General Theory of Information Paves the Way to a Secure, Service-Oriented Internet Connecting People, Things, and Businesses

Mark Burgin
Department of Mathematics
University of California Los Angeles
Los Angeles, CA 90095, USA.
mburgin@math.ucla.edu

Rao Mikkilineni
Ageno School of Business,
Golden Gate University,
San Francisco, CA 94105, USA;
rmikkilineni@ggu.edu

***Abstract*— Two foundational issues are hindering the realization of the full potential of the Internet in connecting people, things, and businesses with the highest levels of stability, safety, security, and performance. First, the symbolic computing model on which the information technologies that support the Internet, is 70 plus years old and is based on Alan Turing's observation of how humans compute numbers using symbols. New insights from the general theory of information (GTI) provide a path to go from symbolic computing to supersymbolic computing based on knowledge structures and operations on them to improve the stability and safety of computing structures. Second, the Internet evolved from a vision that "envisioned a globally interconnected set of computers through which everyone could quickly access data and programs from any site" and the operations and management aspect of a global and complex operational infrastructure required to support stable, safe, secure and highly available services such as scale, performance, and higher-level functionality were add-ons. In this paper, we take the cue from biological systems that manage their stability, security, and resource management to accomplish their goals and apply the tools of GTI to propose and explore a service-oriented Internet architecture that improves the process operation & management in connecting people, things, and businesses with improved safety and security.**

Keywords—Internet, General theory of information, Computing Models, Security, Digital Genome, service-oriented Internet

I. INTRODUCTION

It is not an exaggeration to call today's global economy the Internet economy. On the one hand, since its inception [1, 2], the Internet in a short time changed the way we communicate, collaborate and conduct commerce by connecting people, things, and businesses, and creating conditions for the global economy at scale. On the other hand, the dark side of the Internet [3] has wreaked havoc with spam, malware, hacking, phishing, denial of service attacks, click fraud, the invasion of privacy, defamation, fraud, and violation of digital property rights, etc. The result is the potential and actual damage, in many forms, from harming the national security of nations to inflicting loss of privacy and threatening the security of individuals and their information assets. As Naughton points out [1] p.36, "Asking whether the Net is a good or a bad thing is a waste of time. People once asked the same rhetorical question about electricity and the telephone. A much more interesting question is this: 'What is the Net?" We agree and add another question 'what are the limitations of the current state-of-the-art Internet, and how can we fix them?'

In this paper, we argue that the tools provided by the General theory of information (GTI) allow us to address the security of Internet-based services by filling the gaps in the current state-of-the-art whose foundation is based on a seventy-plus-year-old computing model [4-6]. We discuss the current deficiencies and how to remedy them using the insights about the relationship between information and knowledge, gained from GTI [7-11]. We propose a digital genome-based autopoietic and cognitive service architecture where business processes connect people, and things, and manage themselves even when deployed on a not-so-reliable, and not-so-secure distributed infrastructure. The novel secure service-oriented Internet architecture reuses the current Internet and information technologies using an overlay architecture that is similar to how the mammalian brain introduced higher-level reasoning by repurposing the reptilian cortical columns with a common knowledge representation from the information received through the five senses [12] using the society of genes[13] and the neural networks.

It is necessary to explain the advantages of operational tools, which are suggested in this paper, for the development and maintenance of Internet-based services from the traditional means of information processing used as the operational base for the existing Internet. The traditional information processing systems used now work with symbols and are described by such a theoretical model as a Turing machine. Here we argue that to achieve a higher level of Internet functionality, it is necessary to utilize information processing of structure, which is described by such a theoretical model as a structural machine. It is proved that structural machines are essentially more efficient than the most popular traditional model of computation such as a Turing machine (cf. [14], [15]).

Architectural forms of the Internet are better represented by grid automata in comparison with other theoretical network models such as neural networks, cellular automata, or Petri nets. That is why we suggest using grid automata for the development, control, and maintenance of the Internet and other big computer networks.

In section 2, we discuss the foundational shortcomings of the current information technologies, on which the Internet depends and which do not allow meeting the higher levels of availability, performance, security, and regulatory compliance requirements of the process automation systems that connect people, things, and business processes. In section 3, we focus on the issues of safety and security and identify some ways to improve them taking the cues from how biological systems manage their safety and security with self-regulating processes. In section 4, we use the tools derived from GTI to suggest a new approach to infusing a service-oriented architecture with added autopoietic and cognitive behaviors to current generation information technologies. In section 5, we describe an example use case, a secure, self-regulating video delivery service using a multi-cloud infrastructure. In section 6, we provide some observations that allow delineating several directions for future work both in the theoretical and practical areas.

II. Foundational Gaps in the Current-State-of-the Art

We discuss three foundational issues with the current state-of-the-art information technologies using symbolic, subsymbolic computing models, and the Internet.

A. Limitations of the underlying computing model in dealing with non-functional requirements of distributed computing structures with large fluctuations in either the demand for or the availability of the resources:

On one hand, the success of business process automation using both symbolic and subsymbolic computing combined with ubiquitous access to globally connected computing resources using the Internet has allowed connecting people, things, and businesses at scale, and has made communication, collaboration, and commerce almost at the speed of light. On the other hand, it has also increased the dependence of mission-critical processes on non-functional requirements such as the availability, performance, security, and cost of the computing infrastructure. A process is executed by several distributed software components using computing resources often owned and managed by different providers and the assurance of end-to-end process sustenance with adequate resources, its stability, safety, security, and compliance with global requirements requires a complex layer of additional processes that increase complexity leading to 'who manages the managers' conundrum. Any failure in the system requires information access and analysis from multiple sources which results in a reactive approach to fixing the problems.

The end-to-end process is executed by a structure of distributed software components that are dependent on the infrastructure that provides the resources which are managed by disparate service providers with their own management systems. In essence, the process execution structure behaves like a complex adaptive system that is prone to emergence properties when faced with local fluctuations impacting the infrastructure. For example, if a failure occurs that impacts any one component, action must be taken by external entities to fix the problem.

The concept of the universal Turing machine has allowed us to create general-purpose computers and [14] p. 215 "use them to deterministically model any physical system, of which they are not themselves a part to an arbitrary degree of accuracy. Their logical limits arise when we try to get them to model a part of the world that includes themselves."

In this paper, we discuss how the computer and the computed can be incorporated into the model just as the living organisms do.

B. Integration of knowledge from symbolic and subsymbolic computing structures using supersymbolic computing

Subsymbolic computing with a neural net computing model provides insights into data, but integrating the new knowledge with other processes is cumbersome if not existent. In this paper, we discuss how supersymbolic computing integrates knowledge from both symbolic and sub-symbolic computing structures.

C. Security and safety of services deployed using the Internet

Current state-of-the-art security and safety management of processes deployed using the internet depends on the network, storage, and computing device management which is distributed and provided by several independent operators. The application of security and safety management without end-to-end visibility and control in real-time is prone to be reactive and often too late to react. In this paper, we propose a service-oriented security framework that decouples application security from infrastructure security. GTI provides a framework to address the shortcomings with the addition of autopoietic and cognitive process overlays mimicking living organisms that have developed a mammalian neocortex that manages the system behaviors using the reptilian cortical columns. In the next section, we identify the cues from biological systems to infuse autopoietic and cognitive behaviors into digital computing structures.

III. Lessons from the General Theory of Information

"The single fertilized egg cell develops into a full human being is achieved without a construction manager or architect. The responsibility for the necessary close coordination is shared among the cells as they come into being. It is as though each brick, wire, and pipe in a building knows the entire structure and consults with the neighboring bricks to decide where to place itself." This statement from the book "The Society of Genes" [13] summarizes the power of autopoietic and cognitive processes that biological systems have developed using physical structures that are capable of building themselves, using information received from the five senses, converting into knowledge about themselves and their interactions with the environment, and execute "life" processes that support their sustenance, stability, safety, and optimization of resources in executing their goals.

To achieve this level of functioning in artificial networks, we use the lessons from GTI to design a network architecture aimed at designing and implementing autopoietic and cognitive information processing structures. We discuss how we can utilize this framework to improve digital information processing structures to deploy and manage autopoietic and cognitive applications as services on the Internet.

All e-services and the Internet are based on information acquisition, processing, transmission, and management. Thus, to build efficient e-services and the first-class Internet, it is necessary to have an adequate understanding of information and information processes.

In the general theory of information (GTI), the definition of information in the broad sense is given in the second ontological principle, which has several forms [7].

Ontological Principle O2 (the General Transformation Principle). In a broad sense, information for a system *R* is the potentiality/cause for changes (e.g., formations and transformations) in the system *R* or for prevention of such changes, i.e., for the stability of the system *R*.

Thus, we may understand the information in a broad sense as a capacity (ability or potency) of things, material, as well as mental and abstract, to change other things. Information exists in the form of portions, pieces, or instances of information.

Information in the strict sense is stratified according to the global structure of the world represented by the Existential Triad of the world [8-10], which is composed of the top-level components of the world as a unified whole reflecting the unity of the world. This triadic structure is rooted in the long-standing tradition coming from Plato and Aristotle and consists of three components: the Physical (Material) World, the Mental World, and the World of Structures [8, 9]. The Physical (Material) World represents the physical reality studied by natural and technological sciences, the Mental World encompasses different forms and levels of mentality, and the World of Structures consists of various kinds and types of ideal structures [9].

However, the common usage of the word information does not entail such wide generalizations as the Ontological Principle O2 implies. To define information *per se*, the GTI uses the concept of an infological system IF(*R*) of the system *R* for the information definition. Elements from IF(R) are called infological elements.

Ontological Principle [7] O2a (the Special Transformation Principle). Information in the strict sense or proper information or, simply, information for a system *R*, is the potentiality/cause for changes (e.g., formations and transformations) of the structural infological elements from an infological system IF(*R*) of the system *R* or for prevention of such changes, i.e., for the stability of the system IF(*R*).

According to the Ontological Principles O2 and O2a of the GTI and its additional forms, information plays the same role in the World of Structures as energy plays in the Physical (Material) World.

However, according to the Ontological Representability Principle (Ontological Principle O3) of the GTI, for any portion of the information *I*, there is always a representation *Q* of this portion of information for a system *R* [7]. Often this representation is material, and as a result, being materially represented, information becomes physical. Consequently, a physical representation of information can be treated as the materialization of this information. In addition, the representation of information can be mental projecting its impact on people's mentality.

Moreover, according to the Ontological Embodiment Principle (Ontological Principle O4) of the GTI [7], for any portion of the information *I*, there is always a carrier *C* of this portion of information for a system *R*. This carrier is, as a rule, material, and this, even more, makes information physical. The physical carrier of information can be also treated as the materialization of this information, through which the information influences the material world.

In exploring and designing information processes, researchers come to the following fundamental question. Knowledge and Information – What is the Difference? The general theory of information (GTI) gives a comprehensive answer to this question, which is explained below.

The Ontological Principle O2a implies that information is not of the same kind of essence as knowledge and data, which are structures [9]. Although some researchers announce that information is a kind of data, while others claim that information is a kind of knowledge, from the scientific perspective, it is more efficient to treat information as an essence that has a different nature because other terms represent various kinds of knowledge and information. If we take that matter is the name for all substances as opposed to energy and the vacuum, then relations between information and knowledge bring us to the Knowledge-Information-Matter-Energy (KIME) Square as shown in Figure 1.

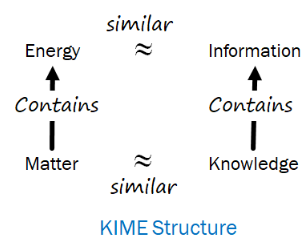

Figure 1. The Knowledge-Information-Matter-Energy (KIME) Square

In other words, we have the following principle [8]:

Information is related to knowledge as energy is related to the matter.

Energy has the potential to create, preserve or modify material structures, and information has the potential to create, preserve or modify knowledge structures. Energy and matter belong to the physical world, while information and knowledge belong to the world of ideal structures and are represented in the mental world.

IV. Hierarchical Autopoietic Operational Networks (HAON)

To achieve a high level of efficiency, reliability, flexibility, and safety, it is practical to build e-services as hierarchical autopoietic operational networks (HAON).

On the first level, HAON has individual machines and local computer networks as well as program systems and Internet services. These machines/services can be modeled by different kinds of abstract automata – finite automata, Turing machines, inductive Turing machines, or structural machines.

On the second level, HAON has the grid (network) array the nodes of which are elements from the first level while links provide channels of interaction between the nodes.

On the third level, HAON has a structural machine that works with the grid (network) of the second level. This machine creates, modifies, or deletes nodes and links, as well as modifies, manages, and controls the whole grid (network) of the second level.

The structural machine on the third level is modeled and mathematically described by an abstract structural machine, which has the following features [11].

A structural machine *M* works with structures of a given type and has three components:

1. The *control device CM* regulates the state of the machine *M*. This control device can be centralized determining the state of the whole machine M or distributed when each of its components called unit control devices regulates the state of some part (component) of the machine *M*. The distributed control device can have unit control devices of two types: a *cluster control device* controls a cluster of processors in the structural machine *M* while an *individual control device* controls a single processor in the structural machine *M*.

2. The *processor PM* performs transformation of the processed structures and its actions (operations) depend on the state of the machine *M* and the state of the processed structures. There are two basic types of the processor *PM*: a *localized processor* is a single abstract device (processor unit or unit processor) while a *distributed processor*, which is also called a *total processor*, consists of a system of *unit processors* or *processor units*

3. The *functional space* Sp_M , in which processors work, consists of three components:

   - The *input space* In_M , which contains the input structure, e.g., a word or a graph, or a system of input structures. In a general case, this system (structure) can be finite, potentially infinite, and in the theoretical context, actually infinite.
   - The *output space* Out_M , which contains the output structure, e.g., a word or a graph, or a system of output structures. In a general case, this system (structure) can be finite, potentially infinite and in the theoretical context, actually infinite.
   - The *processing space* PS_M , in which the input structure(s) is transformed into the output structure(s).

We assume that all structures – the input structure, the output structure, and the processed structures – have the same type. Note that in the classical models of computations, such as Turing machines, all these spaces coincide and are represented by one or several tapes.

The computation of a structural machine *M* determines the *trajectory of computation*, which is a tree in a general case and a sequence when the computation is deterministic and is performed by a single processor unit.

The grid (network) on the second level is a grid array and is modeled and mathematically described by an abstract grid automaton, which is defined in the following way [16].

A *grid automaton* is a system of abstract automata, which are situated in a grid, are called nodes, are connected in a definite manner, and interact with one another using their connections.

A physical realization of a grin automaton is a grid array while the mathematical model of a grid array is a grid automaton [16].

A grid automaton *G* is described by three grid characteristics. The grid characteristics are:

1. The space organization or structure of the grid automaton *G*. This space structure may be in the physical space, reflecting where the corresponding information processing systems (nodes) are situated, or it may be a mathematical structure defined by the geometry of node relations. There are three kinds of the spatial organization of a grid automaton: a static structure that is always the same; a persistent dynamic structure that eventually changes between different cycles of computation; and a flexible dynamic structure that eventually changes at any time of computation.
2. The topology of *G* is determined by the type of the node neighborhood. A neighborhood of a node is the set of those nodes with which this node directly interacts. In a grid, these are often the nodes that are physically the closest to the node in question.
   For example, if each node has only two neighbors (right and left), it defines linear topology in *G*. When there are four nodes (upper, below, right, and left), the *G* has a two-dimensional rectangular topology. three-node characteristics.
3. The dynamics of *G* determine by what rules its nodes exchange information with each other and with the environment of *G*. For example, when the interaction of Turing machines in a grid automaton *G* is determined by a Turing machine, then *G* is equivalent to a Turing machine. At the same time, when the interaction of Turing machines in a grid automaton *G* is random, then *G* is much more powerful than any Turing machine.

Interaction with the environment separates two classes of grid automata/arrays: open grid automata/arrays interact with the environment through definite connections, while closed grid automata/arrays have no interaction with the environment. For example, Turing machines are usually considered as closed automata because they begin functioning from some start state and tape configuration, finish functioning in some final state and tape configuration, and do not have any interactions with their environment.

The node characteristics are:

1. The *category* of the node. For example, one category comprises finite automata, while another category encompasses a Turing machine. In general, a node can belong to the category of grid automata, that is, a node of a grid automaton can be a grid automaton. The category of a node also defines to some extent the structure of this node.
2. The *external dynamics* of the node determine the interactions of this node. According to this characteristic, there are three types of nodes: *accepting nodes* that only accept or reject their input; *generating nodes* that only produce some input, and *transducing nodes* that both accept some input and produce some input. Note that nodes with the same external dynamics can work in grids with various dynamics.
3. The *internal dynamics* of the node determine what processes go inside this node. For example, the internal dynamics of a finite automaton are defined by its transition function, while the internal dynamics of a Turing machine are defined by its rules. Differences in the internal dynamics of nodes are very important because a change in producing the output allows us to go from conventional Turing machines to much more powerful inductive Turing machines of the first order.

Grid automata represent the higher level of the theoretical models of computation. For instance, in comparison with cellular automata, which are uniform, a grid automaton can contain different kinds of automata as its nodes. At the same time, finite automata, Turing machines, and inductive Turing machines can belong to one and the same grid automaton. In comparison with systolic arrays, connections between different nodes in a grid automaton can be arbitrary like connections in neural networks. In comparison with neural networks and Petri nets, a grid automaton contains, as its nodes, more powerful machines than finite automata. As a result, many models of distributed computations such as neural networks, cellular automata, systolic arrays, Petri nets, and many others are important special kinds of grid automata. A significant property of grid automata is their ability to form various hierarchical structures because a node can also be a grid automaton, in which a node can be a grid automaton, and so on. In grid automata, interaction and communication become as important as computation. This peculiarity results in a variety of types of grid automata, their architecture, functioning modes, space organization, and temporal forms

V. An Example Use Case: Secure Video Delivery Using a Multi-Cloud Infrastructure connected by the Internet

The theory and practice of autopoietic and cognitive digital automata are discussed in [4, 5]. In this paper, we focus on how application layer security is decoupled from the IaaS and PaaS security that is managed by the cloud providers who offer these services. As described above, the first level of HAON functional node consists of individual machines and local computing networks organizing *Infrastructure as a Service* (IaaS) and *Platform as a Service* (PaaS), which are often provided by a cloud service provider. The application components that belong to the service in the grid array are executed using the IaaS and PaaS services. On the second level, the grid network is deployed and managed using the application components and the IaaS and PaaS nodes. On the third level the structural machine that deploys and manages the life processes of the grid networks/arrays.

Figure 2 shows a hierarchical autopoietic and cognitive application network specified in a digital genome and deployed in a distributed network consisting of two cloud services from different vendors.

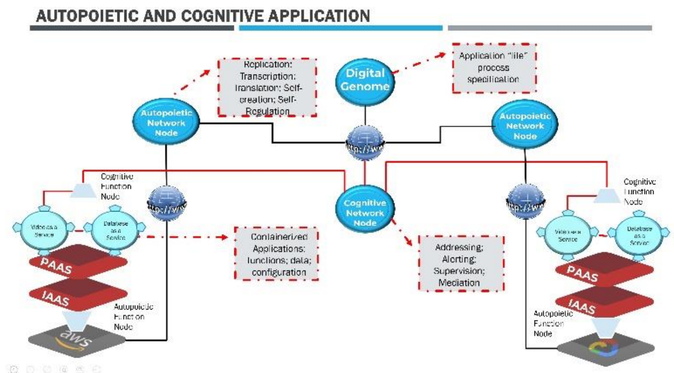

Figure 2: A structural machine framework deploying a video service application using infrastructure as a service (IaaS) and Platform as a Service (PaaS) from multiple clouds.

The autopoietic network node contains the knowledge about the application components (their executables, configurations, and associated data), where to get the resources (IaaS and PaaS services), how to deploy them, monitor them, and manage them. The cognitive network node contains the knowledge about the application "life" processes that fulfill both functional, and non-functional requirements. This knowledge involves the end-to-end service goals, and how to regulate individual component behaviors to optimize global application behavior.

Each functional node contains a knowledge structure that executes the local processes when information received from other functional nodes is processed. The information either creates new knowledge that causes local behavioral changes and potential communication of information exchange to other functional nodes. All the nodes that are wired together can also fire together to exhibit the collective behavior defined in life processes. The second-level grid network nodes manage the autopoietic and cognitive behaviors of the downstream computing structures. The structural machine, which we call the digital genome, contains the knowledge of the system behavior in the form of life processes that are deployed and executed by the grid networks downstream.

Figure 3 shows two functional nodes that deliver a video service where users can access a video stored in a database. Each service component is managed by the cognitive and autopoietic managers that execute the processes specified in the genome to maintain system-level stability, security, and safety. The first differentiation of this approach is the system-level knowledge of the life processes that include both the computer and the computed. The second differentiation comes from the structure of the knowledge network where each function node contains entities, their relationships, and behaviors.

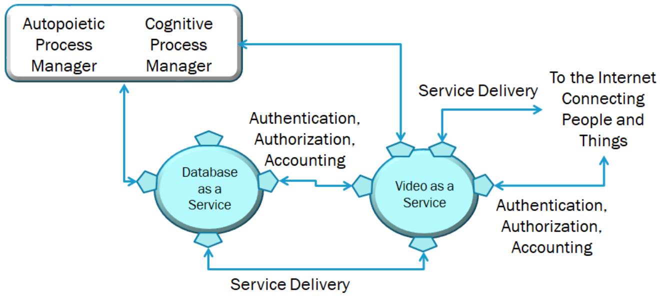

Figure 3: Functional node interaction

The schema of the knowledge structures and the operations are discussed in [6]. Suffice it to say that the operations on the knowledge structures allow the restructuring of the knowledge network without disrupting the end-to-end service behavior [4, 5]. The authentication, authorization, and accounting communication between application components use crypto-secure protocols independent of what IaaS and PaaS layer security protocols are. The service-oriented knowledge network provides a higher level of stability and security at the service level independent of individual functional node stability and security. For example, cognitive and autopoietic managers can detect local fluctuations that are affecting the functional nodes and take predictive actions based on best practice policies encoded in the genome.

The operations on the knowledge network perform restructuring of the components by adding, deleting, or reconnecting the functional nodes.

We believe that infusion of autopoiesis and cognitive process management using the tools derived from GTI provides safer and stable processes that connect businesses, people, and things.

VI. CONCLUSION

In this paper, a new three-tier Internet architecture is suggested and its properties are explored. It is oriented at improving the security and safety of the Internet functioning and providing better tools for interaction between consumers and Internet services. The suggested architecture is based on biological analogies, innovative computational models, such as structural machines and grid automata, as well as on the far-reaching general theory of information. The advantages of the suggested technological Internet architecture are explained in comparison with the traditional one.

The suggested approach leads to three directions of its further development. One of them is the advancement of the theoretical models getting a more exact picture of information processes in big and small computational and communication networks.

Another direction goes into the technological realization of the discussed here theoretical findings and practical ideas.

One more direction is the synthesis of suggested Internet architecture with name-oriented networking or named data networking (NDN) because many of the Internet's problems are related to names and naming [17, 18].

References

[1] Naughton, John. A brief history of the future. Weidenfeld & Nicolson, 2015.

[2] Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, Stephen Wolff, "Brief History of the Internet" 1997. https://www.internetsociety.org/internet/history-internet/brief-history-internet (Accessed on 04/24/022)

[3] Won Kim, Ok-Ran Jeong, Chulyun Kim, Jungmin So, The dark side of the Internet: Attacks, costs, and responses, Information Systems, Volume 36, Issue 3, 2011, Pages 675-705,

[4] Mikkilineni, R. Infusing Autopoietic and Cognitive Behaviors into Digital Automata to Improve Their Sentience, Resilience, and Intelligence. Big Data Cogn. Comput. 2022, 6, 7. https://doi.org/10.3390/bdcc6010007

[5] Mikkilineni, R. A New Class of Autopoietic and Cognitive Machines. Information 2022, 13, 24. https://doi.org/10.3390/info13010024

[6] Burgin, M. and Mikkilineni, R. From Data Processing to Knowledge Processing: Working with Operational Schemas by Autopoietic Machines, Big Data Cogn. Comput. 2021, v. 5, 13.

(https://doi.org/10.3390/bdcc5010013

[7] Burgin, M. Theory of Information: Fundamentality, Diversity and Unification, World Scientific: Singapore, 2010.

[8] Burgin, M. Information in the Structure of the World, Information: Theories & Applications, 2011, 18, 16 - 32

[9] Burgin, M. Structural Reality, Nova Science Publishers, New York, 2012

[10] Burgin, M. Theory of Knowledge: Structures and Processes, World Scientific, New York/London/Singapore, 2016

[11] Burgin, M. Information Processing by Structural Machines, in Theoretical Information Studies: Information in the World, World Scientific, New York/London/Singapore, 2020, pp. 323–371

[12] Hawkins, J. A Thousand Brains: A New Theory of Intelligence; Basic Books: New York, NY, USA, 2021. [Google Scholar]

[13] Yanai, I.; Martin, L. The Society of Genes; Harvard University Press: Boston, MA, USA, 2016. [Google Scholar]

[14] P. Cockshott, L. M. MacKenzie and G. Michaelson, "Computation and Its Limits," Oxford University Press, Oxford, 2012, p. 215

[15] Burgin, M. Superecursive Algorithms; Springer: New York, NY, USA, 2005. [Google Scholar]

[16] Burgin, M. From Neural networks to Grid Automata, in Proceedings of the IASTED International Conference "Modeling and Simulation", Palm Springs, California, 2003, pp. 307-312

[17] O'Toole, J.; Gifford, D.K. Names should mean what, not where, in EW 5. In Proceedings of the 5th Workshop on ACM SIGOPS European Workshop: Models and Paradigms for Distributed Systems Structuring, Mont Saint-Michel, France, 1992, pp. 1–5.

[18] Balakrishnan, H.; Lakshminarayanan, K.; Ratnasamy, S.; Shenker, S.; Stoica, I.; Walfish, M. A Layered Naming Architecture for the Internet. *ACM SIGCOMM Computer Communication Review*, 2004, 34, 343–352