

# Economic Analysis of Law Review

## Data Capitalism: Efficiency as a sociability degree function

*Capitalismo de dados: eficiência como função de grau de sociabilidade*

Dora Kaufman<sup>1</sup>

USP

### RESUMO

O objetivo deste artigo é abordar a relação entre sociabilidade e eficiência nos modelos de IA, em como a economia contemporânea trouxe a noção de eficiência para nossas vidas pessoais. Inicialmente, introduzimos os conceitos básicos de três conceitos principais do artigo: capitalismo de dados, dados e aprendizado profundo. A seguir, descrevemos a evolução exponencial da tecnologia de armazenamento, processamento e transmissão, mostrando que, ao longo dos anos, a capacidade de transformar dados analógicos em dados digitais aumentou exponencialmente. Essa capacidade aumentou a eficiência dos processos operacionais com a medida de eficiência calculada e controlada contra o potencial máximo dos dados digitais produzidos nessas interações. Para empresas tradicionais, competir com rivais digitais envolve re-arquitetar a organização e o modelo operacional da empresa. A compartimentalização em silos compromete a eficiência dos modelos de acionamento AI que exigem base de dados integrada. A transformação digital requer um enorme investimento em gerenciamento, tempo e recursos financeiros. No entanto, é a única maneira de permanecer competitivo e sobreviver no mercado do século XXI. O compromisso de identificar e medir as preferências e hábitos do usuário e, em seguida, prever o comportamento, é a lógica por trás das plataformas e aplicativos de tecnologia, redes sociais online, comércio eletrônico e mecanismos de pesquisa. As plataformas digitais são projetadas para estender a vida útil de seus usuários, gerando maior envolvimento e mais dados. A originalidade deste artigo é correlacionar sociabilidade e eficiência econômica no atual ambiente de negócios com uma abordagem tecnológica e social.

**Palavras-chave:** capitalismo de dados; inteligência artificial; ambientes tecno-sociais inteligentes; eficiência.

### ABSTRACT

The purpose of this paper is to address the relationship between sociability and efficiency in AI-drive models, in how contemporary economics has brought the notion of efficiency into our personal lives. Initially we introduced the basics of three key concepts of the article: data capitalism, data and deep learning. Next, we describe the exponential evolution of storage, processing and transmission technology showing that over the years, the ability to transform analog data into digital data has expanded exponentially. This capacity increased the efficiency of the operational processes with the measure of efficiency calculated and controlled against the maximum potential of the digital data produced in these interactions. For traditional firms, competing with digital rivals involves rearchitecting the firm's organization and operating model. The compartmentalisation in silos compromises the efficiency of AI-drive models which demand integrated data base. The digital transformation requires huge investment in management, time and financial resources. However, it is the only way to remain competitive and survive in the 21st century market. The commitment to identify and measure user preferences and habits, and then to predict behaviour, is the logic behind technology platforms and applications, online social networks, e-commerce and search engines. Digital platforms are designed to extend the lifespan of their users, thereby generating greater engagement and more data. The originality of this paper is to correlate sociability and economic efficiency in the present business environment with a technological and social approach.

**Keywords:** data capitalism; artificial intelligence; smart techno-social environments; efficiency.

**R:** 21/02/18 **A:** 21/06/20 **P:** 31/08/20

<sup>1</sup> E-mail: dkaufman@usp.br

## 1. Introduction

Each economic model has its own persuasion mechanisms, extrapolating consumption with cultural and behavioural impacts, including access to information. In the industrial economy, which is characterised by mass production and the mass consumption of goods and services, advertising has predominated as a means of convincing and influencing individuals' choices and preferences through mass communication via traditional media outlets (television, radio, newspapers, magazines). The networked information economy is characterised by the customisation of products and services, with the segmented communication being directed at the public in specific niches, based on their similar profiles, thus minimising dispersion and costs.

In the data economy (or data capitalism)<sup>2</sup>, "personalisation" is in the basis of mediation of products, services and information; artificial intelligence (AI) algorithms promote assertive communication strategies that are based on the captured, mined and analysed knowledge of the personal data that are generated during digital interactions.

The dramatic increase in the supply of data created a corresponding demand for improved data storage and analytics technologies, and, in a few decades, major technological advances have coalesced in new fields such as big data and reawakened seemingly dormant fields such as artificial intelligence and machine learning. (FRISCHMANN, SELINGER, 2018, p. 115).

In this new logic, the greater the interaction between individuals, i.e., their sociability and communication, the greater the generation of personal data, the greater the data collection and storage activity, the greater the market concentration and the greater the power of large data "hub" firms - Big Techs<sup>3</sup>. These firms accumulate data through their multiple network connections extracting value through analytics and AI.

The AI-drive models correlate efficiency and sociability: the greater the sociability (social interactions) the greater the generation of data, which implies increased efficiency of those models; or as Frischmann and Selinger (2018) observe, a virtuous circle is established: sociability / social interactions generate data, data enhance technological systems, and technological systems enhance sociability / social interactions. The authors coined the term "the smart techno-social environment" to designate the impact that techno-social engineering is having on sociability.

The study of the technological and social fields in general are dealt with separately, which somewhat undermines the arguments and conclusions. The current close relationship between sociability and technology strongly recommends that the two universes be considered in the same approach<sup>4</sup>.

The purpose of this paper is to address the relationship between sociability and efficiency in AI-drive models (even if they are not yet prevalent), in how contemporary economics has brought the notion of efficiency into our personal lives, which is translated into its ability to produce data, by way of technological and social arguments. To do so we explore the synergies of the authors: one a technologist and the other a social scientist.

<sup>2</sup> Data economics considered synonymous with data capitalism.

<sup>3</sup> Big Techs: Google, Facebook, Apple, Microsoft, IBM, Amazon, Alibaba, Tecent and Baidu.

<sup>4</sup> Having in mind: "despite all the attention given to new devices, technology does not determine human behaviour; humans determine how technologies are used" (RAINIE; WELLMAN, 2012, preface s/n).

## 2. Data Capitalism: An Overview

AI technologies represent a new frontier for digital business. In addition to being gradually incorporated into applications, services, products and processes, they are mediating communication and social relationships in all their dimensions. Around 2012, the positive results of the predictive models of deep learning made it possible to identify patterns and correlations based on large amounts of data. While on the one hand they bring benefits that effectively improve 21st century life, on the other they pose ethical and regulatory challenges that are still far from being addressed.

In the field of communication, these technologies introduce unprecedented forms of mediation, characterised by the direct interference of intelligent agents in the online information flow. “The recent explosion of data on the Internet and Web replaces the idea of 'freedom' with the idea of 'relevance'. [...] sophisticated artificial intelligence algorithms individualize Google queries, i.e., results vary depending on the profile of the person seeking the information” (KAUFMAN, 2019, p.46). With specific regard to content selection algorithms, Russell (2019) warns that they are not intelligent, but have the potential to influence billions of people.

Typically, such algorithms are designed to maximize *click-through*, that is, the probability that the user clicks on presented items. The solution is simply to present items that the user likes to click on, right? Wrong. The solution is to change the user's preferences so that they become more predictable. A more predictable user can be fed items that they are likely to click on, thereby generating more revenue. (RUSSELL, 2019, p.8).

AI technologies positively affect the economy increasing productivity and reducing costs, however, in the labor market they pressure variables such as employment and skills, as shown by several studies (IMF, 2018; WEF, 2016; PEW RESEARCH, 2018; MCKINSEY, 2018). Potential unemployment is not the only effect on the labour market, which is perhaps the most socially sensitive; the growing perception is that advances in AI and robotics will radically transform the workplace in the coming decades (BRYNJOLFSSON; MCAFEE, 2014). Research by McKinsey Consulting in 46 countries (2018) finds that at least 30% of the tasks in 60% of all occupations can potentially be automated, and that on average, 15% of current jobs will be replaced or eliminated, the incidence of this being higher in more advanced economies. According to Frey and Osborne, Chui, Manyika and Miremadi and the World Bank, advances in automation threaten 45-57% of all US jobs (IMF, 2018). The White House Council of Economic Advisers predicts that automation will affect 83% of jobs earning \$20 an hour or less (IMF, 2018). The Organisation for Economic Cooperation and Development (OECD), in turn, estimates that 9% of jobs will be affected in the same category. These divergences reflect the respective different perceptions about the interference of social, legal and regulatory frameworks.

There is a high degree of uncertainty in these studies; as intelligent machines are in their infancy, prudently they should be regarded more as trend indicators than actual forecasts.

While there are very different views about the long-term impact on employment as a whole, there is wide argument that the development of Artificial Intelligence and robotics is set to have an enormous impact on the future of human work - driving up productivity, but in the process narrowing or completely shutting down many traditional jobs.(CAMERON, 2017, p.8).

Regardless of the pace and intensity of the digital transformation, as technology advances workers will be relocated to tasks not susceptible to mechanisation, that is, tasks that require human

skills. "For workers to win the race, they will have to acquire creative and social skills" (FREY; OSBORNE, 2013, p. 48).

Mayer-Schönberger and Ramge (2018) coined the term "data capitalism", which highlights the ongoing transition from financial capitalism to data capitalism in a reconfiguration of the economy that is comparable to that of the Industrial Revolution. In the first type of capitalism, information, which was difficult and expensive to acquire, converged around "price", while in the second, information is multiple, complex, fast and cheap. For the authors, price loses its centrality; economic agents use data to identify better matches and explore various dimensions by way of three key technologies: a) a standard language to compare and share data on goods and preferences; b) the ability to identify matches in various dimensions and select the appropriate transactions; and c) the effective capture and use of preferences (assertiveness). These three technologies "have in common that they facilitate the translation of rich data into effective transaction decisions. Underscoring the central role of data, these technologies not only improve our ability to choose based on data, but the technologies themselves are founded on data" (MAYER-SCHÖNBERGER; RAMGE, 2018, p. 64). The real revolution is not the machines that calculate data, but the data itself and how to use them, ponder Mayer-Schönberger and Cukier (2013).

In data-rich markets we have no need to focus on causality, but we can discover patterns and correlations in data that provide us with invaluable insights for decision making (the downside is the risk to privacy). Throughout history, with different perspectives, transactions between agents have been governed by the supply and demand of products and services, regulated by the "price" factor: if supply increases when demand is stable then the price of goods tend to decrease, and if demand is high when supply is stable, then the price of goods tend to increase. The same dynamics can be observed from the viewpoint of the quantity of a good: excess demand generates a shortage of the good and the equilibrium price trend rises; conversely, excess supply generates an abundance of the good and the tendency is for the equilibrium price to fall. The assumption behind these models is that the market is efficient and that agents make rational decisions. In traditional economies, the flow of information converges on price, and consumer preferences are measured from among several other economic variables, i.e., price connects supply and demand.

Data is replacing price as a structural element of the producer-consumer relationship, and currency as a means of payment (MAYER-SCHÖNBERGER; RAMGE, 2018). We are already paying for a lot of data services (Google search, Facebook benefits - social relationships and business platforms), and soon this prerogative could extend to credit card dues, bank fees and telephony costs, industries that concentrate large volumes of customer data. Big Tech firms are the most visible part of the data economy, but not the only one. The drop in revenue in the voice function puts pressure on telecom firms to look for alternative products, and apparently disruptive innovation can be seen in the use of their users' data, particularly in mobile telephony. In 2012 Spanish operator Telefonica created a separate unit - *Telefonica Digital Insights*<sup>5</sup> - to market anonymous and aggregated subscriber location data to retailers and others. In 2016, the leading global payments firm MasterCard created a business unit - *Local Market Intelligence*<sup>6</sup> - to identify and market trends taken from its customer database (903 million cardholders in 210 countries<sup>7</sup>). Banks, perhaps the industry with the most access to personal data, are not yet using their customers' data base to their full extent (they are focusing on reducing costs by migrating from physical to digital

<sup>5</sup> Available at: <https://mobilemarketingmagazine.com/telefonica-digital-launches-telefonica-dynamic-insights>. Accessed on: 2/12/2019.

<sup>6</sup> Available at: <https://newsroom.mastercard.com/press-releases/mastercard-empowers-local-businesses-with-big-data-digital-tools/>. Accessed on: 2/12/2019.

<sup>7</sup> Available at: <https://www.statista.com/statistics/618137/number-of-mastercard-credit-cards-worldwide-by-region/>. Accessed on: 2/12/2019.

platforms). There is strong evidence that banks will soon reinvent themselves as information intermediaries, by maintaining valuable transfer and storage functions.

In the data economy it is essential to label and categorize information, that is, by recording individual product and service references digitally and in detail. The lack of an anthology reduces the number of transactions by limiting the ability to find a match, i.e., the lack of identification filters compromises market efficiency; Mayer-Schönberger and Ramge's prediction is that data itself will drive data anthologies. Amazon was first set up in the mid-1990s, and when its founder and CEO, Jeff Bezos, realized how impractical it was to launch an online store with numerous products, he analysed a list of twenty possible product categories and opted for books: in addition to being a commodity, there were three million books in print worldwide, and the publishers' seasonal catalogues had all been scanned (STONE, 2013).

The commitment to identify and measure user preferences and habits, and then to predict behaviour, is the logic behind technology platforms and applications, online social networks, e-commerce and search engines. This same logic permeates the AI-drive models; mastery of these techniques and access to an expressive database is a competitive factor for firms. Digital platforms are designed to extend the lifespan of their users, thereby generating greater engagement and more data, according to Elizabeth Churchill, director of user experience at Google and president of the Association for Computing Machinery (ACM).<sup>8</sup> One of the challenges is to detect the technical, cultural and social factors that motivate and intensify platform interactivity. Based on data and predictive AI models, firms seek to reduce the unforeseen (and/or the time between the unforeseen event and the appropriate response to it) through more assertiveness. Or even better, the firms can preempt bad events.

Endorsing economists Ariel Eizrachi and Maurice Stucke's contention that machine learning systems are undermining competition, Mayer-Schönberger and Ramge (2018) refute the view that the solution is solely about opening algorithms.

Algorithms alone aren't enough to enable small competitors and new entrants to compete against incumbents because algorithms aren't the raw material. [...] Rather than algorithms transparency, regulators wanting to ensure competitive markets should mandate the sharing of data. (Ibid., p. 167).

The comparative advantage lies in possessing the data, not in algorithm knowledge. If and when data from large players become available to smaller competitors, innovation will tend to spread with data no longer being a barrier to entry. Smart technologies enable the analysis of large numbers of historical operating records, macroeconomic indicators and legislation, that is, large amounts of statistical data. These data can then be segmented into their strategic parameters (by region, by type of customer / consumer, among others), future scenarios can be projected and the accuracy of decisions can be improved (decision - making as a science). Customer / consumer behaviour is not totally random, there are certain hidden patterns in the data: "Pattern recognition is the name of the game - connecting the dots of past behaviour to predict the future" (PASQUALE, 2015, p.20). Amazon illustrates the potential of data mining: with a more accurate understanding of user preferences, they will be able to break away from the traditional retail business model, shopping-then-shipping, to a new model, shipping-then-shopping<sup>9</sup>.

---

<sup>8</sup> Intelligent Systems Conference Opening Lecture, September 4-5, 2019, London.

<sup>9</sup> The patent was granted by the US authorities in December / 2013 and describes the "early transportation" method (Wall Street Journal, January / 2014).

Russell (2019) argues that the way intelligent agents are constructed depends on the nature of the task being faced, which in turn depends on three factors:

First, the nature of the environment the agent will operate in - a chessboard is a very different place from a crowded freeway or a mobile phone; second, the observations and actions that connect the agent to the environment - for example, Siri might or might not have access to the phone's camera so that it can see; and third, the agent's objective - teaching the opponent to play better chess is a very different task from winning the game. (RUSSELL, 2019, p. 43).

Implementation of AI-drive models in firms is a complex challenge. It requires the commitment of its leaders, the acquisition and processing of raw data (labelling, normalisation and standardisation), the choice of models that are appropriate to the business challenges, and infrastructure and teams. It is about delivering new and personalised experiences to customers and partners, and increasing back-office efficiency.

### **What are data**

Data represent the accumulated knowledge about society and are the protagonists of the so-called data capitalism (MAYER - SCHONBERGER; RAMGE, 2018), in which governments and firms, particularly the Big Techs, control much of the generation, mining and use of such data. Business models are based on the possession of and access to big data and the ability to extract information from this data using AI technologies.

Especially in the last twenty years or so, people have increasingly started to ask themselves what they can do with all this data. With this question the whole direction of computing is reversed. Before, data was what the programs processed and spit out - data was passive. With this question, data starts to drive the operations; it is not the programmers anymore but the data itself that defines what to do next. (ALPAYDIN, 2016, p.11).

Any interaction with digital technologies leaves “traces”. Some are voluntary, like online social media posts – on Facebook, Twitter, Instagram – while others are involuntary, such as the information stored in digital databases of credit card, digital banks, transport vouchers, mobile phone communications, online access to medical examination results, or simply when a digital camera captures our image in a public space. These “tracks” can be used by the original platforms, “reused” by third parties, or combined by merging data sets for a variety of purposes.

The data that feeds the dominant business model in the contemporary economy doesn't just come from platforms operated directly by digital giants. The report by the US Presidential Council of Science and Technology Advisers shows that while specialized companies have always collected information to better understand consumers, now, in addition to the mass of information collected through the use of social networks, there is a “data fusion” which is when data from different sources are brought into contact and new facts emerge. (ABRAMOVAY; ZANNATA, 2019, p. 431).

About 80% of globally accumulated data is unstructured (text files, social networks, SMS-text messaging, geolocation data, chats, phone recordings, MP3 files, scanned photos, audio and video files, among others).<sup>10</sup> As Mayer-Schonberger and Cukier remind us (2013, p. 96), “The enthusiasm over the ‘internet of things’ - embedding chips, sensors, and communications modules

---

<sup>10</sup> Examples of structured data: Electronic spreadsheets (Excel, Numbers), databases, XML files (NF-e), and CSV (comma-separated values) files.

into everyday objects - is partly about networking, but just as much about ‘datafying’ all that surrounds us”. Figure 1 shows the exponential growth curve of global data from 2005 to 2019.

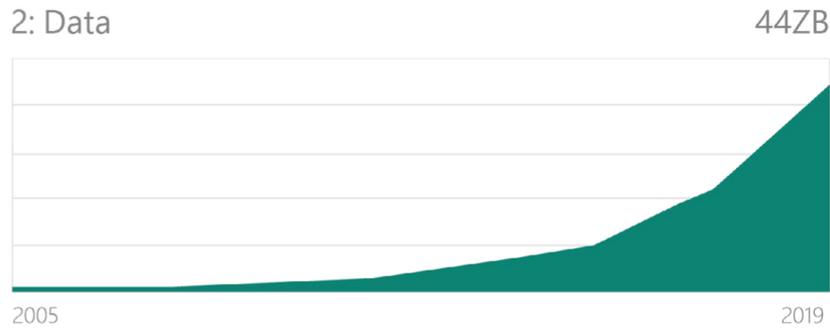


Fig. 1 – Data Growth Curve. Source: IDC Digital Universe Forecast, 2014.

Considering only one of the major technology companies, the volume of data is already astronomical: Google now processes over 40,000 search queries every second on average, which translates to over 3.5 billion searches per day and 1.2 trillion searches per year worldwide. The volume of searches increased 17,000% year-on-year between 1998 and 1999, 1000% between 1999 and 2000, and 200% between 2000 and 2001. Google's search continued to grow at rates of between 40% and 60% between 2001 and 2009, when it began to decelerate. It has stabilized at a rate of 10% to 15% in recent years (source: Google Search Statistics) (Fig. 2).

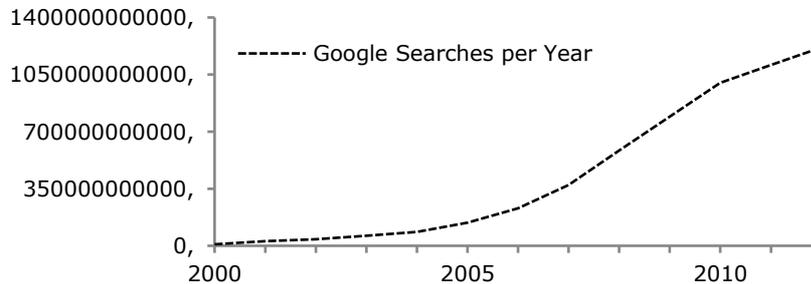


Fig. 2. Google Searches per year. Source: Google Search Statistics.

Agrawal, Gans and Goldfarb (2018) highlight three functions performed by data: a) first, we have input data, which feeds the algorithms and is used in the forecasting process; b) second, we have training data, used to improve the algorithms; and c) third, we have feedback data, used to improve the performance of algorithms based on user experience. Media Lab MIT computer scientist, Alex Pentland (2015), argues that big data offers a chance to see society in all its complexity; for Pentland, once we have developed a view of more precise patterns of human life, we can hope to understand society in a way that is more appropriate to our complex, interconnected network of human beings and technology.

### Deep learning: a brief introduction

The so-called deep learning process is inspired by the functioning of the brain and so it is also known as neural networks. For several decades the dominant approach in the field of AI was based on logical computing programmes, thus marginalising the view based on machine learning.

Yann LeCun remembers that in the 70s and early 80s "you could not publish a paper that even mentioned the words 'neural networks' because it would immediately be rejected by your peers" (FORD, 2018, p.122).

A small group of researchers, however, insisted on the path of neural networks, particularly Geoffrey Hinton, Joshua Bengio and LeCun, and they gained definite recognition in 2012 by winning the ImageNet Competition<sup>11</sup>.

The first two years of the competition saw measured improvement, as the error rate dropped from 28% in 2010 to 26% in 2011. [...] But a paradigm shift happened in 2012, when an inelegant and underdog submission became the undisputed winner of the ImageNet Challenge. The submission was a deep neural network, and it came in with an error of 16%, far below the previous year's rate of 26%. (GERRISH, 2018, p. 135).

In subsequent years, deep learning became ubiquitous and received significant investments from leading technology firms, which resulted in massive profits being generated. "In the last few years, deep learning has generated enough profit for Google to cover the costs of all its futuristic projects at Google X, including self-driving cars, Google Glass, and Google Brain," comments Terrence Sejnowski (2018, p. ix). Deep Learning is used today in more than 100 Google services, from Street View to Gmail's automatic replies.

Deep learning is about forecasting and permeates many of the 21st century's activities. When we enter a query in Google, it selects a personalised response and those ads that are appropriate to the user's profile. It also translates text from another language and filters out unsolicited emails (spam). Amazon and Netflix recommend books and movies using the same process, Facebook uses deep learning to decide which updates to display in its News Feed, and Twitter does the same for tweets. When we access a computing device in any of its formats, we are probably accessing a deep learning process. The current exponential growth in data has made it difficult to write computer programmes. Amazon cannot code the preferences of its customers in a computer program, just as Facebook does not know how to write a program for identifying the best updates in its News Feed. Netflix may have 100,000 DVD titles in stock, but if customers do not know how to find their preferences it is of no use.

We use machine learning when we believe there is a relationship between observations of interest but do not know exactly how. Because we do not know its exact form, we cannot just go ahead and write down the computer program. So our approach is to collect data of example observations and to analyse it to discover the relationship. (ALPAYDIN, 2016, p. 29).

The large amount of data is not the only restrictive factor. Humans find facial image recognition relatively easy, for example, but they cannot explain it (tacit knowledge), which makes it impossible for a computer programme to be written. When analysing different pictures of a person's face, an AI system captures the specific pattern for that person and then checks that pattern against a given image.

From the society standpoint, the reason for deep learning's success is its predictive ability. Anticipating future scenarios and their likelihood is the challenge of any activity. Traditional statistical models are based on samples and error reduction methods that focus on causality. In addition to being costly, these models are not feasible on a large scale (big data). By correlating large amounts of data, AI algorithms are able to more assertively estimate the likelihood of a

---

<sup>11</sup> <http://www.image-net.org>

tumour being a particular type of cancer, or the likelihood of an image being of a dog, or the likelihood of a piece of equipment needing to be replaced on a particular date.

In the second decade of the 21st century, the convergence of various technologies has promoted results that are superior to any previous predictions (albeit falling short of science fiction!). Intelligent systems are performing tasks that until recently were human prerogatives, and in some cases, with faster and better results. But it is only a decade of “revolution”, and machines are still restricted to predicting scenarios (predictive capacity) based on large data sets and performing specific tasks under the direct supervision of computer science experts. Without detracting from their merit, these apparent advances, which are so highly acclaimed by the media, mean that “simple” successful implementations are not really conceptual (or scientific) breakthroughs. “They are applications of conceptual breakthroughs that have happened long before - from the earliest deep learning systems and seizure networks dating back to the late 1980s and early 1990s,” explains Stuart Russell (FORD, 2018, p. 44).

What fires the enthusiasm of the AI community is that artificial intelligence is the core business of technology giants like Google, Facebook, Amazon, Baidu, and Ali-baba, and is also being driven by other powerful industries, such as finance and retail, in addition to governments. All major automakers are investing heavily in driverless cars; in 2017 General Motors paid \$1 billion for Cruise Automation, a Silicon Valley start-up with a focus on driverless cars, and invested an additional \$600 million in research. (SEJNOWSKI, 2018, p. 5).

Fuelled by unprecedented volumes of investment and grounded on an increasingly solid theoretical basis, there is no disagreement as to the enormous benefits to mankind of the advance of AI. On the other hand, it does raise ethical and social issues of great impact, such as the elimination of jobs in labour-intensive sectors and the expansion of economic activities that require less manpower.

We are at a tipping point where it is still possible to balance the negative and positive effects by promoting progress with less inequality. The challenge is to build an inclusive man-machine partnership for the whole of society. “And if the idea that really put a glimmer in researchers’ eyes bears fruit, machine learning will bring about not just a new area of civilisation, but a new stage in the evolution of life on Earth” (DOMINGOS, 2015, p. 22).

### **3. The Exponential Evolution Of Storage, Processing And Transmission Technology**

Over the years, the ability to transform analog data into digital data has expanded exponentially. For reference: in 1986 only 1% of existing information was digitised (books, movies, data, and so on); this percentage currently exceeds 95% of the total data volume (HILBERT, LÓPEZ, 2012) and estimates point to 175 Zettabytes of data being stored worldwide by 2025 (IDC White Paper, 2018).

Until 2018, the “input” of data in organizations was carried out by employees called data entry clerks, whose function was to scan the information manually, whether for administrative, financial or industrial operation. The information came basically from the manufacturing process

that fed Production Planning and Control systems. - PCP<sup>12</sup> (CHIAVENATO, 2005). Although the manufacturing process had a level of mechanisation and automation, it relied on manual processes for entering the industrial production information to be further processed electronically, thus increasing the possibility of mistakes and adding time and a potential loss in efficiency.

Voice, image, impulses, vital signs, temperature, texture, odours, altitude, proximity, light, flow, smoke; these and many other pieces of information are now transformed into digital signals by advanced sensing systems, thereby exponentially increasing the ability to perceive these inputs and generate output information related to a specific physical quantity (RAVI, 2018). The result is the creation and exponential storage of large volumes and a wide variety of data.

Both physical and logical storage technologies continually outperform each other. Microsoft, for example, announced last November its new Project Silica, which deals with a company-created glass unit that is 7.5 x 7.5 cm and 2 mm thick. The size makes it large enough to store 75.6 GB<sup>13</sup>.

As an illustration of the evolution of the sensing and consequent digitalisation of information, not long ago traffic police manually registered infractions on paper to be later digitised and subsequently processed. This task is now performed digitally by way of radar (radio detection and ranging), known as "sparrows", the occurrence of the violation is registered by way of different sensing technologies<sup>14</sup>. In 2018 in the city of São Paulo there were nearly 11 million infractions, of which 77.15% were processed digitally<sup>15</sup>. Electronic processes allow predictive, analytical and regulatory actions to be implemented.

Traditional call centres are another well-referenced example, where customer interaction was (and in part still is) by way of exhaustive calls until the caller was connected to an attendant who interacted with the person based on previously defined scripts. When the call was successfully handled and completed, this process recorded occurrences in the company's computer system, following which the data were processed and, after much effort, analysed. Sensors, voice recognition systems, image, social engineering and many other mechanisms are currently used to "hear" the voice of the digital form of client and instantly produce large datasets, generating efficiency in the processes that support actions in real time (predictions, historical analyses), as well as producing future projections and actions.

The usefulness and timeliness of exponentially accumulated data become relevant when such data can be processed and transmitted to appropriate destinations, whether local or remote. Digital systems, therefore, can be graphically illustrated in the form an equilateral triangle, as shown in Figure 3.

---

<sup>12</sup> PCP process that channels and absorbs information to allow decisions to be taken about what to do, when to do it and how much to do in terms of production, making the planning of machines possible.

<sup>13</sup> Available at: <https://news.microsoft.com/innovation-stories/ignite-project-silica-superman/>. Accessed on: 2/12/2019.

<sup>14</sup> Available at: <https://www.sciencedirect.com/topics/engineering/radar-systems>. Accessed on: 2/12/2019.

<sup>15</sup> Source: Portal Mobilidade Segura – São Paulo Municipal Administration.

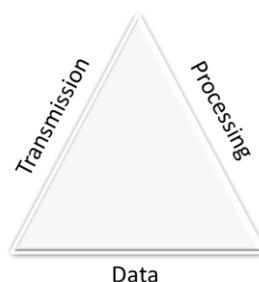


Fig. 3: Digital Systems Representation. Source: authors.

The exponential growth of processing power and increase in cost reductions following Moore's law<sup>16</sup> has created the conditions needed for sophisticated algorithms to handle a huge volume of structured or unstructured data of unprecedented diversity. The emergence of the Graphics Processing Unit (GPU) in 2007 was a definite milestone, expanding processing capacity by 1000 times over a 10-year period, and enabling the advancement of neural networks and deep learning models (OLENA, 2018). It is noteworthy that the technologies used to transform analog information into digital data and store them, in addition to their exponential processing capacity, were accompanied by expansion in transmission capacity; an extensive connection network has made it possible to disseminate innovative technologies beyond the environment of laboratories or specialised organisations. Long-distance data transmission technology increased in 1990 from 2.5G to the current 10Tb, and is projected to reach 100Tb. Similarly, local area network (Ethernet), which in the 1990s reached 100Mb, in 2019 reach as much as 1Tb (MIYAMOTO, SANO, KOBAYASHI, 2013). It is also important to mention mobile transmission technologies, which in the 1990s were 100 Kbps “Narrow Band” and currently, with the launch of 5G, reach as much 1Gbps (the cellular network will allow speeds of 1 Gbps and conversations between drones) (CONESULNEWS, 2018).

Communication interactions, which were once predominantly offline, have migrated to the online environment and currently demand real-time technologies. The concepts of asynchronous and synchronous<sup>17</sup> become confused by the demand for the immediacy that is required by digital interactions. The transmission capacity associated with a vast connectivity network allows businesses or organisations to use their capabilities and exponential storage processing capacity to provide unprecedented, distributed computing solutions (ISMET AKTAŞ, 2019), thus enabling a technological infrastructure that welcomes the extraction of maximum digital efficiency in the interactions required by its core activities, whether commercial, social, scientific, academic, industrial or otherwise. To measure such efficiency, as much data should be assessed as potential interactions and from that stage the scan target can be obtained. For example, we describe a case related to a call centre<sup>18</sup> in order to evaluate the maximum scan efficiency scenario (Fig. 4):

Call centre efficiency gain:

- 100 position call centre

<sup>16</sup> Available at <http://www.moorelaw.org/>. Accessed on: 02/12/2019.

<sup>17</sup> Available at <https://techdifferences.com/difference-between-synchronous-and-asynchronous-transmission.html>

<sup>18</sup> A call centre is a physical structures that aims to centralize the reception of telephone calls, automatically distributing them to the attendants to serve end users with different purposes (market research, sales, retention and other services). Corporations / institutions use call centres as relationship channels with their clients / general public.

- Each position receives an average of 25 calls per hour
- Each call, average length 2.4 minutes, would be the equivalent of approximately 432 spoken words (3 words per second), which, if the entire conversation is scanned, would take around 8 minutes (average scan time: 52 words per minute on a computer<sup>AMP1</sup>). In this case, what could be typed would be somewhere around 100 of the 432 words generated during the interaction, with losses in both quantity and potentially in interpretation.
- Voice recognition technology has the potential to capture 100% of what the customer says, digitalise it and process it in real time, helping the attendant solve the customer's request. It also ensures the quantity, variety, and data integrity of these interactions, which is useful for studies, analyses, and potentially for predicting or for preparing prediction models.
- Considering in this case that 1 word has an average of 6 characters, or 6 bytes (AMP2), a call centre operating at maximum digital efficiency should capture around 100Mb of conversation from its customers every day.
- A traditional call centre, where operators input and register systems manually, would have a maximum digital efficiency of 28%, i.e., it generates a "digital entropy" of 72%

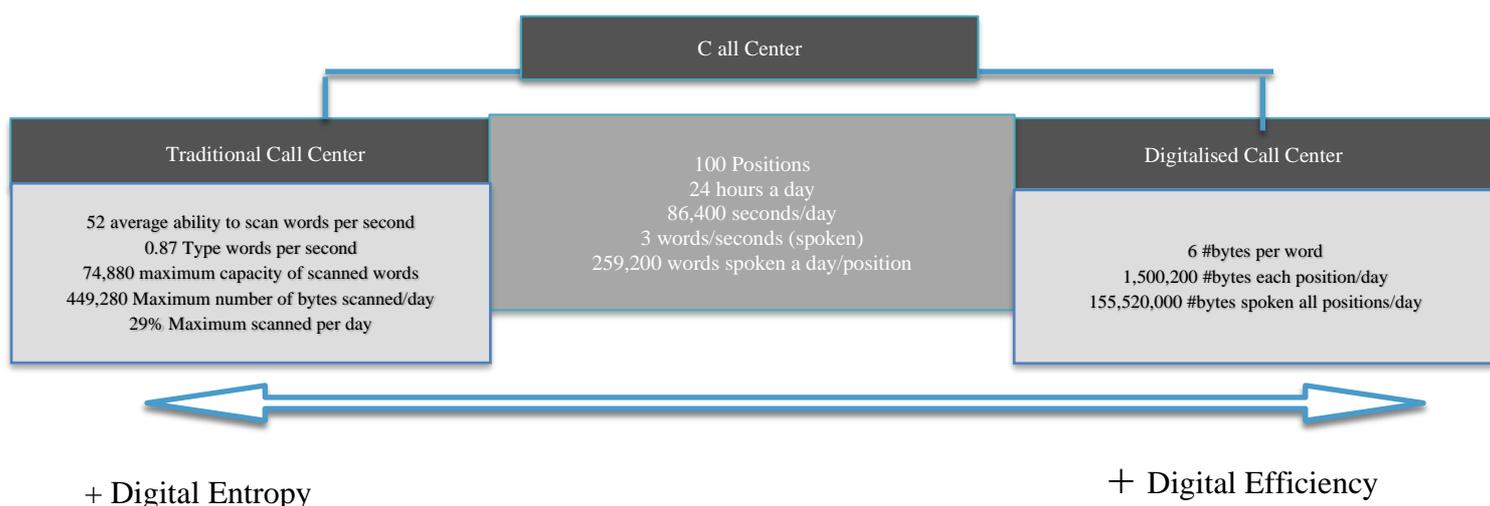


Fig. 4 – Call Centre. Systems: Source: authors.

A fully digital call centre system (ChatBot), with 100% of the interactions digitised, would be operating at maximum efficiency. This is applicable to any social interaction, considering that digital data-capture, processing, storage and transmission systems provide us with the opportunities for more efficient operations, feedback mechanisms and systemic evolution. The digital system will work at full capacity, better and more efficiently, the more data is generated in the H-H, M-M or M-M interaction.

## 4. Final comments

The business models in data capitalism demand large amounts of data, which in turn is generated from H-H, H-M, M-M interactions. The essential prerequisite is to promote interactions by way of coupled technologies that capture and process these interactions and digitally record the data. The greater the capacity for capturing, storing, processing and transmitting data, the more efficient the AI-driven models are.

The measure of efficiency should be calculated and controlled against the maximum potential of the digital data produced in these interactions. The information generated via H-H, H-M and M-M interactions that is not captured, or is lost because it is not digitalised, can be considered system entropy. The efficiency of AI-drive models is a function, in part, of the quantity and quality of the data, which are derived from the different interactions during their processes, services or systems. To achieve the maximum potential of these models and take advantage of the technological evolution that includes the advancement of AI, by their very nature H-H, H-M and M-M relationships must employ mechanisms that operate at maximum efficiency when transforming analog information into digital information. The more data extracted from interactions, the more efficient it will be for digital systems, and the more efficient the data-based business models like Google, Facebook, Amazon, Netflix, Spotify will be, to name the most popular.

Therefore, considering that we live in a digital age, when there is any interaction with a customer, patient, teacher, student, supplier, carrier, or others, this interaction will only be digitally effective if it generates data sufficiently and consistent at some point (in real time or later), through centralised or distributed processing and thus makes these relationships into useful information.

For traditional firms, the current business environment is a major challenge. “Competing with digital rivals involves more than deploying enterprise software or even building data pipeline, understanding algorithms, and experimenting. It requires rearchitecting the firm’s organization and operating model” (IANSITI, LAKHANI, 2020. p. 62). The divisions across functions and products, typical of the industrial economy, it's not efficiency in the data economy. The compartmentalization in silos compromises the efficiency of AI-drive models which demand integrate data base.

The good news is that traditional operating constraints could be overcome. Referring to digital companies, Iansiti and Lakhani (2020) argue:

Every time we use a service from one of those companies, the same remarkable thing happens: Rather than relying on traditional business processes operated by workers, managers, process engineers, supervisors, or customer service representations, the value we get is served up by algorithms. (Ibid, p.62).

Previous systems were monolithic, limiting operational growth to hardware technology. Currently, the systems are horizontal, replicable and unlimited growth microservices. The infrastructure can be outsourced via fixed-term contracts. We are moving from a model where companies did everything to an outsourcing scenario at all stages through APIs (AWS / Amazon, Azure / Microsoft, IBM Cloud, Alibaba Cloud, Cloud Google). Another advantage is that AI technologies make it possible to predict unexpected events and anticipate procedures to eliminate or reduce negative impacts.

Scale, scope, and learning have come to be considered the essential drivers of a firm's operating performance. [...] After hundreds of years of incremental improvements to the industrial model, the digital firm is now radically changing the scale, scope, and learning paradigm. AI-drive processes can be scaled-up much more rapidly than traditional processes can, allow for much greater scope because they can easily be connected with other digitized businesses, and create incredibly powerful opportunities for learning and improvement. (IANSITI; LAKHANI, 2020, p. 64).

For traditional firms digital transformation is not a simple process, it requires huge investment in management, time and financial resources. However, it is the only way to remain competitive and survive in the 21st century market.

## 5. Bibliography

- ABRAMOVAY, Ricardo; ZANNATA, Rafael. Dados, vícios e concorrência: repensando o jogo das economias digitais. *Estudos Avançados*, v. 33, n. 96, p. 421-446, 2019. Available at: <https://www.revistas.usp.br/eav/article/view/161303>. Accessed on: 09/15/2019.
- AGRAWAL, Ajay; GANS, Joshua; GOLDFARB, Avi. *Prediction machines: the simple economics of Artificial Intelligence*. Boston, MA: Harvard Business Review Press, 2018.
- ALPAYDIN, Ethem. *Machine learning*. Cambridge, MA: MIT Press, 2016.
- BRYNJOLFSSON, E.; McAfee, A. O negócio da Inteligência Artificial. *Harvard Business Review*, 6 de nov., 2017. Available at: <http://hbrbr.uol.com.br/o-negocio-da-inteligencia-artificial/>. Accessed on: 03/12/2018.
- CAMERON, Nigel M. de S. *Will robots take your job?* Cambridge, UK: Polity Press, 2017.
- FREY, Carl Benedikt; OSBORNE, Michael. *How Susceptible are Jobs to Computerisation?* 2013. Available at: [https://www.oxfordmartin.ox.ac.uk/downloads/academic/The\\_Future\\_of\\_Employment.pdf](https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf). Accessed on: 10/20/2018.
- FRISCHMANN, Brett; SELINGER, Evan. *Re-Engineering Humanity*. UK: Cambridge University Press, 2018.
- GALLOWAY, Scott. *Os Quatro: Apple, Amazon, Facebook e Google - O segredo dos gigantes da tecnologia*. São Paulo: HSM, 2017.
- IANSITI, M.; LAKHANI, Karim R. Competing in the Age of AI: How Machine Intelligence changes the rules of business. HBR, From the January–February 2020 Issue. Available at: <https://hbr.org/2020/01/competing-in-the-age-of-ai>. Accessed on: 02/01/2020.
- IMF Working Paper: *Should We Fear the Robot Revolution?* 21/05/2018. Available at: <https://www.imf.org/en/Publications/WP/Issues/2018/05/21/Should-We-Fear-the-Robot-Revolution-The-Correct-Answer-is-Yes-44923>. Accessed on: 10/31/2018.
- KAUFMAN, Dora. *A inteligência artificial irá suplantar a inteligência humana?* São Paulo: Estação das Letras e Cores, 2019.

- MCKINSEY GLOBAL INSTITUTE. The new frontier: Agile automation at scale By Federico Berruti, Geet Chandratre, and Zaid Rab - Large-scale automation of business processes requires a new development approach. 2018. Available at: <https://www.mckinsey.com/business-functions/operations/our-insights/the-new-frontier-agile-automation-at-scale>. Accessed on: 11/02/2018.
- MAYER-SCHONBERGER, Viktor; CUKIER, Kenneth. *Big Data: a revolution that will transform how we live, work, and think*. New York, NY: Houghton Mifflin Harcourt, 2013.
- MAYER-SCHÖNBERGER, Viktor; RAMGE, Thomas. *Reinventing capitalism in the age of big data*. London: John Murray, 2018.
- PASQUALE, F. *The black box society: the secret algorithms that control money and information*. Cambridge, MA: Harvard University Press, 2015.
- PEW RESEARCH CENTER. In Advanced and Emerging Economies Alike, Worries About Job Automation: many fear robots, computers will eliminate jobs, increase inequality, 2018. Available at: <http://www.pewglobal.org/2018/09/13/in-advanced-and-emerging-economies-alike-worries-about-job-automation/>. Accessed on: 10/25/2018.
- RAINIE, Lee; WELLMAN, Barry. *Networked: The Social Operating System*. Cambridge: MIT Press, 2012.
- RUSSEL, Stuart. *Human Compatible: artificial intelligence and the problem of control*. USA: Viking, 2019.
- SCHONBERGER, Viktor Mayer; RAMGE, Thomas. *Reinventing Capitalism in the Age of Big Data*. NY: Basic Books, 2018.
- STONE, Brad. *The everything store: Jeff Bezos and the age of Amazon*. New York, NY: Little Brown, 2013.
- WORLD ECONOMIC FORUM. *The Future of Jobs: Employment, Skills and Workplace Strategy for the Fourth Industrial Revolution*, 2016. Available at: [http://www3.weforum.org/docs/WEF\\_Future\\_of\\_Jobs.pdf](http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf). Accessed on: 10/25/2018.
- AMP1: Estamos digitando mais rápido nos smartphones e quase chegando no nível dos teclados físicos by Victoria Song, 2019.
- AMP2 - Calculadora Conversão Carâters para Bytes – ExtraConversion - Available at: <http://extraconversion.com/pt/armazenamento-de-dados/caraters/caraters-para-bytes.html>. Accessed on: 12/20/2019.