The brain is processing information, not data: Does anybody knows about that?

Emanuel Diamant

VIDIA-mant, POB 933, Kiriat Ono 5510801, Israel

E-Mail: emanl.245@gmail.com;

Abstract: Discriminating and opposing "data" and "information" (as it is emphasized in the paper's title) for most of the scientific community sounds like something odd and unnatural. Raised in the spirit of Shannon's Information Theory, most of the scientific community is convinced that data and information are inseparable. Nevertheless, over the last decade we witness a growing recognition that Shannon's Information Theory is wrong, or speaking more politely, is limited only to data communication issues. Today, distinguishing data and information processing is gradually becoming a popular and widespread trend. However, because this trend is missing a firm theoretical underpinning, it looks a bit messy and inconsistent. Despite of this, the paradigm shift in contemporary science is clearly evident – from a data processing (computational) approach we are firstly moving to an information processing (cognitive) approach. ("Cognitive" here implies "capable of information processing"). Undeniably, Computational biology, Computational neuroscience, Computational linguistics (and so on) are being replaced today by Cognitive biology, Cognitive neuroscience, Cognitive linguistics, and so on. However, this tendency is hampered by a lack of understanding about what is "information processing". Subsequently, a question "what is information?" immediately rise up. A consensus answer to it does not exist.
I believe I have the answer. But instead of repetitive explanations about what is information, I prefer to bring an informational perspective to the everyday practice of scientific exploration, especially biological and neuroscience explorations. Maybe this will be more advantageous.

Keywords: Data and Information, Information duality, Physical and Semantic information, Information materialization.

1. Introduction.

There is a joke that the value of a great scientific theory is determined by the time it impedes the advent of another new theory. In this regard, Shannon's Information Theory is a real champion that has dominated in many scientific fields for more than a half of a century.

The notion of "Information" was at first introduced by Shannon in his seminal 1948 paper "A Mathematical Theory of Communication", [1]. The original aim of the theory was to solve a purely technical problem: to increase the performance of a communication system. In his theory, Shannon defines information in terms of signal's statistical properties and the uncertainty of receiving a particular signal among those that are possible. He has explicitly set aside any discussion about signal's value or meaning.

In the year 1949, he wrote: "These semantic aspects of communication are irrelevant to the engineering problem… It is important to emphasize, at the start, that we are not concerned with the meaning or the truth of messages; semantics lies outside the scope of mathematical information theory", [2].

However, in biology, as in many other modern sciences, semantic aspect of a message is of a paramount importance. But, fascinated with the achievements of Information theory in the communication sphere, various scientific communities were eager to apply it almost in every other research field. That forced Shannon to issue an additional warning (in 1956): "In short, information theory is currently partaking of a somewhat heady draught of general popularity. It will be all too easy for our somewhat artificial prosperity to collapse overnight when it is realized that the use of a few exciting words like information, entropy, redundancy, do not solve all our problems", [3].

Yet the mainstream sciences continue to ignore Shannon's warnings. Therefore, even today, the interrelations between "information" and "data", "information" and "semantics", "semantics" and "knowledge" remain undefined, blurred and intuitive (due to the heritage of the Information Theory).

It must be mentioned (in this regard) that the first attempt to clarify the relations between "information" and "semantics" was made about 60 years ago by Yehoshua Bar-Hillel and Rudolf Carnap, [4]. As to my knowledge, they were the first who coined the term "Semantic Information". They have sincerely believed that such a merging can be possible: "Prevailing theory of communication (or transmission of information) deliberately neglects the semantic aspects of communication, i. e., the meaning of the messages… Instead of dealing with the information carried by letters, sound waves, and the like, we may talk about the information carried by the sentence", [4].

However, they were not successful in their attempt to unite the mathematical theory of information and semantics. The mainstream thinking at that time was determined by The Mathematical Theory of Communication, which does not distinguish between data and information. By today's standards, the distinction between data and information is irrelevant and meaningless. For this reason, the issue of information handling (particularly semantic information handling in biological data streams, which inundate today's biological research) remains neglected and unsettled. Therefore, it will be our duty to try to address this challenging issue.

2. What is information?

As it was said above, Shannon defines information as the entropy of a discrete set of probabilities, as an opportunity to reduce uncertainty of a received data transmit. My definition of information relies on the Kolmogorov's view on the matter [5].

A slightly modified and an extended version of Kolmogorov's description sounds today (in my words) like this: **"Information is a linguistic description of structures observable in a given data set".**

To make the scrutiny into this definition more palpable I propose a digital image to be considered as a given data set. A digital image is a two-dimensional set of data elements called picture elements or pixels. In an image, pixels are distributed not randomly, but, due to the similarity in their physical properties, they are naturally grouped into some clusters or clumps. I propose to call these clusters **primary or physical data structures**.

In the eyes of an external observer, the primary data structures are further arranged into more larger and complex agglomerations, which I propose to call **secondary data structures**. These secondary structures reflect human observer's view on the grouping of primary data structures, and therefore they could be called **meaningful or semantic data structures**. While formation of primary (physical) data structures is guided by objective (natural, physical) properties of the data, the ensuing formation of secondary (semantic) data structures is a subjective process guided by human conventions and habits.

As it was said, **Description of structures observable in a data set should be called "Information".** In this regard, two types of information must be distinguished – **Physical Information and Semantic Information**. They are both language-based descriptions; however, physical information can be described with a variety of languages (recall that mathematics is also a language), while semantic information can be described only by means of natural human language. (More details on the subject you can find in [6]).

Those, who will go and look in [6], would discover that every information description is a top-down evolving coarse-to-fine hierarchy of descriptions representing various levels of description complexity (various levels of description details). Physical information hierarchy is located at the lowest level of the semantic hierarchy. The process of sensor data interpretation is reified as a process of physical information extraction from the input data, followed by an attempt to associate this physical information about the input data with physical information already retained at the lowest level of a semantic hierarchy. If such association is achieved, the input physical information becomes related (via the physical information retained in the system) with a relevant linguistic term, with a word that places the physical information in the context

of a phrase, which provides the semantic interpretation of it. In such a way, the input physical information becomes named with an appropriate linguistic label and framed into a suitable linguistic phrase (and further – in a story, a tale, a narrative), which provides the desired meaning for the input physical information.

The segregation between physical and semantic information is the most essential insight about the nature of information. Another insight is that, because of the subjective nature of semantic information, its creation cannot be formalized. Semantic information hierarchy, thus, cannot be learned and has to be provided to the system always from the outside, always as a gift, a grant, an offering. The next important outcome from the definition given above is the understanding that information descriptions are always reified as a string of words, a piece of text, a narrative.

Bearing in mind all these new peculiarities, we can proceed to the revision of today's brain information processing approaches.

3. Rethinking brain information processing

It is a widespread agreement that neuroscience will be the main research topic of the century. As part of this general agreement, there is a remarkable spread of use of terms and expressions related to brain's processing performance. Although the meaning of these terms and expressions is not always perfectly defined. You cannot oppose this fashion without palpable arguments that will support your claims. That is exactly what I am intended to do. From the point of view of the theory just exposed above, the current use of information processing terms is ambiguous and misleading. Therefore, I feel myself obliged to try and to clarify at least some of these commonly held misunderstandings.

3.1 Sensory information processing

Despite its wide and frequent use, the term "sensory information" (645 000 Google hits) is incorrect and misleading – sensors are data gathering devices. Raw data, not information. Data, as you already know, by itself is meaningless. That is, data is semantic information devoid. I stress here "semantic information" because that is what we have in mind and what we are usually looking for when we are talking about information. However, in the spirit of Shannon's information theory, people are familiar only with one sort of information that is not distinguishable from data. The duality of information is unknown to them. So, people are looking for Shannon's information, and they are proud with what they have at hand and what they are doing with confidence.

Some accidentally selected quotes will illustrate what is going on here:

"The fundamental concept behind the Bayesian approach to perceptual computations is that the **information provided by a set of sensory data** about the world is represented by a conditional probability density function over the set of unknown variables – the posterior density function", [9].

"The ability to detect, discriminate and identify sensory signals is limited by how efficiently **information in sensory representations** is put to use in the control of behavior. A stimulus activates a population of neurons in various areas of the brain. To guide behavior, the brain must correctly decode this population response and **extract the sensory information** as reliably as possible", [10].

"More specifically, neurons in the lateral intraparietal area (LIP) have been shown **to accumulate sensory information** provided by earlier visual cortex when a decision is being formed", [11].

 "Nervous systems are standardly interpreted as information processing input–output devices. They **receive environmental information from their sensors as input**, subsequently process or adjust this information, and use the result to control effectors, providing output", [12].

According to the theory given above, there is only one sort of information that can be derived from the sensory data – physical information. Only physical information is further processed in the brain nervous

system aimed to get a suitable interpretation of it, its proper semantic meaning. You can call that – physical information is endowed with semantics.

"Information as processed data" (a popular synopsis of Shannon's information) does not take part in brain semantic information processing. You can easily accept that if you recall from your own experience that:

- We get the meaning of a written word irrelevant to the font size of the letters or their style (irrelevant to data features).
- We recognize equally well a portrait of a known person on a huge-size advertising billboard, on a magazine front page, or on a postage stamp – perceptual information is dimensionless, (while data features are not).
- We get the meaning of a scene irrelevant to its illumination. We look on the old black-and-white photos and we do not perceive the lack of colors.
- The same is true for voice perception and spoken utterance understanding – we understand what is being said irrelevantly to who is speaking (a man, a women, a child). Irrelevant to the volume levels of the speech (loudly or as a whisper).
- Blind people read Brail-style writings irrelevant to the size or the form of the touch-code.

Recall that according to the definition given in this paper, physical information is a **description of structures** observable in a given data set. Only "description of structures" take part in further information processing. Original data features are become dissolved in the descriptions and do not take part in a scene understanding process.

Do I need to remind you that today all classical data interpretation tasks are busy with data features elucidation, delineation and treatment? Neuroscience and brain research are not free from these flaws. I hope my humble explanations will help to avoid these hurdles in the future.

3.2 About Information flow processing

Contemporary brain research paradigm is anchored to several basic generally accepted truths: Brain is processing information. Neurons are functional units that facilitate this processing. Despite their discrete anatomical structure, neurons are not single functional units – successful information processing requires tight cooperation between coworkers. For that reason, neurons are interconnected into a network, in which they communicate between each other conveying, exchanging, carrying – in one word, conjointly processing information. This passage of information between interconnected (at various levels of organization) neurons has received even a special brand name "Neural Information Flow" and has become a subject of increasing study and investigation, [13], [14].

From the standpoint of interneuron communication, neurons could be seen as a chain where two subsequent units are connected by means of a synaptic gap junction. Each single neuron is composed of three constituting parts: the dendrites (at the input), the cell body or soma (the main part), and the axon (the throw out part). Understanding the functional roles of these neuron compartments, and how information is flowing through them was the primary goal of the neuroscience for the most part of the last century. However, for some reasons, among the three parts facilitating the flow of information across a neuron, the axon part was the most studied and scrutinized, [15].

The reason for this is well known – axonal information flow was the only one that could be discerned by an external observer. Axonal information flow is associated with propagation of electrical charge changes (called Action Potential pulses) moving backward (rarely) and forward (most of the time) along the axon length. Textbooks typically indicate that the transition of action potentials is recorded by microelectrodes placed close to the neuron body either outside the cell (extracellular recordings) or within the cell (intracellular recordings).

Since its first introduction in the late twenties and until the late sixties action potential recording was the only technique available for interneuron information flow examination and investigation. Shannon's Information Theory (1948) and Hodgkin-Huxley's model of axon conductness (1952) have provided action potential propagation studies with a firm theoretical underpinning, making them the most dominant and vibrant

research branch of the neuroscience. Even the newest most ambitious Human Brain Project recently launched by the European Commission has a special program aimed on exploration of action potential recordings, [16].

But, in all these exciting stories one vexing point remains unresolved: what is the information that is conveyed via the axon "cable"? How trains of action potentials encode this information? To what extent learning how action potentials are generated and transmitted leads to an understanding about what information is conveyed along the axon pathway. Nobody has an answer to these questions and, it is much worse than that, nobody even ask them. But without answering these questions, brain information processing turns out simply as a storytelling festival.

There are good reasons to believe that the Information theory, explained in the previous sections of this paper, can provide some useful prompts to resolve the inconveniences we face in such a state of affairs. The theory claims that information always looks like as a linguistic text, a string of words and phrases. That is what we have to look for when we talk about neuron information flow. That is what we have to look for when we are talking about axonal action potential propagation. And would we not be so eager to look for India, we would more easily accept that that is America arising in front of us.

Indeed, the advent of electron microscopy techniques (dated by the early nineties of the last century) has promptly moved neuron information flow studies into the era of molecular communications. The molecular hypothesis implies that bioactive molecules (lipids, proteins, a variety of RNAs, and even DNAs) could be information carriers in the case of interneuron communication, [18].

Indeed, it is now generally accepted that chemical synapses perform their communication duties by streaming across the synaptic gap a flow of vesicles released from the presynaptic axon terminal. Vesicles are small structures containing proteins, lipids, and nucleic acid cargo, as well as neuromodulators and neurotransmitter chemicals released by a presynaptic neuron. Thus, information flow in the input dendrite part, in the soma, and at the axon terminal is accomplished by means of molecular information carriers. Only information transfer along the axon pathway is implemented as action potential spikes flow?! Can such a thing be?!

Nature is conservative, it is hard to believe that within one single path of information flow (soma – axon – axon terminal – synaptic gap – dendritic ship) most of the track would be realized as a molecular freight transportation and only one intermediate segment will perform in a completely different way, as a spike trains communication section. That seems incredible.

New evidence indicates that "the afferent axons carry, not only minute-to-minute operative information in the form of spike trains, **but also instructions in the form of specific large signaling molecules**. We have hardly any idea how this is done: nor how far down the chain of information processing this process extends", [18]. Well, it is hard to confess that for a very long time action potential propagation was considered as the only and the main neuron information conveying carrier. The idea that information encoding can be performed as train of spikes was a misleading and wrong idea that unfortunately dominated and continues to dominate neuroscience research for more than half a century.

On the other hand, the idea that information is encoded in large molecular arrangements fits very well our definition of information as text strings and word sequences. It does not matter that these text sequences are written with nucleotide letters and amino acid signs. The idea "information is a text string" is convincingly supported by the newest neuron information flow research findings.

And what is then the fate of the action potential? Discharged, dismissed, disqualified? Action potential was discovered as a sign of neuron activity, and as such it has faithfully served neuroscience for many years. It is now rejected only as a means of information representation and encoding. But as a sign of neuron activity it will continue to serve the research community in many valuable applications such as neuron synchronous behavior studies, electroencephalography for brain activity monitoring, EEG recordings for cortical connectivity evaluation, and so on.

From the standpoint of molecular communication, wagon trains loaded with molecular cargo are moving along the axon pathway. And action potential spikes can be seen as the noise that follows this wagon trains passage. This noise would say you nothing about the cargo the train is carrying. But it can provide you with some information about the intensity of the traffic (for example), and that can be also a useful aspect of the transportation flow, as it is explained just above.

3.3   Reconsidering text processing

Reifying information as a text string is an important novelty introduced in this paper. It is reasonable to expect that it will have a tremendous impact on the progress of biological science in general and boost the brain research and neuroscience in particular. Contemporary biological science witnesses huge volumes of experimental data produced by means of next generation high-throughput technologies. To be useful, such data has to be processed and interpreted in a human-like fashion. That is, valuable semantic information about collected pieces of data has to be revealed as a result of this processing. There is no doubt, that just semantic information in form of a linguistic text (exactly as it is repeatedly emphasized in this paper) is eagerly expected here.

Several quotations from a list of recent publications will promptly illustrate the current state of affairs:

 "…much of the past decade of molecular biology has been devoted to deciphering the meaning of this code. On this premise, the ENCODE Project Consortium sought to discover a complete catalog of all functional elements in the human genome, **analogous to delineating sentences and words that comprise the human genome,** and understanding the type of function each element plays", [19].

"Today, everyone speaks about the genetic code – genes encoded in DNA that serve as the information-bearing molecules for all biological entities… **Syntax-based quantitative approaches that focus on nucleic acid sequences as information-bearing molecules will have** problems explaining why identical nucleic sequences may have **different and, in extreme cases, even contradictory meanings**", [20].

"…it provides a **foundation for DNA processing as it can implement all basic text processing operations on DNA molecules** including insert, delete, replace, cut and paste and copy and paste**",** [21].

"Meanwhile, the linguistic approach has also lost its metaphorical character, and the similarity between **linguistic languages/codes and genetic storage media are not only accepted, but are fully adapted in** bioinformatics, biolinguistics, protein linguistics, biohermeneutics and biosemiotics", [22].

Despite wide use in these quotations of modern buzz-terms (like "information", "linguistic expressions", "text processing"), I am not sure that people in biologic research community (as well as people in the other research communities) are aware about the real meaning of these terms and the consequences of their use. With the Shannon's definition of information at hand, they do not really discriminate between data and information. Their next generation high-throughput technologies provide them with huge amounts of raw data and they are busy (as usual) with extraction of information (essentially – semantic information, but they do not know about that) from this data. Sorry, that does not work. Raw data processing can give them only physical information descriptions. People know nothing about information duality, about physical and semantic information dichotomy and mutual interdependency. "Processed data is the information we look for" – that is what they have been thought for near a half of a century by the zealots of Shannon's information theory. My purpose in this paper is not to preach on behalf of a new doctrine. My purpose is to draw people's attentions to the flaws of their currently running mega-projects, devoid of any signs of understanding about physical and semantic information coexistence.

To make my arguments more palpable I would like to remind you a beautiful story about Craig Venter's experiments. Busy with creating a synthetic bacterium, Craig Venter and his colleagues took a DNA sequence as a computer file, modified it, made a new DNA sequence that was placed into a cell, which then reproduce itself as a new bacterium. Sequencing it, "showed our genome to have the 1 077 947 base pairs, exactly as intended, including nineteen expected differences from the native genome, as well as the four watermark sequences, a critical part of the proof that the DNA was synthetic" [7].

The watermark sequences intentionally inserted into bacterial genome have the purpose to verify that the given genome indeed belongs to a new artificial organism and is not assembled from DNA "parts" of other natural bacteria. The watermarks were English texts – aphorisms of well-known public figures (like Robert Oppenheimer, Richard Feynman, and James Joyce), which have been written in specially designed codes composed from the four DNA "letters" (G, A, T, C). "We have put a significant effort into the design of our watermarks to ensure that we could safely code for complex messages in the DNA sequence. We designed a (much more complete) system that would enable us to code the entire English alphabet together with punctuation, numbers, and symbols. This cipher was the key to decoding the watermarks" [7].

The most exciting feature of the experiment was that the watermark texts were undistinguishable from the rest of the genomic sequences. And like the rest of the genomic sequence, they were viable, that is, have reproduced and duplicate themselves in billions of times of bacterial cell replications.

But in the light of our discussion, the critically important outcome of Craig Venter's experiment is that the reverse deciphering of watermark sequences (uncovering the original text messages) cannot be done under any imaginable condition. Only one condition is compulsory – some preliminary knowledge about the encoded text has to be granted to the deciphering programmer in advance (e.g., that the texts are written in English and some of the encoded words were explicitly designated) [8].

That fits well the hypothetical assumption, which stems from the information theory briefly exposed in this paper: Semantic information cannot be derived from physical information (from the input data that bears it). In Craig Venter's case, the English text of a watermark is the semantic information while the result of genome sequencing is the DNA data (nucleotide letters). Bearing that in mind, we can look again at the flourishing business of Bioinformatics and some other well-known research enterprises (like the ENCODE project, for example) where extensive statistical data processing is done aimed to reveal semantic (functional) information text descriptions.

The need for a referential knowledge base (the referential semantic information hierarchy, the preliminary knowledge in the Craig Venter's experiment), which is a foundational requirement for a successful physical information interpretation (delivering the desired meaning for the input data) has also to be mastered carefully. First, the difference between knowledge and information (so often emphasized in almost every research application) can be now finally settled – knowledge (the referential knowledge base) is the memorized information kept in the semantic information hierarchy! (And nothing else!) As it follows from our information definition (and illustrated by the above Craig Venter's experiment), the reference knowledge base (the semantic information hierarchy) cannot be learned, it must be granted or delivered from the outside. That pose in doubt all numerous claims about various learning techniques successfully developed for various biotic research trials.

4 Concluding remarks

In this paper, I have tried to explain the devastating results of mixing data and information, as well as the price for neglecting information dichotomy and information reifying as a text string of words and letters. I have chosen some representative cases from biological and neuroscience research practice to explain the consequences of such ignorance and misuse of information elucidation principles. Only few examples have been brought here while the bulk of mistreated cases was left behind the scope of this paper. They are still awaiting for a revision, but I think it will be interesting just to name some of them for the sake of future research roadmaps forecast.

1. Memory. As information becomes reified as text strings, the repositories for its storage and accommodation have to be also materialized. Dendritic spines might look as a promising candidate for this purpose.

2. Genetic code is a huge repository of information texts accumulated in the course of evolution. Epigenetic code is a collection of narratives acquired in the course of personal interaction with the surrounding world.

3. The reference knowledge base (the reference information) needed for input data interpretation (cognitive input data processing) can be acquired in two ways – by vertical information transfer from the predecessors (genetic inheritance) and by horizontal information transfer from the social contemporaries (epigenetic acquirement).

4. Mirror neurons can be seen as physical information holding entities located at the low level of the system's reference knowledge base and activated when the input physical information is in search for its in-system memorized referents.

5. Exons and introns can be seen as physical and semantic information (texts) analogs, similar to data (attributes) and procedure (methods) components of Object-oriented programming paradigm.

6. Thinking. If information is reified as text and thinking is assumed to be a way of memorized information processing, then thinking must be seen as text processing. That is, thinking is information text processing – text editing, text modification, tuning, adjustment, decomposition and reassembling. What does it really mean? – I don't know, and nobody knows yet. Our computers are data processing machines. How to do information texts processing is still a subject for future investigations.

7. Language. Text is essentially written in a language. It seems quite reasonable to think that natural evolution has changed the language used for information descriptions in course of its evolutionary development. Therefore, there must be a mix of languages interchangeably used by any living being. May be the Long Term Potentiation of memory is a kind of text reshaping and rewriting from one language to another. May be genetic and epigenetic memories are written in different languages (Nucleotide codes are not a language, if you remember Craig Venter's experiments).

8. Ontology. There is an intuitive confession that interpretation requires a reference knowledge base. Essentially, that is right. But contemporary computational (data processing) approaches use ontologies as a form of knowledge representation and as a tool for turning data into knowledge. Knowledge is memorized information – data cannot be converted to information. So, the purpose of ontologies "to describe the semantics of data", [23], is misconceived.

That is only a short list of possible future research directions, which I hope are soon to come. Meanwhile, an intuitive turn to an information processing approach is truly visible in our scientific landscape. From traditional well-established data processing way of thinking, we are swiftly moving to an information processing way of thinking. The first is known as a "computational approach", the second is dubbed as a "cognitive approach", where "cognitive" implies "capable of information processing".

It is hard not to notice how "Computational biology", "Computational linguistics", "Computational intelligence" (and so on) are swiftly turn into "Cognitive biology", "Cognitive linguistics", "Cognitive intelligence" (and so on). Ask Google – such examples could be multiplied endlessly.
It does not matter that "what is information" for the majority of the scientific community is still unknown (as well as that the data-information interchange is a slip, that information is a composition of physical and semantic information and that information is always represented as a text string). Never mind, I hope my humble efforts will help them to change their minds.

References

[1] Shannon, C. (1948). **A Mathematical Theory of Communication**, 1948, Published by the Board of Trustees of the University of Illinois. Used with the permission of University of Illinois Press.
http://www.mast.queensu.ca/~math474/shannon1948.pdf

[2] Shannon, C., Weaver, W. (1949). **The Mathematical Theory of Communication**, University of Illinois Press, 1949. http://raley.english.ucsb.edu/wp-content/Engl800/Shannon-Weaver.pdf

[3] Shannon, C. (1956). **The Bandwagon**,
http://www.monoskop.org/images/2/2f/Shannon_Claude_E_1956_The_Bandwagon.pdf

[4] Bar-Hillel, Y. & Carnap, R. (1952), **An outline of a theory of semantic information**, Technical report No.247, October 27, 1952, http://www.survivor99.com/lcg/information/CARNAP-ILLEL.pdf

[5] Kolmogorov, A. (1965). **Three approaches to the quantitative definition of information**, Problems of Information and Transmission, Vol. 1, No. 1, pp. 1-7, 1965.
http://alexander.shen.free.fr/library/Kolmogorov65_Three-Approaches-to-Information.pdf

[6] Diamant, E.; Brain, Vision, Robotics and Artificial Intelligence. http://www.vidia-mant.info

[7] Venter, J. Craig, **Life at the Speed of Light: From the Double Helix to the Dawn of Digital Life,** Penguin, 2013, https://books.google.co.il/books?id=GWDD9yMb33cC&source=gbs_navlinks_s

[8] Shirriff K., **Using Arc to decode Venter's secret DNA watermark,** Ken Shirriff's blog
http://www.righto.com/2010/06/using-arc-to-decode-venters-secret-dna.html

[9] David C. Knill and Alexandre Pouget, **The Bayesian brain: the role of uncertainty in neural coding and computation,** TRENDS in Neurosciences Vol.27 No.12 December 2004
https://bi.snu.ac.kr/Courses/aplc12/11-2.pdf

[10] Mehrdad Jazayeri & J Anthony Movshon, **Optimal representation of sensory information by neural populations**, Nature Neuroscience, volume 9, number 5, May 2006,
http://ccnss.org/ccn_2013/materials/pdf/movshon/jazayeri-movshon-2006.pdf

[11] Tianming Yang & Michael N. Shadlen, **Probabilistic reasoning by neurons,** Nature 447, 1075-1080 (28 June 2007), http://www.cns.nyu.edu/events/spf/SPF_papers/yang2007probabilistic.pdf

[12] Fred Keijzer, Marc van Duijn and Pamela Lyon, **What nervous systems do: early evolution, input–output, and the skin brain thesis,** Adaptive Behavior · April 2013,
http://www.researchgate.net/publication/258123212

[13] Smith V.A., et al, **Computational Inference of Neural Information Flow Networks,** PLOS Computational Biology, November 2006 | Volume 2 | Issue 11 | e161
http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0020161

[14] Salinas E., and Sejnowski T., **Correlated neuronal activity and the flow of neural information,** Nature Reviews Neuroscience, Vol. 2, No. 8. pp. 539-550, (01 August 2001),
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.164.8514&rep=rep1&type=pdf

[15] Dominique Debanne, et al, **Axon Physiology,** *Physiol Rev* 91: 555–602, 2011;
http://physrev.physiology.org/content/physrev/91/2/555.full.pdf

[16] Gordon M. Shepherd, et al., **The Human Brain Project: neuroinformatics tools for integrating, searching and modeling multidisciplinary neuroscience data,** Trends in Neurosciences, Vol. 21, No. 11, 1998. http://www.neuroteoretica.org/doc/Shepherd_Hum_brain_project.pdf

[17] Koch C., and Segev, I., **The role of single neurons in information processing,** Nature Neuroscience supp. Vol. 3: 1171-1177, November 2000. http://lobster.ls.huji.ac.il/idan/files/Koch_Segev_NN2000.pdf

[18] Smythies J., and Edelstein L., **Transsynaptic modality codes in the brain: possible involvement of synchronized spike timing, microRNAs, exosomes and epigenetic processes,** Front. Integr. Neurosci., 04 January 2013. http://journal.frontiersin.org/article/10.3389/fnint.2012.00126/full

[19] Michael Hoffman, et al., **Integrative annotation of chromatin elements from ENCODE data**, Nucleic Acids Research, 2012, 1–15,
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.303.3972&rep=rep1&type=pdf

[20] G. Witzany, **Life is physics and chemistry and communication,** Annals of the New York Academy of Sciences, 2015 - Wiley Online Library, http://www.biocommunication.at/pdf/publications/chap%201-1.pdf

[21] Shabi U., et al., **Processing DNA molecules as text,** Syst Synth Biol. 2010 Sep;4(3):227-36. http://www.ncbi.nlm.nih.gov/pubmed/21189843

[22] Witzany, G., **Biocommunication and Natural Genome Editing**, World J Biol Chem. 2010 Nov 26; 1(11): 348–352. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3083936/

[23] Sklyar, N., **Survey of existing bio-ontologies**, University of Leipzig, 2001. http://lips.informatik.uni-leipzig.de/files/2001-30.pdf