Information Engineering: Strategic decision based on data science

Prof. Carlos Alberto Martins da Rocha¹, Prof. Dr. Luis Borges Gouveia²

¹Professor of Television Journalism and Cinema. Department of Social Communication at the Federal University of Paraná (Brazil)
PhD student in Information Science

²PhD in Computer Science from the University of Lancaster (2002). Full Professor at the University Fernando Pessoa. (Portugal)

Received: 27 Apr 2021;
Received in revised form: 27 May 2021;
Accepted: 10 Jun 2021;
Available online: 18 Jun 2021
©2021 The Author(s). Published by AI Publication. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Abstract—The contemporary society of the 21st century is composed and grounded in data. This scenario was made over the years and centuries of history. The technology, however, is recent in human history and began about 400 years ago, simultaneously with the emergence of modern science. We are no longer an industrial society, but an information society. This is a conception of a new social organization, based on information technology. The Information Society is supported by the acquisition, storage, processing and distribution of information by electronic means, in Information and Communication Technologies. The interaction between individuals and institutions became predominantly digital. Up-to-date managers must make data-driven decisions. Those who are able to unite domain expertise with data science can make more accurate strategic decisions. Knowledge is linked to its context in a broader way, with social and cultural implications as well. Industry 4.0 makes it possible to gather and analyze data between machines, faster and more efficiently, thus having an effect on the competitiveness between companies and even regions, even altering the economy. Information Science is a vast growing field of research and development of solutions, including, and mainly, for Knowledge Management and Strategic Management, in decision-making by public and private institutions.

I. INTRODUCTION

The contemporary society of the 21st century is composed and grounded in data. This scenario took place over the years and centuries of history. In summary, human knowledge since the dawn of man, with the discovery of primitive techniques, originated in the discovery of fire, in the polishing of stones and in the cooking of food, even in the Paleolithic period (VARGAS, 1985). Knowledge managed to be better passed on from one generation to another when writing rudimentary came to represent human thought and language. Initially pictorial, then cuneiform. This all started more than 3,000 years before Christ. Afterwards, many techniques were created and improved in order to optimize the communication, information and, why not say, training process.

Knowledge is the result of a codification/decoding process since its beginning:
The ways of recording knowledge went through changes in the alphabets adopted throughout human history and the respective domain of this system of encoding and decoding recorded messages.

According to Nobrega (2009), language is alive and constantly changing, even today. In the author’s opinion, the press is a daily example that accompanies movement and transformation, including the inclusion of slang in the journalistic repertoire. It is undeniable that information, based on facts, is the basis of a pyramid of human knowledge. Barreto (2003) shows us that information, knowledge and intelligence are amalgamated structures in order to organize, to transfer and constitute human knowledge.

The technology, however, is recent in human history, according to VARGAS (1985). It began about 400 years ago, simultaneously with the emergence of modern science. But it only took shape with the Industrial Revolution. Technique and technology merge in a choreography of complementary actions and converging results. They make it possible to combine technical knowledge, which allows us to have a precise look at a specific solution. Meanwhile, scientific knowledge drives us away, forces us to see a problem from the perspective of a wide-angle lens, which allows for context analysis and
the advancement of knowledge. Science, in itself, is a search for truth. The search for answers to phenomena, natural or artificial (man-made).

II. METHOD

Vargas (1985) also shows us that the origin of science lies in a type of theoretical knowledge invented by the Greeks in Ionia (6th century BC). The “epistēme theoretiké”:

“The Greek word episteme can be translated to know. But the knowledge that was established in the Greek world in the sixth century BC was, by themselves, adjectived as theoretiké; this is theoretical knowledge. The Greek word theory comes from the verb theorein which means ‘to see’. Therefore, ‘epistēmethoretiké’ is a knowledge acquired by the ‘eyes of the spirit’, capable, according to the Greeks, of discovering reality o it actually is. Theory is thus linked to the truth; in Greek, ‘truth’ is said to be aletheia what is discovered.” (VARGAS, 1985, P01)

Classical science, from ancient Greek, still remains intrinsic to the essence of modern science, theoretical knowledge. The man in his childhood, youth and old age remains in his essence the same man, despite changing his figure. It is what remains that allows us to logically reason and construct theory. The basis of science. The ideas in Plato's theory and the changing appearances of Aristotle are substances, which are behind, for example, Geometry (3rd century BC), with its points, lines and planes; postulates and theorems.

Libraries are the representation of the need for information and knowledge. Humanity accumulates and loses knowledge because of the way it stores and preserves information. The organization of information, therefore, is a science that every day proves to be more necessary and even vital for organizations and people. First they were crystallized and rigid catalog information, today they are digital libraries, with fluid organization to be able to encompass the content. Scientific work is based on content management, with methods and strategies, production and organization. Castells (2001) already tells us that we are no longer an industrial society, to be an information society. This conception of a new social organization, based on information technology, pointed out in the article by Gouveia (2004), which highlights not only the use in related activities, such as banking and mobile phones, for example, but also in video systems, linked to surveillance and traffic control. Gouveia (2009) also reveals that the Information Society is supported by the acquisition, storage, processing and distribution of information by electronic means, in Information and Communication Technologies (ICT). He also points out that the beginnings of centuries have been periods of great changes and transformations for Western civilization, and the 21st century is no exception. As a product, immense changes in the habits of the individual and in nature; in the activities of organizations and the use of information as a strategic resource.

Saracevic (1996) traces that the origin of Information Science can be identified after World War II, in the wake of the scientific and technical revolution, in an MIT article, by Vannevar Bush, in 1945:

“In this important article, BUSH did two things: (1) succinctly defined a critical problem that had been on people's minds for a long time, and (2) proposed a solution that would be a technological tweak, in keeping with the spirit of the time, in addition to strategically attractive. The problem was (and basically still is) ‘the massive task of making a growing body of knowledge more accessible’: BUSH identified the problem of information explosion - the irrepressible exponential growth of information and its records, particularly in science and The solution he proposed was to use the incipient information technologies to combat the problem. And he went further, proposed a machine called MEMEX, incorporating (in his words) the ability to associate ideas, which would duplicate 'mental processes artificially.' The anticipation of the birth of IC and even of artificial intelligence is quite evident.”(SARACEVIC. 1996. P02)

III. RESULTS

The information society changes its focus and uses information as a strategic and central resource for all human activity (CASTELLS, 2001). The interaction between individuals and institutions becomes predominantly digital. Currently, not only social and interpersonal relationships are mediated by social networks, present in the vast majority of mobile phones. Business relationships between companies, such o banking institutions and even the simple ordering of a pizza at night, are already a digital reality due to the applications available to practically everyone. This social structure that we are now experiencing makes users providers of online digital data (digital footprints), but at the same time they provide feedback and strengthen the system, as evidenced in the text “Big Data: The Management Revolution”:

“After purchasing online, customer understanding increased dramatically. Online retailers could track not only what customers buy but also what
else they were looking at; how they navigated the site; how much they were influenced by promotions, reviews and page layouts; and similarities between individuals and groups. Before long, they developed algorithms to predict what individual customers would like to read next.” (MCAFEE. 2012, P04).

In analogue society, strategic management decisions were based on scarce data, costly to obtain or not available in digital formats. People made management decisions based on accumulated experience and patterns of inference, deduction and induction, or intuition. Also according to McAfee (2012), to date there are few data scientists. The technologies are new and it is very easy to confuse causal correlations and find wrong patterns. Up-to-date managers must make decisions based on data. Those who are able to unite domain expertise with data science can make more accurate strategic decisions and, possibly, we can infer, determine investments and future actions more confidently.

IV. DISCUSSION

But this strategic knowledge management is based on a succession of concepts and principles, which clearly demonstrate the importance of knowing how to deal with and understand the correct feeding of information systems, from the simple task/action of data entry. Initially, as Weinberger (2010) points out, the proposal of the data-information-knowledge-wisdom hierarchy, characterized in the DIKW (Data, Information, Knowledge and Wisdom) pyramid, seemed like a great idea, based on the logic of Computer Science, where one learns that information is a refinement of mere Dice. Information, therefore, is the value we extract from the data.

This graphic model was used by several authors and sought to reflect not only the concepts of knowledge management, but the functions and attributions related to the information flow in an organization. The amount of data generated and available, exposed or not, leads to a chain of activities relevant to Information Science, which seeks to systematize and organize in order to generate information and knowledge. In short, they can be used in strategic intelligence planning.

Management and information science seek to systematize and usually represent the hierarchical structure that leads to decision making, based on a pyramid symbol. The base is composed of data, which pass information, which, in turn, generates knowledge and which, finally, generates

![Fig.3: Pyramid DIKW.](image1)

*Font: authored by the author himself*

![Fig.4: Knowledge hierarchy. ROWLEY (2006).](image2)

*Font: [http://www-public.imtbs-tsp.eu/~gibson/Teaching/Teaching-ReadingMaterial/Rowley06.pdf]*
intelligence, where strategic decision-making occurs (or should at least occur).

According to Remor, Fialho & Queiroz (2017), the data (1st step of the pyramid) is composed of no assigned value and individual (of things and people). The information level (2nd step of the pyramid) is where the data becomes meaningful and can be categorized and measured. At the level of knowledge (3rd step of the pyramid) is where the information can be absorbed and memorized; where learning is positioned. At the top (4th step of the pyramid) is where understanding, reflection, more precisely: wisdom; and where the decision-making process is expected to take place.

“... the view of greater consensus in the literature, about the structure of the hierarchy, perceives data as simple facts that become information as data is combined in structures that make sense or have purpose, which subsequently become knowledge as information is put into practice. context and can be used to make predictions.” (REMOR, FIALHO & QUEIROZ. 2017. P04)

However, the changes in stages of this pyramid encounter problems of clear definition, for example, when it comes to the change of knowledge and wisdom. Other forms of representation of the structure of knowledge, although less widespread, incorporate other factors and even portray the participation of human interaction, as this model proposed by Choo (2006) points out.

![Fig.5: Data, Information and Knowledge. CHOO (2006)](http://choo.fis.utoronto.ca/FIS/ResPub/KOart.html)

This graph shows that the higher, on the scale, or closer to the pyramidal top (by direct analogy), the more there is a need for human participation and interaction. Now, going back to the authors Remor, Fialho & Queiroz (2017), we have that “wisdom would be the ability to project the consequences of an act, and assess the scenario taking into account what is desired.” Still, in conclusion, and for that reason the graphic of Choo (2006), we see this statement that the subjective value competences are of predominantly human characteristics, with Ackoff (1989) apud Remor, Fialho & Queiroz. (2017):

“Wisdom would be the ability to project the consequences of an act, and assess the scenario taking into account what is desired. Ability to inquire about what is not understood, and through that, goes far beyond understanding itself, it would be the essence of ‘philosophical proof’. It is the process by which we can discern or judge between right and wrong, good and bad, says Ackoff (1989). The author personally believes that computers will never possess the capacity for wisdom.” (REMOR, FIALHO & QUEIROZ. 2017. P04)

According to Weinberger (2010), knowledge is when information is transformed into instruction. Citing the official guide to good practices in Knowledge Management, from “The European Committee for Standardization” (CEN), “Knowledge is the combination of data and information, to which is added expert opinion, skills and experience, to result in a valuable asset which can be used to aid decision making. Knowledge may be explicit and/or tacit, individual and/or collective.” Still, continuing in CEN’ text, “Knowledge Management is the management of activities and processes for leveraging knowledge to enhance competitiveness through better use
and creation of individual and collective knowledge resources.”

Knowledge is a much more complex process than a game of assembling pieces and discovering hidden images in it. “Knowledge is not determined by information, as it is the knowledge process that first decides which information is relevant and how it should be used” (Weinberger, 2010. P3). Knowledge is linked to its context in a broader way, with social and cultural implications as well.

The search for the desired results requires current organizations to constantly plan and consequently manage existing data sources. Organizational knowledge management is strategic for more accurate decision making. The volume of data available in this Information Society, in this Digital Century, is incomparable with any analog storage available in the past. It takes investment and mastery of digital tools to store and collect data, process and obtain information. Then, generate knowledge that can then be a strategic asset. Finally, increase the “wisdom” to be able to feed back this information system or choose to discard data that could generate distortions of reality.

Due to the large volume of data available in these times of the Information Society, the need for investments in skills and people is undeniable, but much depends on machines and processes.

Computers

This current social reality, which mixes the real world and allocated and processed data, in a virtual world with processing capacity increasing with each generation of new chips, forces us to think about the rapid technological evolution of computers, but more specifically the evolution of processors, as these are the main hardware components. This is exactly where software is processed and executed. There is no room for a detailed description of the entire evolution of the generations of processors and computers here, but it is worth highlighting Moore's Law:

“On April 14, 1965, Intel founder Gordon Moore published an article in Electronics Magazine about increasing the processing power of computers. Moore states in the article that this capacity would double every 18 months and that growth would be steady. This theory became known as 'Moore's Law' and remains valid to this day.” (ALMEIDA, 2009, P01)

The consequence of this technological race led us from the first commercial microchip, launched in 1971 by Intel (Intel4004), to a series of new processors, whether produced by Intel itself or by competitor AMD, which already achieve high performance in processing capacity, when compared to the precursor Intel4004 core. Currently, Intel, for example, makes explicit on its official website that:

“We believe that data is dramatically shaping the future of all humanity. Intel is working tirelessly to unlock the potential of data, leading to more capable and efficient networks and pervasive AI in smart devices. Moore's Law set the pace for the digital revolution and continues to inspire us to this day.” [emphasis added] (INTEL, 2021)3.

Microchips that have 18 colors (processing cores) on a single i9 chip are already produced and are on sale for consumers in general, as in Intel's X-series (i9-10980XE)2. However, the Information Society always needs more and more processing, as the volume of data grows daily on an exponential scale. The days of computing based only on a binary system, of 0 and 1, already demonstrates exhaustion with the appearance of the chip with quantum processing, which can use more variables and execute a much higher number of instructions per second. Released only for commercial use by the Canadian company D-Wave4, this new paradigm of computational processing is now free to operate with other sets of variables and consequently new horizons of use in the real world. D-wave itself already indicates that this processor can be used for “airline programming, election modeling, quantum chemistry simulation, automotive design, preventive health, logistics and much more”. [our emphasis] D-WAVE4: This company has even been selected to demonstrate its real usability for the Australian Department of Defense5. What may appear to the common user as something distant is already a reality in research spaces and in some corporations, as IBM, Intel and Google are already investing in the development of the new quantum chip model.

---

3 https://www.intel.com/content/www/us/en/company-overview/company-overview.html
2 https://www.intel.com.br/content/www/br/pt/support/articles/000005505/processors.html
4 https://www.dwavesys.com/
5 https://www.dwavesys.com/press-releases/nec-d-wave-and-australian-department-defence-collaborate-quantum-computing-initiative
The most advanced quantum chip available today [2021] is the latest generation of D-wave, with a capacity of 5,000 qubits.

“As an example, a 128-bit AES (symmetric) cipher has 2128 (about 1038) possible keys. A classical computer, which generally executes 1 trillion instructions per second, would take about 10.79 quadrillion years to test every possibility. Conversely, for an n-bit cipher, a quantum computer operates on the order of 2n = 2n/2. For an 128-bit cipher, this is 264 (about (1019) steps and it would take about 6 months to test every possibility.” (KIRSCH, 2015, P8)

However, the innovations and applications of new processors also permeate other frontiers, as already published in the scientific magazine “Scientific American Brazil”7, which shows us another option that also goes beyond the binary set (0 and 1), but starts using A, T, C and G, the components of DNA, to perform mathematical calculations and logical deductions. In addition to the broad spectrum of processing, based on 4 information sources (A, T, C and G), this type of processor is under development with a focus on disease detection and can be easily inserted, or incorporated into the human body, as it is a biological material (with a non-silicon-based substrate) and has the potential for numerous application possibilities at the cellular level. "Because biomolecules are able to access data encoded in other biomolecules, they are compatible with living systems as electronic computers will never be." (Scientific American Brazil, 2020)

### Industry 4.0

Costa (2017) points out that the third Industrial Revolution began in 1970, due to the digital revolution, as a result of the increase in the use of semiconductors, automation and robotization of production, in addition to advances in communications and the internet (COSTA, 2017).

Now, according to Rüssmann, Lorenz, Gerbert & Waldner (2015), we are experiencing the fourth wave of industrial revolution, driven by technological advances, where machines, parts and IT systems will be connected beyond a single company, using cyberphysical systems, which can connect with each other using the Internet. Industry 4.0 makes it possible to gather and analyze data between machines, faster and more efficiently, thus having an effect on the competitiveness between companies and even regions, even altering the economy.

---

6 https://www.dwavesys.com/d-wave-two%E2%84%A2-system
7 https://sciam.com.br/computadores-de-dna-ganham-vida/
Although Rüßmann, Lorenz, Gerbert & Waldner (2015) point out 9 bases for Industry 4.0, Costa (2017) presents a summary proposal, based on only 3 pillars: Internet of Things, Cyber-Physical Systems and Big Data. In this investigation, the aspect that is outlined is the importance of using data, in a treated and polished way, for analysis and consequent strategic decision-making, in public or commercial institutions.

**BIG DATA**

Big Data is a fundamental technology for I4.0, as pointed out by Ottonicar, Atayde and Santa-Eulalia (2021), and the trend is for companies to increasingly use these systems. This happens because not only will people be producers of data and information, but also biological objects and elements. Also, as the authors claim, Big Data banks store a huge amount of data generated by the internet, mainly by the Internet of Things, and feed strategic (or intelligent) decision-making.

The scale of industrial evolution takes us to this level with possibilities for the production of many different data, whether individual or legal entities, which can be collected, processed and analyzed through the Internet and will certainly have impacts on education, and especially on universities. Whether in the curriculum or in bureaucratic routines, but certainly also in fields of research and technological innovation, among them Information Science.

Still in the argument of Ottonicar, Atayde & Santa-Eulalia (2021):

“I4.0 has several technologies that make it up in the organizational environment, including Big Data. This technology helps to store a large amount of data produced by people and objects. This technology and the advent of I4.0 bring changes to universities and can be leveraged to develop new ways of learning.” (OTTONICAR, ATAYDE & SANTA-EULALIA. 2021. P 159)

One of the main concerns is the issue of individuality and privacy, as, according to PIMENTA (2013), we leave digital footprints of political trends, purchase and sale relationships, and other private data that are collected through surveillance technologies and monitoring. As noted by McAfee (2012), the developed algorithms are able to predict the actions of users, based on their preferences. However, as a tool, when well used, it can contribute to citizen and democratic education, with public access to data, information and even analysis of these, such as state spending and other areas of society.
We can call this virtual universe, where the data are, the Big Data, the datasphere, as described by Magaly Prado (2021): “The datasphere has been the object of study by this researcher since 2005” (PRADO, 2021. P140). Thus, personal data are already available in the datasphere and with consequences that go beyond responsible and democratic use.

Panama Papers
The concentration of large volumes of data can also have drastic consequences for the security system, which needs constant improvement against intrusions or simple leaks. A classic example in the world of communication is the case called Panama Papers, when the German newspaper SüddeutscheZeitung received and shared a volume of data on financial transactions from the Panamanian law firm Mossak Fonseca, with the International Consortium of Investigative Journalists (ICIJ). Explained in detail by the BBC (2016), which is part of the 107 organizations that analyze the 11 million documents, it is clear that the company helped clients avoid sanctions and pay taxes and launder money, from many people, and even from 72 heads of state.

Cambridge Analytica
An article in the magazine Direitoem Debate, by FORNASIER and BECK (2020), shows in detail what has also become known worldwide as the case of the misuse of a large volume of data. According to Fornasier and Beck (2020), it was first denounced by journalist Harry Davies, from the English newspaper The Guardian, on December 11, 2015, saying that the company Cambridge Analytica (CA) had illegally collected millions of personal data from the social network Facebook, unbeknownst to users, for use in political and electoral campaigns.

“The Guardian has continued the journalistic investigation since February 2017, with renowned journalist Carole Cadwalladr, as well as the independent online newspaper The Intercept, since March of the same year. On May 7, 2017, Cadwalladr wrote for The Observer, together with The Guardian, an article with a provocative title: The great British Brexit robbery: how our democracy was hijacked, using a source as the main means of accessing information, then anonymous. The article at the time went viral, gaining more than 63,000 shares.” (FORNASIER & BECK, 2020. P3)

It was a scandal in the world press, an example that used Information Science practices on Big Data, security and privacy to manipulate elections in democratic countries such as the United Kingdom and the United States.

“March 17, 2018: the day that sealed CA's fate forever. On that day, three of the world's leading journal and newspaper organizations – The Observer, The Guardian and The New York Times – jointly published the article titled How Trump Consultants Exploited the Facebook Data of Millions ('How Trump Consultants Explored the Facebook data of millions'). The article was only viable after Carole Cadwalladr was very convinced to demonstrate the right, ethical and democratic path – despite having a huge personal cost – of her anonymous source publicly denouncing the CA of its data mining practices (Data Mining) and data processing (interpreted here as the practice of Data Scraping) and efficiently interfere in the results of democratic processes.” [our emphasis] (FORNASIER & BECK, 2020. P4)

V. CONCLUSION
“Big data studies are often linked to extremist positions of exaggerated optimism or exaggerated pessimism. Related to this duality polarized in the consideration of Big Data (on the one hand the miraculous bets of the enthusiasts and on the other the warnings of the extreme danger of radical critics)” (SOUZA & GONZALEZ, 2021. P11); however, it is a solid understanding that Big Data is an important source of data and information for strategic decision making.

The reality that presents itself, with new and more powerful processing technologies, whether in hardware or software, with the formation of huge databases on users of social networks or applications with common access to all mobile phones, available on a large scale through the world, shows us that Information Science is a vast growing field of research and development of solutions, including, and mainly, for Knowledge Management and Strategic Management, in the decision-making of public and private institutions.

VI. REFERENCES
[1] ALMEIDA, R. B. Evolução dos processadores - comparação das famílias de processadores intel e amd. Disponível em: <https://www.ic.unicamp.br/~ducatte/mo401/1s2009/T2/089065-t2.pdf>. Acesso em 28 de maio de 2021.
