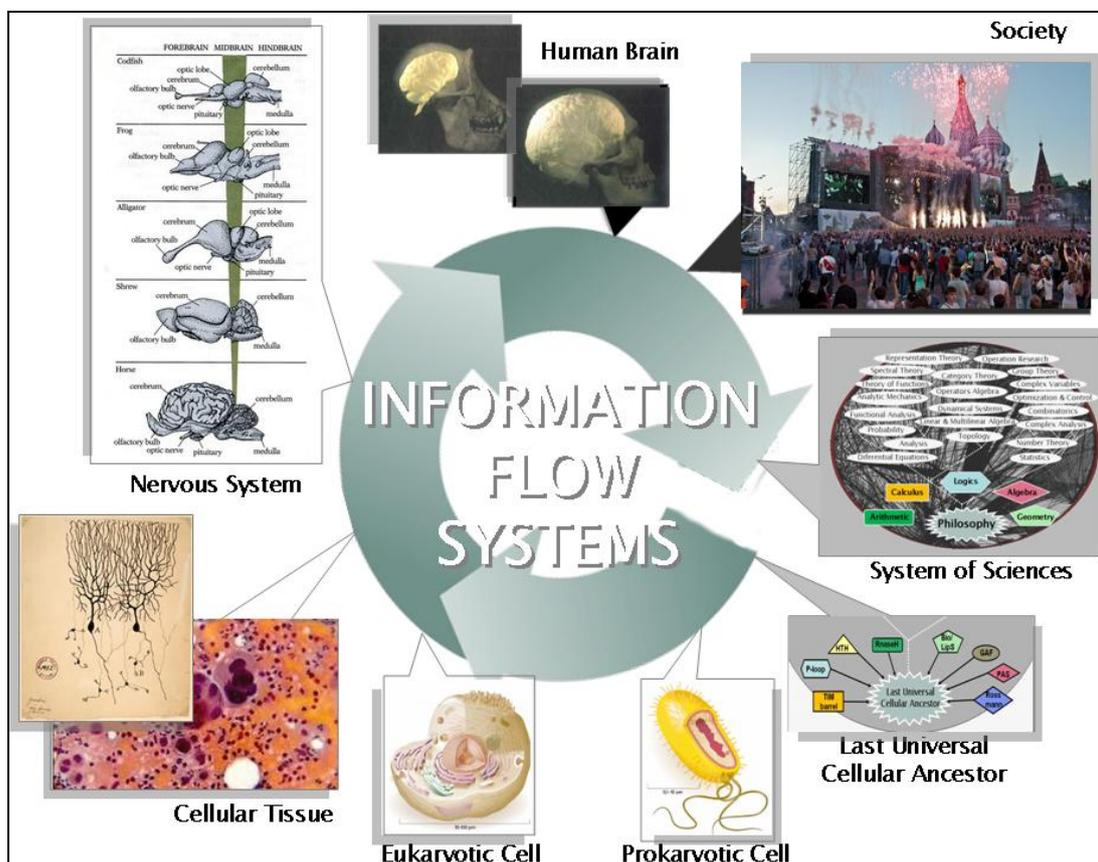
games. In the technological realm, historians have already been aware, at least since C. Gilfillan [44], that innovations stem from combinations of what is already known. More recently, the work of W.B. Arthur [37] is intensely dealing with the evolution of technological systems through the social organization of “knowledge recombination processes”. In the history of science, scholars of interdisciplinarity have been progressively aware of the recombination phenomenon in the relationship between disciplines [45]; a number of new ideas and projects have also been developed during the last two decades [21]. The present approach may also be related to deep theories of science [46], and to the sociology and psychology of science [47,48]. The ideas that have been presented here, which may be considered as germane or as a rough continuation of some of these previous works, are now drafted from an emerging bioinformational perspective—they put together a new recombinatory “scientomic” sense that may be applied upon the inter-multi-pluri-trans-disciplinary games within the sciences.

6. Final Comments: The Information Flow in Complex Organizations

Too many important topics related with information have been addressed in this paper. Perhaps, the most essential idea concerns the evolution of information flow systems—how life at all scales is based on the coupling between self-production and communication. The adage says that “a picture is worth a thousand words”. However, there is also the opposite, as the Chinese say, “1001 words are worth more than a picture”. Therefore, these final comments will attempt the synthesis in both ways.

The image of Figure 5 speaks by itself; it is really a visual synthesis conducting “from genomics to scientomics”.

Figure 5. Information Flow Systems: a visual synthesis.



Continuing with the conceptual summary, we are witnessing during last two decades how an increasing variety of research practices are converging into information science's territories potentially. 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 emerge may not only depend on successful discussions about the philosophy of information; perhaps, the most important adjustment to make is about framing a new perspective or way of thinking within information science itself, interrelating the research practices across the fundamental empirical and theoretical areas of the new science. Tentatively, the new way of thinking could be based on the considerations that follow.

- Rather than continuing with the discussions on how to define the information concept, a conceptual shift from information to “informational entities” is advocated. It implies a new type of scientific observer, a consensus notion on information, and an essentially empirical orientation for the new science.
- Achieving a new understanding on the natural flows of information, basically making sense of how the flows of self-production and the communication flows become intertwined in the workings of the informational entity. It can be extended to a set of canonical informational entities: cells, multi-cell organisms, nervous systems, and societies. The simplest cells—prokaryotes—are the paradigmatic model.
- Placing the emphasis on *applied* informational themes, trying to cooperate in useful research topics with other disciplines and paying attention to relevant social problems; and also contributing to analyze the social pitfalls of contemporary information technologies, both in the mental (stress, learning, social bonding, tunnel vision) and in the physical (energy burden, waste, pollution, pilfering of resources).
- Providing a new vantage point to look at the whole social dynamics of scientific knowledge, an essential problem of today's “information societies”—a Babel Tower of more than 6000 scientific and technological disciplines, where classical information (library) science is helpless in providing any interesting guidance to society.
- Contributing to more sophisticate cultural syntheses, needed in these global “instantaneous” times of continuous information overload and critical demands beyond the bounds of social representation systems. New structures of social bonding, of social problem solving, and of collective intelligence should appear for the sustainable societies of tomorrow.

The “invisible hand” of information has been the great shaper of all the complexity evolved by the different informational realms around, either in the biomolecular and neuronal, or in the social and economical. Information can only be detected in process, symbolically flowing amidst the different agency realms and contributing to modify their self-producing structures adaptively. There are too many conceptual obstacles to grasp this most protean entity, now flowing more and more rapidly amidst individuals and societies, and overwhelming our limited resources and capabilities of individual and collective intelligence. As the prescient Spanish philosopher J. Ortega y Gasset [49] would have put it: mastering information science has become “the challenge of our times”—*el tema de nuestro tiempo!*

Author Contributions

Raquel del Moral contributed to this work with the biological and the scientomic discussion. Jorge Navarro also contributed to the biological part focusing in prokaryotic signaling. Pedro C. Marijuán oriented the information science general discussion and the scientomics approach.

Acknowledgment

The authors acknowledge to the organizers of the Moscow FIS Conference for the opportunity to participate and discuss these views.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Conrad, M. Cross-scale information processing in evolution, development and intelligence. *BioSystems* **1996**, *38*, 97–109.
2. Marijuán, P.C. First conference on foundations of information science: From computers and quantum physics, to cells, nervous systems, and societies. *BioSystems* **1996**, *38*, 87–96.
3. Home page of Foundation of Information Science. Available online: <http://infoscience-fis.unizar.es> (accessed on 26 January 2014).
4. Hjørland, B. Information Science and its Core Concepts: Levels of Disagreement. In *Theories of Information, Communication and Knowledge: A Multidisciplinary Approach*; Studies in History and Philosophy of Science, Volume 34; Ibekwe-SanJuan, F., Dousa, T.M., Eds.; Springer: Dordrecht, The Netherlands, 2014; pp. 205–235.
5. Rosenbloom, P.S. *On Computing: The Fourth Great Scientific Domain*; The MIT Press: Cambridge, MA, USA, 2013.
6. Hobart, M.E.; Schiffman, Z.S. *Information Ages*; The John Hopkins University Press: Baltimore, MD, USA, 1998.
7. Lanham, R.A. *The Economics of Attention*; The University Chicago Press: Chicago, IL, USA, 2006.
8. O'Donnell, J.J. *Avatars of the Word*; Harvard University Press: Cambridge, MA, USA, 1998.
9. Poe, M.T. *A History of Communications: Media and Society from the Evolution of Speech to the Internet*; Cambridge University Press: New York, NY, USA, 2011.
10. Scott, J.C. *Seeing Like a State*; Yale University Press: New Haven, CT, USA, 1998.
11. Wright, A. *Glut: Mastering Information through the Ages*; Joseph Henry Press: Washington, DC, USA, 2007.
12. Ortega y Gasset, J. *La rebelión de las masas* (in Spanish); Tecnos: Madrid, Spain, 2003.
13. Helbing, D. Globally networked risks and how to respond. *Nature* **2013**, *497*, 51–59.
14. Bejan, A.; Peder, J. *Design in nature*; Doubleday: New York, NY, USA, 2012.
15. Holland, J.H. *Signals and Boundaries: Building Blocks for Complex Adaptive Systems*; The MIT Press: Cambridge, MA, USA, 2012.

16. Deacon, T.W. *Incomplete Nature: How Mind Emerged from Matter*; W.W. Norton & Company: New York, NY, USA, 2012.
17. Burgin, M. *Theory of Information: Fundamentality, Diversity and Unification*; World Scientific Series in Information Studies; World Scientific Publishing: Singapore, Singapore, 2010.
18. Lenski, W. Information: A conceptual investigation. *Information* **2010**, *1*, 74–118.
19. Home Page of Special Issue “Information: Its Different Modes and Its Relation to Meaning”. Available online: http://www.mdpi.com/journal/information/special_issues/information_meaning (accessed on 13 February 2014).
20. Del Moral, R.; González, M.; Navarro, J.; Marijuán, P.C. From genomics to scientomics: expanding the bioinformation paradigm. *Information* **2011**, *2*, 651–671.
21. Klein, J.T. Interdisciplinary and complexity: An evolving relationship. *Complexity* **2004**, *6*, 2–10.
22. Logan, R.K. What is information? Why is it relativistic and what is its relationship to materiality, meaning, and organization? *Information* **2012**, *3*, 68–91.
23. Marijuán, P.C. The advancement of information science: Is a new way of thinking necessary? *TripleC Commun. Capital. Crit.* **2009**, *7*, 369–375.
24. Marijuán, P.C.; Pastor, J.; Villarroel, M. The language of cells: A partitional approach to cell-signaling. *Sym. Cult. Sci.* **1998**, *9*, 383–392.
25. Zhong, Y.X. On information Science: An introduction to Principles of Information Science. In Proceedings of the 4th International Conference on the Foundations of Information Science. Beijing, China, 22 August 2010.
26. Marijuán, P.C. Knowledge recombination: on the informational adaptability of cells, nervous systems, and societies. *Inf. Theories Appl.* **2011**, *18*, 3–15.
27. Marijuán, P.C. Information and life: Towards a biological understanding of informational phenomena. *TripleC Commun. Capital. Crit.* **2004**, *2*, 9–19.
28. Marijuán, P.C. Bioinformation: Untangling the networks of life. *BioSystems* **2002**, *64*, 11–118.
29. Balázsi, G.; Heath, A.P.; Shi, L.; Gennaro, M.L. The temporal response of the Mycobacterium tuberculosis gene regulatory network during growth arrest. *Mol. Syst. Biol.* **2008**, *4*, doi: 10.1038/msb.2008.63.
30. Navarro, J. Transcriptional Regulatory Network of M. tuberculosis: Functional and Signaling Aspects. Master Thesis, Universidad de Zaragoza, Spain, 2010.
31. Navarro, J.; Goñi-Moreno, A.; Marijuán, P.C. Varieties of biological information: A molecular recognition approach to systems biology and bioinformatics. *Inf. Theories Appl.* **2010**, *4*, 56–66.
32. Levine, M.; Tjian, R. Transcription regulation and animal diversity. *Nature* **2003**, *424*, 147–151.
33. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 5th ed.; Garland Science: New York, NY, USA, 2007.
34. Marijuán, P.C.; Navarro, J.; del Moral, R. On prokaryotic intelligence: Strategies for sensing the environment. *BioSystems* **2010**, *99*, 94–103.
35. Marijuán, P.C.; del Moral, R. The Informational Architectures of Biological Complexity. In *Computation, Information, Cognition—The Nexus and The Liminal*; Dodig-Crnkovic, G., Stuart, S., Eds.; Cambridge University Press: Cambridge, UK, 2007.
36. Marijuán, P.C.; del Moral, R.; Navarro, J. Scientomics: An emergent perspective in knowledge organization. *Knowl. Organ.* **2012**, *39*, 153–164.

37. Arthur, W.B. *The Nature of Technology: What It Is and How It Evolves*; Free Press: New York, NY, USA, 2009.
38. Kirkpatrick, J. Multi-Inter-Disciplinarity in Bioengineering, Biomaterials, and Nanomedicine. In Proceedings of the CIBER-BBN: Networking Biomedical Research Centre for Bioengineering, Biomaterials & Nanomedicine, Zaragoza, Spain, 17–18 February 2009.
39. Landes, D.S. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*; W.W. Norton & Company: Brooklyn, NY, USA, 1998.
40. Börner, K. *Atlas of Science: Visualizing What We Know*; The MIT Press: Cambridge, MA, USA, 2010.
41. Michel, J.B.; Shen, Y.K.; Aiden, A.P.; Veres, A.; Gray, M.K.; The Google Books Team; Pickett, J.P.; Hoiberg, D.; Clancy, D.; Norvig, P.; *et al.* Quantitative analysis of culture using millions of digitized books. *Science* **2011**, *331*, 176–182.
42. Ostwald, W. Kombinatorik und Schaffende Phantasie. In *Forschen und Nutzen: Wilhelm Ostwald zur wissenschaftlichen Arbeit* (in German); Akademie-Verlag: Berlin, Germany, 1978; pp. 28–30.
43. Hapke, T. Combinatorics and Order as a Foundation of Creativity, Information Organization and Art in the Work of Wilhelm Ostwald. A Poster Presented at the Conference of Analogous Spaces—Architecture and the Space of Information, Intellect and Action, 15–17 May 2008, Ghent University, Ghent, Belgium.
44. Gilfillan, S.C. *The Sociology of Invention*; Follett: Chicago, IL, USA, 1935.
45. Dogan, M.; Pahre, R. *Creative Marginality: Innovation at the Intersections of Social Sciences*; Westview Press: Boulder, CO, USA, 1990.
46. Kuhn, T. *The Structure of Scientific Revolutions*; The University of Chicago Press: Chicago, IL, USA, 1962.
47. Leydesdorff, L. *A Sociological Theory of Communication: The Self-Organization of the Knowledge-Based Society*; Universal-Publishers: Parkland, FL, USA, 2001.
48. Feist, G.J. *The Psychology of Science and the Origins of the Scientific Mind*; Yale University Press: New Haven, CT, USA, 2006.
49. Ortega y Gasset, J. *El tema de nuestro tiempo* (in Spanish); Tecnos: Madrid, Spain, 2002.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).