

Using the information and communications technology data deluge from a semantic perspective of a dynamic challenge: What to learn and what to ignore?

-Part 3-



Prof. Eng. Ph.D. Victor GREU

Abstract

The paper analyses the Data Deluge premises and main approaches toward a **dynamic and semantic search of balance between generating data and extracting knowledge**, focusing on some semantic criteria of the search, using Information and Communications Technologies (ICT) context at the planetary scale, which could influence Earth environment and eventually humankind evolution with an unprecedented speed and impact. The main reasons for this kind of analyses come from the ICT generated global and exponential changes, which are more and more difficult to be controlled or at least estimated, regarding all their positive or negative consequences, just because of the complexity and speed of their vast penetration. The semantic analysis have to face a dynamic environment, because the relevant data/facts are fast changing (not from centuries or decades, but from months or even days), so we have to refine our search looking for new semantic terms or changing the old meaning (associated information/ knowledge) of the usual ones. It is emphasized that, although all technologies and ICT applications largely contribute to DD, lately, AI and its main branches, machine learning (ML), deep learning (DL) or cognitive learning (CL), are prominent in striving to extract information/knowledge from data, or even to compete with the human intelligence by cognitive features.

The work identified criteria of the search, which are linked with all (positive or negative) consequences of ICT on Information society (IS) toward Knowledge Based Society (KBS) and generally of World changes, just because of the complexity and speed of their vast proliferation, more visible by the dramatic climate changes, Earth resources fading, social imbalance or actually by the vulnerability of humankind to diseases spreading in the globalization context. **Such consequences should be alarm signals and optimization criteria** in our decision processes, as individuals, but mainly for the incumbents. These signals and the ICT impact on IS/KBS at planetary scale could indicate **the main criteria for the mentioned semantic search** in the ICT/AI/ML/DL/CL fast implementation processes, **in order to meet Earth ecosystem needs and expectations for a sustainable progress.**

An important paper conclusion is that **the semantic approach of decision processes, based on the generated data and extracted information, is a complex optimization target which just could match and benefit from the dynamic ICT/DD/ML/DL/CL context, where the data signification, implications and consequences are fast appearing and changing, because of this context potential to produce changes, but we have to be sure which of them are positive and which negative, to what extent and time period.** The most difficult and complex problem will be, every time, to **understand and define criteria for what is positive and what negative**, as all our **Earth ecosystem and even humankind behaviour and evolution is changing, but also our criteria should!** Such conclusions are sustained by some concrete examples of the massive expansion of ICT, like the case emergent 5G (among others), which could produce, besides positive, real and unprecedented benefits, new, but less desired effects. Here, the good news is that AI/ICT itself could compensate such negative effects, as it is revealed also by some concrete applications cases.

The second section of the paper completes the previous technical elements with other improvements of learning from the real-time complex processes of ICT/IS/KBS that could be observed just in the most prominent processes of learning at planetary scale, by ...ML/DL.

Although we speak about how **the machines are learning**, these could reveal not only the state-of-the-art of ICT/AI/ML/DL/CL and their limits (for example, ML and Wolpert famous "no free lunch" theorem), but indirectly shows the limits of humans that have the responsibility to design them, just considering the mentioned criteria.

Another conclusion is that all we see or perceive on these ICT/DD evolutions is not simply to understand nor to manage, but we have to timely analyse and try to optimize, as the price is so big. In fact, the complexity obliges us to also change the rules of learning we are used to practice, because the old rules could not be completely useful in the deep and ever-changing context, i.e., optimally and dynamically refine the criteria and the models of our thinking, in order to rapidly adapt to the new realities, which exceed the ICT/DD context, as we can see in these days, when the main challenges of our Earth ecosystem seem to come from other side (human and social health), but we also have to agree that any dramatic evolution at Earth scale have to be seen in the globalization context and we must use the power of ICT/DD/AI to optimally face the challenges (including the emotional impact on human life and behaviour, which ultimately could also induce, on short or long time, important changes, that we not really or completely perceive now).

The final conclusion is that, for **learning to better learn**, the emphasis must be secondary on the fact that **the human contribution to ICT/AI/ML/DL/CL processes is needed for better results “against” the difficulty/complexity, but primary in order to be sure that the “desired outputs” are designed to match the priority criteria we have above mentioned as essential for a sustainable development of ICT/IS/KBS and humankind evolution on Earth. More than these, learning to better learn, for a systemic approach of refining knowledge, is necessary to be the goal of all responsible people, beginning with the incumbents, the specialists and teachers, as everybody has to adapt to changes (Intelligence is the ability to adapt to change - Stephen Hawking) and use new technologies with appropriate criteria of optimization, i.e. timely analyse what and how we should learn or ignore.**

Keywords: Data Deluge, Internet of Things, artificial intelligence, machine learning, deep learning, cognitive learning, semantic based retrieval, green ITC, 5G, information society, knowledge based society

JEL Classification: L63; L86; M15; O31; O33

“In the age of technology there is constant access to vast amounts of information. The basket overflows; people get overwhelmed; the eye of the storm is not so much what goes on in the world, it is the confusion of how to think, feel, digest, and react to what goes on.”

Criss Jami, Venus in Arms

1. A dynamic and semantic search of balance between generating data and extracting knowledge

Sometimes, continuing to re-analyse a context could seem a relative time consuming, but always it depends on complexity and target. Recalling the evolution of information storage methods [3] would not be, generally, useful, unless we consider that it is important to deeply analyse current information before striving to acquire something new. We concluded that people “*had to store data in their brains memory cells*”, especially in the ancient times.

Now, when the Information and Communications Technology (ICT) is the main driving factor of the progress of the Information Society (IS) toward Knowledge Based Society (KBS), we are facing the unprecedented challenges of storage and manage the increasing amount of data that ICT, in the actual phase of **Data Deluge** (DD), generates when involved at Earth scale in all humankind fields of activity.

Still, other authors also consider that today it is useful to review the time picture of this evolution of data storage and management [1]:

“ONCE UPON A TIME, information was deposited only inside human brains, and ancient bards could spend hours retelling stories of conflicts and conquests. Then external data storage was invented. Small clay cylinders and tablets, invented in Sumer some 5,000 years ago, often contained just a dozen cuneiform characters, equivalent to a few hundred bytes (10^2 B). The Oresteia, a trilogy of Greek tragedies by Aeschylus (fifth century BCE), amounts to about 300,000 B....A radical shift came with Johannes Guttenberg’s printing press, using movable type. By 1500, less than half a century after printing’s introduction, European printers had released more than 11,000 new book editions...Then, in the 19th century, came photographs, sound recordings, and movies... during the 20th century came ... new storage modes, notably magnetic tapes and long-playing records...Beginning in the 1960s, computers expanded the scope of digitization to medical imaging (a digital mammogram is 50 MB), animated movies (2–3 gigabytes), intercontinental financial transfers, and eventually the mass emailing of spam (more than 100 million messages sent every minute). Such digitally stored information rapidly surpassed all printed materials.

Shakespeare's plays and poems amount to 5 MB, the equivalent of just a single high-resolution photograph....Printed materials have thus been reduced to a marginal component of overall global information storage. By the year 2000, all books in the Library of Congress were on the order of 10^{13} B (more than 10 terabytes) but that was less than 1 percent of the total collection (10^{15} B, about 3 petabytes) once all photographs, maps, movie, and audio recordings were added ..."

Coming back to our days, the mentioned evolution could help us to perceive the whole picture, in order to better manage its complex consequences:

"And in the 21st century this information is being generated ever faster. In its latest survey of data generated per minute in 2018, Domo, a cloud service, listed more than 97,000 hours of video streamed by Netflix users, nearly 4.5 million videos watched on YouTube, just over 18 million forecast requests on the Weather Channel, and more than 3 quadrillion bytes (3.1 petabytes) of other Internet data used in the United States alone. By 2016, the annual global data-creation rate surpassed 16 ZB (a zettabyte is 10^{21} B), and by 2025, it is expected to rise by another order of magnitude—that is, to about 160 ZB or 10^{23} B. And according to Domo, by 2020 1.7 MB of data will be generated every second for every one of the world's nearly 8 billion people."

It is still important to observe, even preliminary, the extent and the features of the mentioned consequences, i.e. the ways these are involved in the ICT **Data Deluge** (DD), starting from the same amazing presented figures [1]:

"These quantities lead to some obvious questions. Only a fraction of the data flood could be stored, but which part should that be? Challenges of storage are obvious even if less than 1 percent of this flow gets preserved. And for whatever we decide to store, the next question is how long should the data be preserved. No storage need last forever, but what is the optimal span? The highest prefix in the international system of units is yotta, $Y = 10^{24}$. We'll have that many bytes within a decade."

Noticing such facts, we consider them not only a realistic insight, but a comprehensive base to further continue our analysis toward a **dynamic and semantic search of balance between generating data and extracting knowledge**.

If the dynamic feature of the ICT/DD context is more than obvious, some semantic criteria of the search should better complete the image of these complex and huge processes, at the planetary scale, which influence Earth environment and eventually humankind evolution with an unprecedented speed and impact.

The main reasons for this kind of analyses come from the generated global and exponential changes, which are more and more difficult to be controlled or at least estimated, regarding all their positive or negative consequences, just because of the complexity and speed of their vast penetration.

In a simpler expression, the semantic analysis have to face a dynamic environment, because, as we could have seen above, the relevant data/facts are fast changing (not from centuries or decades, but from months or even days), so we have to refine our search looking for new semantic terms or changing the old meaning (associated information/ knowledge) of the usual ones.

The road to our search aim is also obvious in the above approach [1] conclusions, as a reasonable rhetoric question:

“And once we start creating more than 50 trillion bytes of information per person per year, will there be any real chance of making effective use of it? It is easier to find new prefixes for large databases than to decide how large is large enough. After all, there are fundamental differences between accumulated data, useful information, and insightful knowledge”.

Here we find the point of our previous approach [3], toward semantic signification of the plethora of applications and advances generated by the ITC/DD context, paving the “*long and winding road*” from data to information and eventually to knowledge.

We can see, all around and every day, the iceberg tips of ICT exponential development, including Artificial Intelligence (AI), Internet of Things (IoT), Cloud, Big Data, 3D Printing, Robotic Process Automation, Hardware Robotics, Blockchain, Augmented/Virtual Reality etc.[13][14][15][22].

Although all these technologies and ICT applications largely contribute to DD, lately, AI and its main branches, machine learning (ML), deep learning (DL) or cognitive learning (CL), are prominent in striving to extract information/knowledge from data, or even to compete with the human intelligence (HI) by cognitive features [17][18][20][12].

We can start our semantic search with a metaphoric approach, very well expressed by [2]:

“Miguel Street is a winding, narrow route through the Glen Park neighbourhood of San Francisco. Until a few years ago, only those living along the road travelled it, and they understood its challenges well. Now it's packed with cars that use it as a shortcut from congested Mission Street to heavily travelled Market Street. Residents must struggle to get to their homes, and accidents are a daily occurrence. • The problem began when smartphone apps like Waze, Apple Maps, and Google Maps came into widespread use, offering drivers real-time routing around traffic tie-ups. An estimated 1 billion drivers use such apps in the United States alone. • Today, traffic jams are popping up unexpectedly in previously quiet neighbourhoods around the country and the world. Along Adams Street, in the Boston neighbourhood of Dorchester, residents complain of speeding vehicles at rush hour, many with drivers who stare down at their phones to determine their next manoeuvre. London shortcuts, once a secret of black-cab drivers, are now overrun with app users. Israel was one of the first to feel the pain because Waze was founded there; it quickly caused such havoc that a resident of the Herzliya Bet neighbourhood sued the company.”

Our metaphor, laying from the “*long and winding road*” to the above *traffic jams*, is relevant not only for the complexity and difficulty of the **search of balance between generating data and extracting knowledge**, but more than this, for what we repeatedly claimed to be all (positive or negative) consequences of ICT/IS/KBS and generally of World changes, just because of the complexity and speed of their vast proliferation, more visible by the dramatic climate changes, Earth resources fading, social imbalance, or actually by the vulnerability of humankind to diseases spreading in the globalization context.

Without dramatizing, **such consequences should be alarm signals and optimization criteria** in our decision processes, as individuals, but mainly for the incumbents.

Anyway, these signals and the above iceberg tips of ICT impact on IS/KBS at planetary scale, could indicate **the main criteria for the mentioned semantic search** in the

ICT/AI/ML/DL/CL fast implementation processes, **in order to meet Earth ecosystem needs and expectations for a sustainable progress.**

An example of such semantic issues, appearing in one of the most prominent “tips” of ICT, as the emergent IoT is considered, is given by [4]:

“IoT systems, such as smart homes, smart buildings, smart cities, smart planet, smart farms, smart agriculture, smart factory, smart manufacturing, smart industry, smart parking, smart transportation, smart grid, etc., have been steadily developed and deployed...Many IoT systems are data intensive and are for the purpose of monitoring for fault detection and diagnosis of critical systems. A large volume of data steadily come out of a large number of sensors in the monitoring system. Thus, we need to consider how to store and manage these data. Existing time series databases (TSDBs) can be used for monitoring data storage, but they do not have good models for describing the data streams stored in the database ... we develop a semantic model for the specification of the monitoring data streams (time series data) in terms of which sensor generated the data stream, which metric of which entity the sensor is monitoring, what is the relation of the entity to other entities in the system, which measurement unit is used for the data stream, etc...”

Again we see a confirmation of our prior opinions [9][16] that, in this incredible advances of (ICT) technologies, for their optimal implementations, we mostly need better *models* matching them, along with refined criteria for decisions, based on ... retrieving data for refined knowledge, as it is finally concluded:

“Semantic based retrieval is especially important in a largescale integrated IoT-Edge-Cloud system, because of its sheer quantity of data, its huge number of computing and IoT devices that may store the data, and the dynamics in data migration and evolution. With better data semantics, data streams can be more effectively tracked and flexibly retrieved to help with timely data analysis and control decision making anywhere and anytime.”

The above example (a largescale integrated IoT-Edge-Cloud system) is quite relevant, as an emergent huge part of the dynamic ITC/IS/KBS context, where the semantic analysis is necessary, but not simple, as it is supposed to investigate all available data processes by considering *the dynamics in data migration and evolution*, aiming to provide better, *timely data analysis and control decision*.

Although it is not very obvious, as a “long and winding road”, **the semantic approach of decision processes, based on the generated data and extracted information, is a complex optimization target which just could match and benefit from the dynamic ICT/DD/ML/DL/CL context, where the data signification, implications and consequences are fast appearing and changing, because of this context potential to produce changes, but we have to be sure which of them are positive and which negative, to what extent and time period.**

Perhaps the most difficult and complex problem will be, every time, to **understand and define criteria for what is positive and what negative**, as all our **Earth ecosystem and even humankind behaviour and evolution is changing, but also our criteria should!**

Another example will clearly express such complicate challenge for our decision processes, in the emergent and important context of **green ITC** [5]:

“Energy costs for telecom operators around the world are already high: at the end of 2018, they accounted, on average, for around 5 percent of operating expenditures....This

growing energy challenge is, in large measure, a result of the exponential growth in traffic that new 5G services are likely to deliver. Although the 5G-new-radio standard is more energy efficient per gigabyte than are the 4G standards, the proposed 5G use cases and new spectrum bands will require many more mobile sites, outstripping potential energy efficiencies. Each 5G site will need two to three times more power than the 4G-equivalent site, according to industry estimates. At the same time, as more services are provided at the edge, the number of data centres will need to rise. “

We observe this way that the massive expansion of ICT, by the emergent 5G (among others), will produce, besides positive, real and unprecedented benefits, new, but less desired effects. Here, the good news is that AI/ICT itself could compensate such negative effects:

“Some savings lie in deploying artificial intelligence (AI) and the Internet of Things (IoT): some in structural and architectural transformations, and some in cheaper and more sustainable energy sourcing. ...Nevertheless, our work suggests that many operators can reduce energy costs by at least 15 to 20 percent in the space of just one year—and more over a longer period...Artificial-intelligence-driven sleep and shutdown running systems that are not in constant use consumes significant amounts of energy. Typically, radio access network (RAN) accounts for about 60 percent of the power used at a mobile site. Data-traffic loads are intermittent, though, so that different parts of RAN can be put briefly into sleep mode, even during periods of peak traffic. ...AI expands the potential for such energy-saving opportunities across the network.”

On the other side, it is important to notice that, as the ICT applications are more complex, which is the communications networks case, the control and optimal decisions are more difficult, because they need to analyse huge amount of data (just a part of DD) in order to extract the useful information for efficacy:

“The ability to analyse vast amounts of data relating to traffic patterns, real-time demand, and network-resource availability allows for quick, automated decisions on the parts of the system that can be put into sleep mode or shut down. For example, this could involve shutting down frequency carriers or shutting down a site momentarily in areas where there is overlapping coverage. We estimate that such energy-conserving AI tools can deliver 5 to 7 percent savings for some operators, in addition to savings that accrue from stand-alone, site-level efficiency measures. And the potential will surely rise further. As open and cloud native approaches to building RAN gain ground, additional AI solutions that not only save energy but also minimize related customer-experience issues, such as latency, are emerging. They can also be used on all networks, from 2G to 5G...The AI tools for managing energy efficiency in data centres (where the cloud RAN will be located) are considerably more mature. Google, for example, has reported 30 percent energy savings using AI at its data centres, illustrating just how high telecom operators could set their sights.”

More than these, the previous mentioned large scale integrated IoT-Edge-Cloud system finds here a natural confirmation that reflects the emergent reality of the ICT complex and eventually integrated (AI and IoT) subsystems:

“...By placing sensors at various points to gauge grid-power input, fuel levels, the number of hours the generation set has been running, battery voltages, and consumption by different types of equipment and then analysing the data, operators could uncover potentially costly anomalies. What’s more, the IoT platform could provide real-time alerts when they

occur. Like AI, the IoT also makes it much easier to optimize consumption—with air conditioners, for example. Many operators have already moved heat generating equipment outdoors to take advantage of natural convection cooling. Nevertheless, our analysis shows that, on average, about 20 percent of a telecom operator's sites and other facilities still use air conditioning. On mobile sites where this is the set case, air conditioning accounts for 30 percent of energy consumption. Here, the installation of sensors would allow monitoring and remote adjustment of the site temperature. The sensors could even alert the operations centre if a door was left ajar or the maintenance crew altered the setting, forcing the air conditioning to work unnecessarily hard. The overall result could be a 1 to 3 percent reduction in a site's energy consumption.”

Finally it is also confirmed our repeated opinion [12][16][19][21][23], that only a systemic approach, at Earth ecosystem scale, could provide the desired optimal solution that, in the globalized World, would bring benefits and sustainable (survival) progress for humankind:

“...Costs are not the only concern, however. Telecom operators already account for 2 to 3 percent of total global energy demand, often making them some of the most energy-intensive companies in their geographic markets. As operators' energy consumption expands, so will their carbon footprint, hurting not just the environment but also their reputation and standing, particularly among the expanding class of socially responsible investors.”

We have to recall here that the global ICT carbon footprint could exceed all the contribution of all airplanes flights and for worst, Moore Law consequences could determine also a higher increasing rate when compared with airplanes [6][17][9][23].

This way, we have just arrived to a point where it is clear that the **dynamic and semantic search for a sustainable balance between generating data and extracting knowledge**, has in fact global implications in the ICT/IS/KBS context, as the above (mainly technical) elements only (partially) showed the complex mechanisms and soft/hard approaches that could lead to a balance, but a further analysis could reveal that ... **there is more to learn toward optimum.**

2. Learning how to better learn ... for wisdom progress and survival

When generally speaking about learning, one could think about reading or arithmetic, other could imagine a foreign language and so on. All these are something difficult, to remember anyway, for every human age.

In the AI age of ICT, people could think to teach robots to learn human behaviour and still, perhaps here we are approaching something rarely we hear for humans: **learn to learn.**

Even so, we could remember that, sometimes, we are telling to our students (on communications engineering) that **the faculty could not teach you everything, but rather we wish to teach you to learn, in this overwhelming and fast changing world of ICT!**

Along with an amazing and exponential expansion, we witness today how the complexity of ICT applications, systems, products and services are increasing with an unprecedented speed, spreading all over the World of things or beings, as another air, sometimes perfumed, sometimes strange or even almost difficult to breath.

With other words, in all we see, all we built, all we use and all we dream, the technology (mostly ICT) is present and influence our lives, our future and generally humankind and Earth evolution.

Such context is nearly natural, after centuries and lately decades of impressing progress, but the processes complexity, pace and consequences are more and more difficult to understand, to control and eventually to forecast in the available shorter and shorter time to react.

One could simply ask: Why ICT (and no other) is mostly worrying us and what this has to do with learning?

It is largely agreed that no other technological field has the induced power, penetration, speed and complexity on IS/KBS and human life, as it is also confirmed by [7]:

<< “*The scope of electrical engineering has been growing continuously through the years, but so too has the depth of complexity and required knowledge across this ever-larger landscape. There are many more highly trained engineers worldwide now than there were a few decades ago, so new applicable knowledge accumulates at a faster pace, while it seems that older, irrelevant knowledge leaves the field more slowly. There is more to know, and it is more demanding and complex. I took a cursory look, for example, at the mathematics in information theory or the physics in quantum computing or in electronic and optical devices, and I said to myself, “I could do this if I wanted, but I choose not to.”* >>

Perhaps one of the less discussed dimension/criteria of those ICT evolutions is the emotional impact on human life and behaviour, which ultimately could also induce, on short or long time, important changes, that we not really or completely perceive now:

“*I remember when engineering seemed much simpler. As I write this, I’m noticing across the room the little blinking node on my Wi-Fi mesh network. Not so long ago AM/FM radios and TVs were the ubiquitous home electronics exploiting the electromagnetic spectrum. Those were the days when radios were just simple devices whose only standardization was the frequency band and a fairly simple modulation scheme. Take a look, though, at the specification for the IEEE 802.11ac family of Wi-Fi transceivers. There are many, many pages of complicated minutiae needed to describe the protocols, controls, signal formats, and so forth. Then, if we dig down deeper, the math for multiuser MIMO (multiple-input, multiple-output) needs to be understood. The encryption is based on elliptic curve cryptography—try checking that out. Everywhere we look at increasing depth, the complication turns to turgid complexity. That is the world we live in now.*”

This expressive introspection, in *the world we live in now*, is not only an answer to the above question, but also a bridge to the next, but more difficult, question we have to answer or least to think to, aiming criteria like **wisdom progress and survival**:

“*How is it that we can survive and indeed thrive in such a world? The obvious answers are specialization to narrow the field, and the layering of knowledge to reduce complexity to only that needed for a given purpose. In this latter pursuit we are aided by software that empowers the user while hiding the underlying complexity. I think that most engineers seldom will have the need, or the privilege, to examine the complex math and physics beneath so much of what we do...So, am I kidding myself that I could do this stuff if I chose to? I hope not, but I can’t be sure. The problems are time and motivation. I have the engineer’s curiosity and I want to know everything, and it bothers me when I don’t truly*

know how something works. But there is limited time and so very, very much to know. Sadly, I have to choose, and the choices are influenced by motivation. There are things that I have to know, and things that it would be nice to know. Often, the former overwhelm the latter. The way complexity is increasing, I imagine that in the future, computers will be writing papers for the IEEE societies' Transactions, but those papers will be so complex that only other computers will be able to understand them. A lot of the fun will be gone."

Well, our opinion, pointed in the section title, is that **we have to learn how to better learn**. Of course, one could say: Easier said than done!

We have repeatedly expressed [9][16][21] that all we see or perceive on these ICT/DD evolutions is not simply to understand nor to manage, but we have to timely analyse and try to optimize, as the price is so big.

Now, for having more chances to reach this goal, we can see that the complexity obliges us to also change the rules of learning we are used to practice, because the old rules could not be completely useful in the deep and ever-changing context.

It is not a paradox, as all the above proved, that, in spite of DD, it is more difficult or even impossible to use old models of extracting knowledge while the semantic search with new criteria is a good start, but ... far from being all!

A first difficult issue is to optimally and dynamically refine the criteria and the models of our thinking, in order to rapidly adapt to the new realities.

In order to be fair, we must emphasize that the World realities exceed the ICT/DD context, as we can see in these days, when the main challenges of our Earth ecosystem seem to come from other side (human and social health), but we also have to agree that any dramatic evolution at Earth scale have to be seen in the globalization context and we must use the power of ICT/DD/AI to optimally face the challenges.

Here we may come back and try to better learn, we humans and our smart machines, from the dramatic changes reflected also in the ICT/DD potential.

When we start to better learn with the ICT/AI/DD help, we first have to analyse/perceive their features, limits and vulnerabilities, especially on estimating the complex context futures [8]:

<< *"Twelve years ago, IEEE Fellow, Marconi Prize winner, and beloved Spectrum columnist Robert W. Lucky wrote about the difficulty of predicting the technological future. We've reprinted his wise words here... Why are we engineers so bad at making predictions? In countless panel discussions on the future of technology, I'm not sure I ever got anything right... How can something like Google feel so inevitable and yet be impossible to predict? I'm filled with wonder at all that we engineers have accomplished, and I take great communal pride in how we've changed the world in so many ways. Decades ago I never dreamed we would have satellite navigation, computers in our pockets, the Internet, cellphones, or robots that would explore Mars. How did all this happen, and what are we doing for our next trick? The software pioneer Alan Kay has said that the best way to predict the future is to invent it, and that's what we've been busy doing. The public understands that we're creating the future, but they think that we know what we're doing and that there's a master plan in there somewhere. However, the world evolves haphazardly, bumbling along in unforeseen directions. Some seemingly great inventions just don't take hold, while overlooked innovations proliferate, and still others are used in unpredicted ways.>>*

We also agree with these pertinent observations, which have the quality to express, from detailed arguments to common sense conclusions, the big picture of ICT evolution and challenges.

Concretely, this is a confirmation of our purpose necessity of analysing these huge processes, where *world evolves haphazardly*, in order to estimate/control their *unforeseen directions*, as eventually to avoid the actual undesired trends (*great inventions just don't take hold, while overlooked innovations proliferate, and still others are used in unpredicted ways*) and leverage the long expected progress in areas/criteria like health, social imbalance and environment protection, also considering the climate changes and Earth resources fading.

Consequently, such conclusions must be used for the **mentioned criteria of optimization**, along with other improvements of learning from the real-time complex processes we face in ICT/IS/KBS.

An interesting example for the challenges of striving to better learn from everything could be observed just in the most prominent processes of learning at planetary scale, by ...ML/DL.

Here we have to notice that, although we speak about how **the machines are learning**, this could reveal not only the state-of-the-art of ICT/AI/ML/DL/CL and their limits, but indirectly shows the limits of humans that have the responsibility to design them, just considering the above mentioned criteria.

A good example for this features and ICT/AI/ML models of working and learning from DD is given by [10]:

“According to A Few Useful Things to Know about Machine Learning, while the discipline by itself is far from simple, it is based on a simple (but not simplistic) principle:
LEARNING = REPRESENTATION + EVALUATION + OPTIMIZATION

Where:

- *Representation means the use of a classifier element represented in a formal language that a computer can handle and interpret;*
- *Evaluation consists of a function needed to distinguish or evaluate the good and bad classifiers; and*
- *Optimization represents the method used to search among these classifiers within the language to find the highest scoring ones.”*

Without entering the sophisticated details of ML/DL/CL algorithms, we have to observe, here transferred, the same method of optimizing the search, we also have approached in this paper, aiming the best classifiers/criteria.

Further, we can see how the most difficult issue of the process is solved, with human contribution (supervised strategy):

“In this context, machine learning can be done by applying specific learning strategies, such as:

- *A supervised strategy to map the data inputs and model them against desired outputs, and*
- *An unsupervised strategy, to map the inputs and model them to find new trends.*

Derivative ones that combine these for a semi-supervised approach and others are also be used. This opens the door onto a multitude of applications for which machine learning can be used, in many areas, to describe, prescribe, and discover what is going on within large volumes of diverse data...Thanks to the success of the application of machine learning within

certain disciplines such as speech recognition, computer vision, bio-surveillance, and robot control, the interest in and adoption of machine language technologies has grown, particularly over the last decade ... There are several scenarios where machine learning can have a key role: in those systems that are so complex that algorithms are very hard to design, or when an application requires the software to adapt to an operational environment, or with complex systems that need to work with extensive and complex data sets. In this way, machine learning methods play an increasing role not just in general in the field of computer science, but also in terms of enterprise software applications, especially for those types of applications that need in-depth data analysis and adaptability. These areas include analytics, business intelligence, and Big Data.”

From here we can observe not only the huge diversity of emergent applications of such methods of “learning”, but also the fact that their increasing complexity ask the human (supervised) contribution, as we must aim to “*model them against desired outputs*”.

If our above idea could seem not very clear, there are realistic and rigorous analyses that point just the deep implications of any error (see critical infrastructures) in the huge applications that ICT/AI/ML/DL/CL will spread over the World [11]:

“First-timers are often surprised by how little time in a machine learning project is spent actually doing machine learning. But it makes sense if you consider how time-consuming it is to gather data, integrate it, clean it and pre-process it, and how much trial and error can go into feature design...Generalization being the goal has another major consequence: data alone is not enough, no matter how much of it you have. Consider learning a Boolean function of (say) 100 variables from a million examples. There are $2^{100} - 10^6$ examples whose classes you don’t know. How do you figure out what those classes are? In the absence of further information, there is just no way to do this that beats flipping a coin. This observation was first made (in somewhat different form) by the philosopher David Hume over 200 years ago, but even today many mistakes in machine learning stem from failing to appreciate it. Every learner must embody some knowledge or assumptions beyond the data it’s given in order to generalize beyond it. This was formalized by Wolpert in his famous “no free lunch” theorems, according to which no learner can beat random guessing over all possible functions to be learned...”

Our opinion is that, for **learning to better learn**, the emphasis must be secondary on the fact that **the human contribution to ICT/AI/ML/DL/CL processes is needed for better results “against” the difficulty/complexity**, but primary in order to be sure that the **“desired outputs”** are designed to match the priority criteria we have above mentioned as essential for a sustainable development of ICT/IS/KBS and humankind evolution on Earth.

More than these, **learning to better learn**, for a systemic approach of refining knowledge, is necessary to be the goal of all responsible people, beginning with the incumbents, the specialists and teachers, as everybody has to adapt to changes (*Intelligence is the ability to adapt to change-Stephen Hawking*) and use new technologies with appropriate criteria of optimization (mentioned above), i.e. timely analyse **what and how we should learn or ignore**.

3. Conclusions

The paper analysed the DD/ICT/IS/KBS premises and main approaches toward a ***dynamic and semantic search of balance between generating data and extracting knowledge***, focusing on some semantic criteria of the search, using ICT/DD context at the planetary scale, which could influence Earth environment and eventually humankind evolution with an unprecedented speed and impact. The main reasons for this kind of analyses come from the ICT generated global and exponential changes, which are more and more difficult to be controlled or at least estimated, regarding all their positive or negative consequences, just because of the complexity and speed of their vast penetration.

The semantic analysis have to face a dynamic environment, because the relevant data/facts are fast changing (not from centuries or decades, but from months or even days), so we have to refine our search looking for new semantic terms or changing the old meaning (associated information/ knowledge) of the usual ones.

On this line it is emphasized that, although all technologies and ICT applications largely contribute to DD, lately, AI and its main branches, machine learning (ML), deep learning (DL) or cognitive learning (CL), are prominent in striving to extract information/knowledge from data, or even to compete with the human intelligence by cognitive features

The identified criteria of the search are linked with all (positive or negative) consequences of ICT/IS/KBS and generally of World changes, just because of the complexity and speed of their vast proliferation, more visible by the dramatic climate changes, Earth resources fading, social imbalance or actually by the vulnerability of humankind to diseases spreading in the globalization context. **Such consequences should be alarm signals and optimization criteria** in our decision processes, as individuals, but mainly for the incumbents. These signals and the ICT impact on IS/KBS at planetary scale could indicate **the main criteria for the mentioned semantic search** in the ICT/AI/ML/DL/CL fast implementation processes, **in order to meet Earth ecosystem needs and expectations for a sustainable progress.**

An important conclusion is that **the semantic approach of decision processes, based on the generated data and extracted information, is a complex optimization target which just could match and benefit from the dynamic ICT/DD/ML/DL/CL context, where the data signification, implications and consequences are fast appearing and changing, because of this context potential to produce changes, but we have to be sure which of them are positive and which negative, to what extent and time period.** The most difficult and complex problem will be, every time, to **understand and define criteria for what is positive and what negative, as all our Earth ecosystem and even humankind behaviour and evolution is changing, but also our criteria should!**

Such conclusions are sustained by some concrete examples of the massive expansion of ICT, like the case emergent 5G (among others), which could produce, besides positive, real and unprecedented benefits, new, but less desired effects. Here, the good news is that AI/ICT itself could compensate such negative effects, as it is revealed also by some concrete applications cases.

The second section of the paper completes the previous technical elements with other improvements of learning from the real-time complex processes of ICT/IS/KBS that could be observed just in the most prominent processes of learning at planetary scale, by ...ML/DL. Although we speak about how **the machines are learning**, these could reveal not only the state-of-the-art of ICT/AI/ML/DL/CL and their limits (for example, ML and *Wolpert famous “no free lunch” theorem*), but indirectly shows the limits of humans that have the responsibility to design them, just considering the mentioned criteria.

Another conclusion is that all we see or perceive on these ICT/DD evolutions is not simply to understand nor to manage, but we have to timely analyse and try to optimize, as the price is so big. In fact, the complexity obliges us to also change the rules of learning we are used to practice, because the old rules could not be completely useful in the deep and ever-changing context, i.e., optimally and dynamically refine the criteria and the models of our thinking, in order to rapidly adapt to the new realities, which exceed the ICT/DD context, as we can see in these days, when the main challenges of our Earth ecosystem seem to come from other side (human and social health), but we also have to agree that any dramatic evolution at Earth scale have to be seen in the globalization context and we must use the power of ICT/DD/AI to optimally face the challenges (including the emotional impact on human life and behaviour, which ultimately could also induce, on short or long time, important changes, that we not really or completely perceive now).

Finally, our conclusion is that, for **learning to better learn**, the emphasis must be secondary on the fact that **the human contribution to ICT/AI/ML/DL/CL processes is needed for better results “against” the difficulty/complexity, but primary in order to be sure that the “desired outputs” are designed to match the priority criteria we have above mentioned as essential for a sustainable development of ICT/IS/KBS and humankind evolution on Earth.** More than these, learning to better learn, for a systemic approach of refining knowledge, is necessary to be the goal of all responsible people, beginning with the incumbents, the specialists and teachers, as everybody has to adapt to changes (*Intelligence is the ability to adapt to change-Stephen Hawking*) and use new technologies with appropriate criteria of optimization, i.e. timely analyse what and how we should learn or ignore.

REFERENCES

- [1] Vaclav Smil, *Data World: Racing Toward Yotta*, IEEE Spectrum (Volume: 56 , Issue: 7 , July 2019).
- [2] J. Macfarlane, *When apps rule the road: The proliferation of navigation apps is causing traffic chaos. It's time to restore order*, IEEE Spectrum 56(10):22-27 · October 2019.
- [3] Victor Greu, *Using the information and communications technology data deluge from a semantic perspective of a dynamic challenge: What to learn and what to ignore?* – (Part 2), Romanian Distribution Committee Magazine, Volume 10, Issue 4, Year 2019.
- [4] Shuai Zhang, Wenxi Zeng, I-Ling Yen, Farokh B. Bastani, *Semantically Enhanced Time Series Databases in IoT-Edge-Cloud Infrastructure*, 2019, <https://arxiv.org/pdf>
- [5] Richard Lee, Dickon Pinner, Ken Somers, and Sai Tunuguntla, *The case for committing to greener telecom networks*, February 2020, McKinsey&Company, Technology, Media & Telecommunications Practice.

- [6]Victor Greu, *Extending information and communications technologies's impact on knowledge based society through artificial and collective intelligence –(Part 3)*, Romanian Distribution Committee Magazine, Volume 9, Issue 3, Year 2018.
- [7]Robert W. Lucky, *Deep Complexities in EE*, IEEE Spectrum, May 2018.
- [8]Robert W. Lucky, *Back To The Elusive Future We looked back for inspiration on looking forward*, IEEE Spectrum, Jan 2020.
- [9]Victor Greu, *The information and communications technology is driving artificial intelligence to leverage refined knowledge for the World sustainable development – (Part 2)*, Romanian Distribution Committee Magazine, Volume 10, Issue 1, Year 2019.
- [10]Jorge Garcia, *Machine Learning and Cognitive Systems: The Next Evolution of Enterprise Intelligence*, 2014, <https://www.wired.com/insights/2014/07/machine-learning-cognitive-systems-next-evolution-enterprise-intelligence-part>.
- [11]Pedro Domingos, *A Few Useful Things to Know about Machine Learning*, <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>
- [12]Victor Greu et all, *Human and artificial intelligence driven incentive-operation model and algorithms for a multi-purpose integrated crowdsensing-crowdsourcing scalable system* - paper submitted to International Conference Communications 2018 (Politehnica University of Bucharest, Romania Military Technical Academy, IEEE Romania), June 2018.
- [13]Bhupinder Kour, *The Rise of Machine Learning and AI is Improving Lives in 2018*, <https://www.smartdatacollective.com/rise-of-machine-learning-ai-improving-lives/>
- [14]Seth Earley, *There Is No AI Without IA*, 2016, IEEE IT Professional (Volume: 18, Issue:3, May-June 2016)
- [15]Niko Mohr, Holger Hürtgen, *Achieving business impact with data - A comprehensive perspective on the insights value chain*, Digital McKinsey 04.2018 Copyright © McKinsey & Company www.mckinsey.com
- [16]Victor Greu, *Searching the right tracks of new technologies in the earth race for a balance between progress and survival*, Romanian Distribution Committee Magazine, Volume 3, Issue1, Year 2012.
- [17]Mike Sirius, *How artificial intelligence and innovation will interact*, June 5, 2017, Idea Drop Ltd, <http://ideadrop.co/artificial-intelligence-innovation/>
- [18]Rohit Akiwatka, *Introduction to cognitive computing*, may 2019, <https://channels.theinnovationenterprise.com/articles/introduction-to-cognitive-computing>
- [19]Victor Greu, *The Exponential Development of the Information and Communications Technologies – A Complex Process Which is Generating Progress Knowledge from People to People*, Romanian Distribution Committee Magazine, Volume 4, Issue2, Year 2013
- [20]***, *Transforming data into knowledge*, Collective Learning group at the MIT Media Lab, 2017, <https://www.media.mit.edu/groups/collective-learning/overview/>.
- [21]Victor GREU, *Information and Communications Technologies are Learning from Nature's "Research" to Push the Performance Limits*, Romanian Distribution Committee Magazine, Volume 5, Issue 1, Year 2014
- [22]Mudassir Khan, *Big data analytics emerging trends, technology and innovations for the future business in the global market*, International Journal of Scientific Research and Review, Volume 8, Issue 2, 2019.
- [23]Victor Greu, *Information and communications technologies go greener beyond IOT-behind is all the earth-Part1*, Romanian Distribution Committee Magazine, Volume 7, Issue 2, Year 2016.