

Information and Communications Technology is Merging Data Science and Advanced Artificial Intelligence Towards the Core of Knowledge Based Society

-Part 2-



Prof. Eng. Ph.D. Victor GREU

Abstract

The paper approaches the complex processes of the information and communication technology (ICT) and Knowledge Based Society (KBS) context, leading to the Data Science (DS) concept, aiming to reveal the essential links between the DS, advanced artificial intelligence (AI) and refining knowledge at Earth ecosystem scale as the core of KBS.

These premises, due to the scientific, technical and economical reasons of optimizing the mentioned processes, further link the ICT/KBS context with the necessary concept of Cyberinfrastructure, we consequently have to approach. We analysed this concept keeping in mind our title point on knowledge (DS, advanced AI and core of KBS), aiming to emphasize its main components, complex interactions and mutual-supporting mechanisms, which are usually shadowed by the well-known resonance of the term infrastructure.

Based on appropriate references for confirmation, we showed that, beyond the usual approach that involves research activities, the Cyberinfrastructure should contain people and scholarly activities more generally, including teaching and learning, i.e., the whole education fields, which in the Big Data era became not only important, but part of the most difficult processes of decision. Another finding is that ICT advances impact not only the humankind activity fields, but also the people's thinking, generally for understanding all practical implications of using AI/ICT features/applications, but, essentially, we consider them an expression of the golden rule AI/ICT have to implement in order to maintain its race toward better performance: re-invent itself all the time by leveraging people's creation skills and thinking. This goal is double fold, aiming apparently to leverage user capacity to fully exploit the designed AI (machine learning algorithms) potential, but eventually this will increase experts' capacities to use their imagination in order to improve (re-invent) AI instruments performance.

Consequently, the complex (mutual-supporting) relations between Cyberinfrastructure components are approached from the human side, considering their crucial role in the actual World context where USA and EU agreed on the strategic importance of the acquired technological advances (i.e., supposing the associated economic, political and social consequences, but also the cyber-security). It resulted that the need for highly skilled people raised the importance and role of the human component for Cyberinfrastructure and then for AI/ICT/KBS, but the complex relations between them represent the most complicate problem to be further considered, by the need to timely analyse all the consequence of ICT exponential development and impact at Earth scale and on human behaviour and evolution, approaching the ICT-human complex relations and impact on KBS.

Continuing to follow the ICT rule of re-inventing itself, we consider that, beyond the new hard/soft performant technologies, in the future we have to look for other leveraging ideas, mainly on improving the output of the computer-human relation by refining knowledge by more subtle criteria of performance. In a concrete manner, the perpetual informational transition where the engineers have to work, is updated to the actual computer-human relation/technology, where the knowledge is refined. The knowledge exchange between researchers is in fact the area that pushes the World ahead with the greatest impact, but also with the most of generation and refining problems, which are worth to be analysed in order to identify prominent mechanisms that could contribute to the optimization of AI/ICT/KBS development in the Big Data era, taking into account the actual challenges and expectations which are naturally associated with the unprecedented advances already reached, but difficult to be improved. Knowing those problems could contribute to the configuration of the actual and the emerging context of developing AI/ICT/KBS, providing this way a useful base for the analyses/decisions, irrespective how these could/will be. One of the main mechanisms of progress, for AI/ICT/KBS, is refining knowledge at Earth ecosystem scale, which is essentially depending on human capacity of creation and integration of the advanced technology in new models, with a larger view of the complex and dynamic processes of generating/perceiving data, information and knowledge in the emerging World of KBS, where the relativity and the probabilistic computing approach [19] are inherent features.

We concluded that understanding the unlimited dynamic AI/ICT/KBS optimization processes is a long and winding road, full of surprises, we have to face with timely analyses and ... imagination.

Keywords: cyberinfrastructure, explainable artificial intelligence, cyberspace, computer-centred humans, knowledge based society, big data, knowledge refining

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

“Artificial intelligence will not destroy this planet, irresponsible human intelligence will”
— Abhijit Naskar

1. From cyberinfrastructure to knowledge

Even at their early steps, humans faced the decisions processes and challenges, but, naturally, the difficulty of such problems could become overwhelming in a probable future, as their complexity and implications increase exponentially in the context created by the society general development, where information and communication technology (ICT) became the main driving factor of the human society by the complex consequences of ICT services, products and applications, when supporting the Information Society (IS) on its way toward the Knowledge Based Society (KBS)[3].

Although this ICT/KBS context features seem evident, it is interesting to observe that the difficulty of decisions increases not only with the complexity and the importance of the problem, but also with the contextual information level/uncertainty involved each time.

Now, the role of the ICT advances, where artificial intelligence (AI) is by far the most expected instrument for data/information high level processing in the data deluge and Big Data context, became essential and led to the Data Science (DS) concept [3].

Based on these observations, the main parts of our title point (DS, advanced AI and core of KBS) could be associated in fact with complex processes of the **ICT/KBS context**, in a **circle evolution and mutual-supporting relations**, with crucial and complicate consequences (mainly benefic) for the Earth ecosystem.

These premises could explain why, due to the scientific, technical and economical reasons of optimizing the mentioned processes, many references perceive this **ICT/KBS context** linked with the necessary concept of *Cyberinfrastructure* [2][5][8][16].

The reason for which we also consider important to analyse this concept, keeping in mind our title point on knowledge (DS, advanced AI and core of KBS), is to emphasize its main components, complex interactions and ***mutual-supporting*** mechanisms, which are usually shadowed by the well-known resonance of the term *infrastructure*.

The extent and the importance, for the link between the AI/CT/KBS context and the stage of what cyberinfrastructure could provide, are also revealed by an actual example [1]:

“IBM named a Leader *Gartner releases 2021 Magic Quadrant for Data Science and Machine Learning Platforms...*

IBM Watson® Studio empowers data scientists, developers and analysts to build, run and manage AI models, and optimize decisions anywhere on IBM Cloud Pak® for Data. Unite teams, automate AI lifecycles and speed time to value on an open multicloud architecture.”

Even the condensed (by the density of black boxes/applications) style of the above text shows something (like an iceberg tip!) about the undergoing complexity that could support large and promising fields of human activity, by *automation of AI and speed time*.

On the other hand, a systemic approach, for such complex and important concept (*cyberinfrastructure*), should begin from definitions.

A good point of starting is to use some reference examples for revealing the content, limits and even challenges of *Cyberinfrastructure* definitions [2]:

“Cyberinfrastructure consists of computational systems, data and information management, advanced instruments, visualization environments, and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible...”

The value of any definition lies in its ability to identify that which it calls out and to distinguish those entities that do not meet the definition. The above definitions of cyberinfrastructure are helpful in that they each have three distinct elements relating to technology, people, and effect...”

Here we have to notice that beyond the usual approach that involves *research activities*, we have to consider *people* and *scholarly activities* more generally, including teaching and learning, i.e., the whole education fields, which in the Big Data era became not only important, but part of the most difficult processes of decision [5][7][8][10][6][3][9][15].

More than this general and comprehensive definition, the links and implications of *Cyberinfrastructure* in KBS are further revealed by quoting *Fran Berman*, director of the *San Diego Supercomputer Center*, as [2]:

“The component parts of Cyberinfrastructure are human, software, hardware, instrument, and other resources coordinated so as to interoperate “end-to-end” and to support multiple users simultaneously. At scale, this complex structure will need to involve appropriate user incentive structures, effective organizational frameworks, policy and privacy constraints, and a wealth of other social mechanisms to ensure stability, performance, and usefulness”

This way we have just arrived to the top of the picture for the systemic approach, confirming our paper title point, as it is clearly expressed that the high level of analyses and decisions, for this complex AI/CT/KBS context, is including *effective organizational frameworks, policy and privacy constraints, and a wealth of other social mechanisms to ensure stability, performance, and usefulness*.

Approaching the “iceberg” inside, i.e., the *Cyberinfrastructure* components and their complex relations, we can make it, essentially, in two steps.

First, considering the hard/soft resources and instruments, we could extend the above prominent example of *IBM Watson Studio*, now having a larger perspective and entering one of the most complex and relevant of its features (*explainable AI*), with impact on AI development and performances [1]:

“Explainable AI is a set of processes and methods that allows human users to comprehend and trust the results and output created by AI algorithms, including its expected impact and potential biases...”

Explainable AI is used to describe an AI model, its expected impact and potential biases. It helps characterize model accuracy, fairness, transparency and outcomes in AI-powered decision making. Explainable AI is crucial for an organization in building trust and confidence when putting AI models into production. AI explainability also helps an organization adopt a responsible approach to AI development.”

With other words, here we could witness one of the subtle ways ICT advances impact not only the humankind activity fields, but also the people’s thinking, as we repeatedly expressed [11][3].

Concretely, we notice that one of the prominent features/applications of *IBM Watson Studio* is oriented by design toward helping users (including the most skilled of them, i.e., the AI experts!) to understand all practical implications of those *black box*-s we above remarked:

“As AI becomes more advanced, humans are challenged to comprehend and retrace how the algorithm came to a result. The whole calculation process is turned into what is commonly referred to as a “black box” that is impossible to interpret. These black box models are created directly from the data. And, not even the engineers or data scientists who create the algorithm can understand or explain what exactly is happening inside them or how the AI algorithm arrived at a specific result.

There are many advantages to understanding how an AI-enabled system has led to a specific output. Explainability can help developers ensure that the system is working as expected, it might be necessary to meet regulatory standards, or it might be important in allowing those affected by a decision to challenge or change that outcome.”

About the benefits of these features/applications, there is a lot to say, but, essentially, **we consider them an expression of the golden rule AI/ICT have to implement in order to maintain its race toward better performance: re-invent itself all the time.**

In fact, here we just switched to the second step of the above-mentioned systemic approach, as, after considering *Cyberinfrastructure* hard/soft resources and instruments, we point the human component, because, in the IBM example, to implement *explainable AI* means to leverage the human potential.

It is very important to observe that this goal is double fold, aiming apparently to leverage user’ capacity to fully exploit the designed AI (machine learning algorithms) potential, but eventually this will increase experts’ capacities to use their imagination in order to improve (**re-invent**) AI instruments performance. On the other hand, the complex (mutual-supporting) relations between *Cyberinfrastructure* components are also revealed.

In order to conclude the short approach of *Cyberinfrastructure*, the top of the picture for the systemic approach, we just mentioned above, should also be confirmed at least by one

more reference that points similar *effective organizational frameworks, policy and privacy constraints, and a wealth of other social mechanisms to ensure stability, performance, and usefulness.*

More than these, **observing the prominent news of our days and trying to see even beyond Covid 19 pandemic or regional crises, we could notice that both EU and USA concluded that the future of the World will strongly depend on the acquired technological advances and their economic, political and social consequences, but the cyber-security was also pointed.**

Here it is worth to recall a reference that is not only relevant and actual in spite of the passed years, but also backed by prestigious organizations [4]:

"One way to think of the Information Technology revolution is to think of cyberspace as a new continent - equivalent to discovery of the Americas 500 years ago. Cyberspace is transforming the old world with new goods and services. It is changing the way we learn, work and play...."

Since long-term research is a public good, it needs to be funded as such: making all the boats rise together. That is why funding should come in part from society: industry is paying a tax by doing long-term research; but, the benefits are so great that society may want to add to that, and fund university research. Funding university research has the added benefit of training the next generations of researchers and IT workers. I do not advocate Government funding of industrial research labs or government labs without a strong teaching component..."

Although this arguments for *Cyberspace funding* support could appear utopian versus the actual mentioned World context, further there are given arguments which must also be considered, at least for an accurate analysis:

"...One might argue that US Government funding of long-term research benefits everyone in the world. So why should the US fund long-term research? After all, it is a social good and the US is less than 10% of the world. If the research will help the Europeans and Asians and Africans, the UN should fund long-term research.

The argument here is either altruistic or jingoistic. The altruistic argument is that long-term research is an investment for future generations world-wide. The jingoistic argument is that the US leads the IT industry. US industry is extremely good at transforming research ideas into products – much better than any other nation.

To maintain IT leadership, the US needs people (the students from the universities), and it needs new ideas to commercialize. But, to be clear, this is a highly competitive business, cyberspace is global, and the workers are international. If the United States becomes complacent, IT leadership will move to other nations."

It is also obvious that the example of the USA could be considered not of universal relevance, but it is still the most complex case and this way important for the World, as it reveals almost all challenges for the complicate problem of *Cyberinfrastructure funding* approach.

In addition, the need for highly skilled people (including *students from the universities*) is pointing again **the crucial importance and role of the human component for Cyberinfrastructure and then for AI/ICT/KBS, but the complex relations between them represent the most complicate problem to be further analysed.**

2. Evolutions to knowledge-centered computers and humans

In our days almost nothing could surprise when speaking about ICT evolutions, including the ways humans are involved, as users or creators of the new products, applications or services emerging from ICT advances like AI, cloud, Big Data, IoT, 5G etc.

Still, the idea of Computer-Centered Humans [5] seems a little shocking and we also repeatedly expressed the need to timely analyse all the consequence of ICT exponential development and impact at Earth scale and on **human behaviour and evolution** [18][9][11].

Fortunately, in a World of Data Deluge, we have to see the details before jump into conclusions, but, in this case, Computer-Centered Humans surely could inspire for a new and useful approach of the ICT-humans complex relations and impact on KBS.

Perhaps the best start is to enter the original details and then to observe the issues associated with the general approach of optimizing the AI/ICT/KBS context and development [5]:

“The technology development of Artificial Intelligence (AI) in the U.S. Department of Defense (DoD) is proceeding at an unprecedented pace, with record breaking levels of funding to support unprecedented breakthroughs. As these technologies progress through technology readiness levels and make their way into the hands of human beings, however, the need for human-centered design practices will become more evident.

It is difficult to overstate just how prevalent discussions involving AI have become across the entire spectrum of warfighting and defense capabilities. But while the speed of research in the technology of AI is clearly increasing, research in human-AI interaction does not appear to be keeping pace. Issues such as usability, interaction modalities, visualization, and knowledge representation techniques are vital parts of a coherent technology integration strategy. While it is critical that the United States keep pace with its near-peer allies and adversaries by developing advanced technologies, it is also vital that these technologies are developed in ways that humans can understand and use appropriately...”

Beyond the defense use of AI, we have to notice again, as we earlier mentioned [11], the necessity to follow the ICT pace (*human-AI interaction does not appear to be keeping pace*) with the appropriate analyses of its consequences.

It is also worth to observe, as being written in a more subtle expression, the importance of using **AI in the critical infrastructures/applications** by humans (*way into the hands of human beings*) and this way revealing why this example is relevant beyond the defense field:

“...two principal reasons why research in the psychology of human-AI interaction and human-centered design work needs to keep pace with technology development in the DoD: to ensure AI systems that are safe and reliable for humans to use, and to ensure these systems are integrated in ways that do not inject new kinds of error and risk to existing systems.”

Although it appears that here the point is mainly on humans/users safety, it is clear that the impact of these powerful technologies is way beyond, covering strategic issues (*error and risk to existing systems*) at Earth scale.

This way our systemic analysis is starting again from the top (as we have mentioned above, when expressing the strategical options for **acquiring technological advances**), explaining also why we have selected a defense example for approaching the ICT-humans complex relations and impact on KBS.

Continuing to follow the ICT rule of re-inventing itself, we consider that this example showed that, beyond the new hard/soft performant technologies, in the future we have to look for other leveraging ideas, mainly on improving the output of the computer-human relation by refining knowledge by more subtle criteria of performance.

An integrative approach of this idea is well reflected below, as it also presents face to face the pros and the cons of refining knowledge from ... old knowledge [7]:

“I was listening recently to some engineering graduates talking about their current research efforts... These engineers were really skilled. I was greatly impressed by their depth of knowledge and facility with analysis. How did they learn so much in their few years of university training? After all, there is so much more to know now than there used to be, and every year it gets ever more overwhelming. • One explanation is that, as new knowledge accumulates, some old knowledge becomes irrelevant and falls off the knowledge stack. Almost all the college course work I took long ago is now useless in itself, although what remains is an engineering mind-set and a mathematical grounding. Perhaps every course should have a sell-by date. Indeed, in retrospect I now realize that a number of the courses I took were already well past their sell-by dates when I took them...”

Confirming the unprecedented increase of knowledge, the challenge of refining knowledge is also recognized as very complicate:

“The problem is that we never quite know when a particular course or book will become obsolete. But the purging of obsolete knowledge is probably insufficient in itself to make room for the new stuff, as there seems to be an exponential increase in knowledge. The complexity of our work is always increasing, similar to the increase in entropy decreed by the second law of thermodynamics. For many decades we’ve been driven by Moore’s Law, which has urged us to embrace and exploit complexities that were unthinkable in previous decades. Even as Moore’s Law wanes, I feel sure that the general law of exponential increase will continue the trend.”

In a concrete manner, the perpetual informational transition where the engineers have to work, is updated to the actual computer-human relation/technology, where the knowledge (accumulated by others) is refined:

“New engineers often enter fields that have become well plowed, and so find themselves pushing against physical and theoretical limitations. The issues are complex, and the incremental gains may be small... I frequently hear about engineers employing machine learning in some new and creative way. As the new engineers come out of school, they are also empowered by the continual rise of new tools. They work with networked computers, whose software embeds the knowledge and techniques of analysis and design accumulated by others. I remember when we used to wander down the halls asking other engineers how to solve some problem. Now we put our questions to the computer, and by doing so, ask all the other engineers and scientists of the networked world. You don’t have to know everything yourself, only how to ask the questions. So, the challenge is great, but these graduates are up

to it. I'm thinking, however, that there must be some value in the experience of older engineers."

We have to notice that, in spite of the pros and cons, the conclusion is pointing the crucial role of knowledge exchange between people, even in the era when we *put our questions to the computer*, because *there must be some value in the experience of older engineers*.

The knowledge exchange between researchers is in fact the area where the knowledge that pushes the World ahead has the greatest impact, but also the most of generation and refining challenges, as it is next and detailed approached [8]:

"Scientists in many disciplines have begun revolutionizing their fields by using computers, digital data, and networks to replace and extend their traditional efforts. The calculations that can be performed and the information that can be archived and used are exploding. In the not-too-distant future, the contents of the historic scientific literature will fit on a rack of disks, and an office computer will provide more computing than all the supercomputing centers together today. The results of today's largest calculations and most sizable collections will take seconds to transmit using the fastest known network technologies. New technology-mediated, distributed work environments are emerging to relax constraints of distance and time. These new research environments are linking together research teams, digital data and information libraries, high-performance computational services, scientific instruments, and arrays of sensors. In many cases these emerging environments for knowledge work are essential, not optional, to the aspirations of research."

Speaking of *emerging environments for knowledge work*, here we have to observe the clear confirmation of the crucial role of this subject, also emphasized by our paper title, as being *essential, not optional*, and the arguments for this necessity is further provided as a comprehensive picture of the future:

"...some shifts in current research practice:

- The classic two approaches to scientific research, theoretical/ analytical and experimental/observational, have been extended to in silico simulation and modelling to explore new possibilities and to achieve new precision.*

- The enormous speedups of computers and networks have enabled simulations of far more complex systems and phenomena, as well as visualizing the results from many perspectives.*

- Advanced computing is no longer restricted to a few research groups in a few fields such as weather prediction and high-energy physics, but pervades scientific and engineering research, including the biological, chemical, social, and environmental sciences, medicine, and nanotechnology.*

- Crucial data collections in the social, biological, and physical sciences are now online and remotely accessible – modern genome research would be impossible without such databases, and soon astronomical research will be similarly redefined through the National Virtual Observatory."*

The mentioned new World research trends/features could impact all science areas, but more than this, they will mainly improve the instruments, efficacy and performance of outputs:

“...The trends represented by these examples will only accelerate. In the future, we might expect researchers to:

- Combine raw data and new models from many sources, and utilize the most up-to-date tools to analyse, visualize, and simulate complex interrelations.*
- Collect and make widely available far more information (the outputs of all major observatories and astronomical satellites, satellite and land-based weather data, three-dimensional images of anthropologically important objects), leading to a qualitative change in the way research is done and the type of science that results.*
- Work across traditional disciplinary boundaries: environmental scientists will take advantage of climate models, physicists will make direct use of astronomical observations, social scientists will analyse interactive behavior of scientists as well as others.*
- Simulate more complex and exciting systems (cells and organisms rather than proteins and DNA; the entire earth system rather than air, water, land, and snow independently).*
- Visualize the results of complex data sets in new and exciting ways, and create techniques for understanding and acting on these observations.”*

Because time is the main enemy of Earth ecosystem when confronted, as we repeatedly mentioned [18][11], with hard challenges as climate changes, resource fading or social unbalances, here the conclusion is pointing the necessity to act now, based on specified benefits:

“...We believe that several key thresholds have recently been reached in the use of IT ... including networking, supercomputing, human interfaces, collaboration environments, and information management:

- Closed-form analytic solutions are available for a decreasing fraction of interesting research challenges; often only a numeric computation can produce useful results.*
- Moore’s law has led to simulations that begin to match the complexity of the real world, with fully three-dimensional, time-dependent modelling with realistic physical models opening up a vast range of problems to qualitative attacks. They range from cosmology to protein folding – problems formerly considered far too complex to address directly.*
- In an increasing fraction of cases, it is faster, cheaper, and more accurate to simulate a model than construct and observe a physical object.*
- Increasingly ubiquitous networking and interoperability of information formats and access make high-quality remote collaboration feasible.*

The practical value and the importance of the same conclusion (act now) is completed by also identifying some associated risks to be considered:

- “There are also significant risks and costs if we do not make a major move at this time:*
- Absent coordination, researchers in different fields and at different sites will adopt different formats and representations of key information, which will make it forever difficult or impossible to combine or reconcile.*
 - Absent systematic archiving and curation of intermediate research results (as well as the polished and reduced publications), data gathered at great expense will be lost.*
 - Effective use of cyberinfrastructure can break down artificial disciplinary boundaries, while incompatible tools and structures can isolate scientific communities for years...”*

Again, we consider important to stress that, in spite of the excess of optimism or utopia, such points of view have the merit of expressing a lot of the details that could contribute to the configuration of the actual and the emerging context of developing AI/ICT/KBS, providing this way a useful base for the analyses/decisions, irrespective how these could/will be.

In fact, these details identify prominent mechanisms that could contribute to the optimization of **AI/ICT/KBS development in the Big Data era**, considering the actual challenges and expectations which are naturally associated with the unprecedented advances already reached, but difficult to be improved [12][14][16][17].

As we already have mentioned, one of the main mechanisms of progress, for **AI/ICT/KBS** is refining knowledge at Earth ecosystem scale [11], which is essentially depending on human capacity of creation and integration of the advanced technology in new models.

This reality is considered, as we have mentioned above, by the EU incumbents, confirming this way that KBS is a strategic goal and intensive studies are made just to optimize the road to it, as it is next expressed [10]:

<<Knowledge and how people acquire knowledge has fascinated human beings from the ancient Greeks to our day. With the emergence of the so-called knowledge economy, knowledge has become one of the most fashionable terms in the political and managerial sphere. As Weiler ... put it: "The politics of knowledge become less and less separable from the politics of production and profit, arguably the most powerful political dynamics in today's world". Halal's ... enthusiastic words show this central role of knowledge in the new way of seeing the organization: "We see now that knowledge is the most strategic asset in enterprise, the source of all creativity, innovation and economic value".>>

Although the management is considered first (at EU level), here we can observe an interesting analysis of the knowledge basics, which could have useful impact for identifying best practice approaches:

<<The main focus will be on knowledge management approaches, since this field of study has been very much involved in relating knowledge to managerial practices. Managing knowledge can be seen as the crucial aspect of the so-called "knowledge economy", and hence it has strong influence on innovative practices ... Three main approaches to epistemology are proposed in a broad sense: innatist/introspection, empiricism/behaviorism and critical philosophy/constructivism. Special emphasis is placed on constructivism since it is argued that it is the predominant view nowadays...The first distinction, commonly addressed in the literature, is between knowledge, information, data and expertise ... Data constitute the bricks from which the pyramid of knowledge is built. Davenport et al. ... define data as "a set of discrete, objective facts about events". Most of the knowledge management literature agrees with this definition (see Tuomi ... for a review). Data constitute all the empirical reality that is presented to us in our daily experience, the "ocean of impressions" in Kant's terms. In managerial terms, data are that which are available without much restriction in the knowledge society, on the Internet, in databases, and in daily activities. Bender and Fish ... point out that data become information only when they have been imbued with meaning, understanding, relevance and purpose...>>

Beyond the appearance of theoretical issues, we consider that presenting the different approaches of generating/perceiving data, information and knowledge could reveal important elements to be used when optimizing these processes in the computer-human relation and generally in the **AI/ICT/KBS**:

<<While general information is data that are organized and structured, contextual information is created by filtering and organizing general information to meet the requirements of a specific community of users. Information is transformed into knowledge when the individual processes it and internalizes it. New information has to be integrated into the individual's existing knowledge structure in order to produce learning. In Piaget's terms, the new information can be either accommodated or assimilated. If the new information is assimilated there is a change in the content of the schema. We add the information into our schema, but the schema does not change dramatically. If accommodation occurs, the new information triggers a change in the structure of the schema. A new schema means that the reality that the information refers to is looked upon in a different way. In this way, schemas "form the basis for comparing and interpreting incoming data" (Shute...). Finally, if one masters a certain subject or area of knowledge, s/he will become an expert...>>

If initially there was no clear link with **AI/ICT/KBS** optimization processes, we consider that finally we have to notice here a valuable set of available rules in the generation and exploit of knowledge, but more than these, we find a practical description of the expertise/expert in this emergent and difficult field of human activity:

<<Alexander..., in her review of psychology research on expertise, maintains that experts are people who:

- possess extensive and highly integrated bodies of domain knowledge,*
- are effective at recognizing the underlying structure of domain problems,*
- select and apply appropriate problem-solving procedures for the problem at hand,*

and

- can retrieve relevant domain knowledge and strategies with minimal cognitive effort.*

Wiig ... refers to the proficiency dimension to explain levels of expertise with regard to knowledge. He proposes seven different categories from beginner to "grand master". He differentiates between expertise and wisdom. Expertise refers to "specialized knowledge and skills in a particular area" ... whereas wisdom involves, in addition to a high level of knowledge in a specific area, certain personal characteristics such as the willingness to learn or to be flexible. Tuomi ... criticizes this classical hierarchical conception of knowledge and information. He argues that all these models consider knowledge as a "higher form of information" ... For Tuomi ..., "data emerge last – only after knowledge and information are available. There are no 'isolated peaces of simple facts'". He turns the pyramid upside-down. Once knowledge is articulated, verbalized and structured, it is transformed into information. Information is transformed into data when it is placed within certain predefined structures...>>

Regarding the (Tuomi) shocking opinion, last presented, we consider that as **an opening to a larger view of the complex and dynamic processes of generating/perceiving data, information and knowledge in the emerging World of KBS, where the relativity and the probabilistic computing approach, we already presented [19], are inherent features.**

With other words, we have to agree that **understanding the unlimited dynamic AI/ICT/KBS optimization processes is a long and winding road, full of surprises, we have to face with timely analyses and ... imagination.**

3. Conclusions

Analysing the role of the ICT advances, where artificial intelligence (AI) is by far the most expected instrument for data/information high level processing in the data deluge and Big Data context, leading to the Data Science (DS) concept, is a way to reveal the essential links between the main parts of our title point (DS, advanced AI and core of KBS) with the complex processes of the **ICT/KBS context**, in a **circle** evolution and **mutual-supporting relations**, having crucial and complicate consequences (mainly benefic) for the Earth ecosystem. These premises, due to the scientific, technical and economical reasons of optimizing the mentioned processes, further link the **ICT/KBS context** with the necessary concept of *Cyberinfrastructure*, we consequently have to approach. We analysed this concept keeping in mind our title point on knowledge (DS, advanced AI and core of KBS), aiming to emphasize its main components, complex interactions and **mutual-supporting** mechanisms, which are usually shadowed by the well-known resonance of the term *infrastructure*.

Using appropriate references for confirmation, we concluded that, beyond the usual approach that involves *research activities*, the *Cyberinfrastructure* should contain *people* and *scholarly activities* more generally, including teaching and learning, i.e., the whole education fields, which in the Big Data era became not only important, but part of the most difficult processes of decision.

Another finding is that ICT advances impact not only the humankind activity fields, but also the people's thinking, generally for understanding all practical implications of using AI/ICT features/applications, but, essentially, **we consider them an expression of the golden rule AI/ICT have to implement in order to maintain its race toward better performance: re-invent itself all the time by leveraging people's creation skills and thinking.**

Concretely, in a systemic approach, after considering *Cyberinfrastructure* hard/soft resources and instruments, we pointed the human component, starting from the IBM example (*Implement explainable AI*) [1], meaning the aim of leveraging the human potential.

It is very important to observe that this goal is double fold, aiming apparently to leverage user's capacity to fully exploit the designed AI (machine learning algorithms) potential, but eventually this will increase experts' capacities to use their imagination in order to improve (**re-invent**) AI instruments performance.

Consequently, the complex (mutual-supporting) relations between *Cyberinfrastructure* components are approached from the human side, considering their crucial role in the actual World context where USA and EU agreed on the strategic importance of the **acquired technological advances (i.e., supposing the associated economic, political and social consequences, but also the cyber-security)**. It resulted that the need for highly skilled people raised **the importance and role of the human component for Cyberinfrastructure and then for AI/ICT/KBS, but the complex relations between them represent the most complicate problem to be further considered**, by the need to timely analyse all the consequence of ICT

exponential development and impact at Earth scale and on **human behaviour and evolution**, approaching the ICT-human's complex relations and impact on KBS.

Continuing to follow the ICT rule of re-inventing itself, we consider that, beyond the new hard/soft performant technologies, in the future we have to look for other leveraging ideas, mainly on improving the output of the computer-human relation by refining knowledge by more subtle criteria of performance.

In a concrete manner, the perpetual informational transition where the engineers have to work, is updated to the actual computer-human relation/technology, where the knowledge is refined. The knowledge exchange between researchers is in fact the area that pushes the World ahead with the greatest impact, but also with the most of generation and refining problems, which are worth to be analysed in order to identify prominent mechanisms that could contribute to the optimization of **AI/ICT/KBS development in the Big Data era**, taking into account the actual challenges and expectations which are naturally associated with the unprecedented advances already reached, but difficult to be improved.

Knowing those problems could contribute to the configuration of the actual and the emerging context of developing AI/ICT/KBS, providing this way a useful base for the analyses/decisions, irrespective how these could/will be.

One of the main mechanisms of progress, for **AI/ICT/KBS**, is refining knowledge at Earth ecosystem scale, which is essentially depending on human capacity of creation and integration of the advanced technology in new models, with **a larger view of the complex and dynamic processes of generating/perceiving data, information and knowledge in the emerging World of KBS, where the relativity and the probabilistic computing approach [19] are inherent features.**

We concluded that **understanding the unlimited dynamic AI/ICT/KBS optimization processes is a long and winding road, full of surprises, we have to face with timely analyses and ... imagination.**

REFERENCES

- [1] *** , *IBM Watson Studio*, 2021, https://www.ibm.com/cloud/watson-studio?p1=Search&p4=43700052310836558&p5=b&gclid=CjwKCAjwiLGGBhAqEiwAgq3q_kVnog-GfCMmo417Tjt9fX_PrUyfs3QWUVB3QD1VAg2MwYhte-183BoC90gQAvD_BwE&gclid=aw.ds
- [2] Craig A. Stewart et al, *What is Cyberinfrastructure?*, SIGUCCS'10, October 24–27, 2010, Norfolk, Virginia, USA, <https://www.researchgate.net/publication/49471518>
- [3] Victor Greu, *Information and communications technology is merging data science and advanced artificial intelligence towards the core of knowledge based society -(Part 1)*, Romanian Distribution Committee (affiliated to the “International Association of the Distributive Trade”-scientific association – A.I.D.A. Brussels) Magazine(international; electronic; covered in RePEc International Data Base), Volume 12, Issue 1, Year 2021.
- [4] Jim Gray, *What Next? A Dozen Information-Technology Research Goals*, Microsoft Research, June 1999, Technical Report MS-TR-99-50
- [5] E. S. Vorm, *Computer-Centered Humans: Why Human-AI Interaction Research Will Be Critical to Successful AI Integration in the DoD*, IEEE Intelligent Systems (Volume: 35, Issue: 4, July-Aug. 1 2020)

- [6]Florin Enache, Victor Greu, Petrică Ciotîrnae, Florin Popescu, *Model and Algorithms for Optimizing a Human Computing System Oriented to Knowledge Extraction by Use of Crowdsourcing*, 2020, 13th International Conference on Communications (COMM), (Politehnica University of Bucharest, Military Technical Academy, IEEE Romania), (COMM 2020 is covered in IEEE Explore Database and ISI Web of Science in the Conference Proceedings Citation Index)
- [7] Robert W. Lucky, *The expiration date of knowledge*, IEEE Spectrum, vol. 56, no. 09, pp. 21-21, Sept. 2019
- [8]Daniel E. Atkins et al, *Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*, 2003, <https://www.nsf.gov/cise/sci/reports/atkins.pdf>
- [9]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 (affiliated to the “International Association of the Distributive Trade”-scientific association – A.I.D.A. Brussels) Magazine(international; electronic; covered in RePEc International Data Base), Volume 10, Issue 4, Year 2019.
- [10]Ernesto Villalba, *The Concept of Knowledge for a Knowledge-based Society From knowledge to learning*, European Commission Joint Research Centre © European Communities, 2007
- [11]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 (affiliated to the “International Association of the Distributive Trade”-scientific association – A.I.D.A. Brussels) Magazine(international; electronic; covered in RePEc International Data Base), Volume 10, Issue 1, Year 2019.
- [12]Gerbrand Tholen, *The problem with a knowledge-based society*, November 26th 2017, <https://blog.oup.com/authors/gerbrand-tholen/>
- [13]Victor Greu, *Extending information and communications technologies’ impact on knowledge based society through artificial and collective intelligence* –(Part 3), Romanian Distribution Committee (affiliated to the “International Association of the Distributive Trade”-scientific association – A.I.D.A. Brussels) Magazine(international; electronic; covered in RePEc International Data Base), Volume 9, Issue 3, Year 2018.
- [14] 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/>
- [15]Victor Greu et all, *Human and artificial intelligence driven incentive-operation model and algorithms for a multi-purpose integrated crowdsensing-crowdsourcing scalable system*, Proceedings of International Conference Communications 2018, (Politehnica University of Bucharest, Military Technical Academy, IEEE Romania), June 2018(COMM 2018 is covered in IEEE Explore Database and ISI Web of Science in the Conference Proceedings Citation Index).
- [16]Fran Berman, *Current Working Definitions Of Cyberinfrastructure*, 2005, https://www.researchgate.net/publication/49471518_What_is_Cyberinfrastructure/link/0912f51085bdceaf0a000000/download
- [17]Gordon Bell, Tony Hey, Alex Szalay, *Beyond the Data Deluge*, Science Vol 323, 6 March 2009.
- [18]Victor Greu, *Searching the right tracks of new technologies in the earth race for a balance between progress and survival*, Romanian Distribution Committee Magazine,
- [19] Victor Greu, *Information and Communications Technology new paradigm of probabilistic computing could inspire our thinking through a future World of uncertainty* --(Part 3), Volume 11, Issue 4, Year 2020.