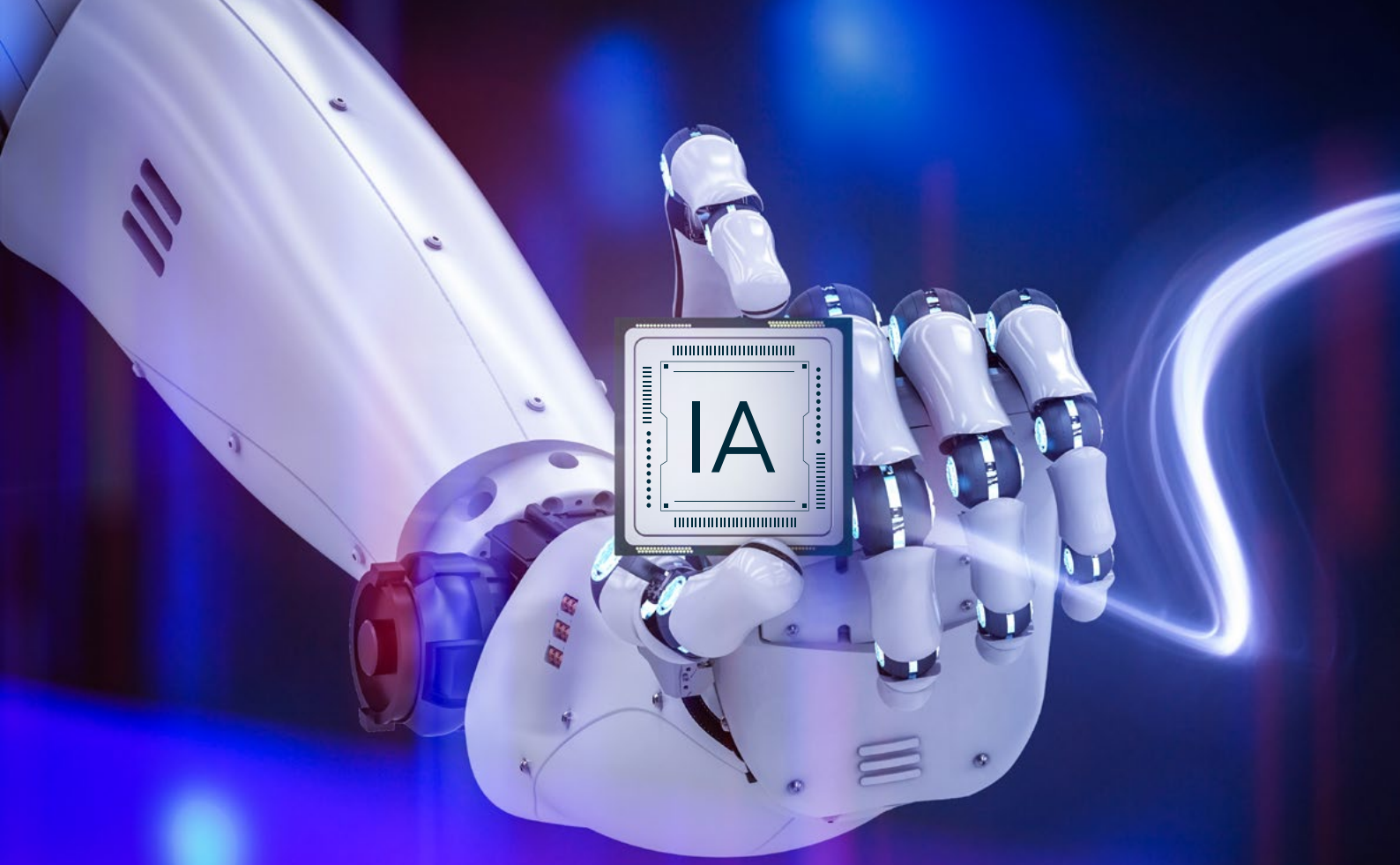


GOOD PRACTICES IN **ARTIFICIAL INTELLIGENCE**

A GUIDELINE

Coordinated by

Eduardo Magrani

Tatiana Campello

DEMAREST

Authors

Eduardo Magrani

Renan Medeiros de Oliveira

Tatiana Campello

Support

Published in November 2021

AUTHORS

Eduardo Magrani

emagrani@demarest.com.br

+55 21 3723 9811

Partner at Demarest responsible for the areas of Intellectual Property and Innovation, Data Privacy, Technology and Cybersecurity. Post-Doctoral degree from the Technical University of Munich (TUM) in artificial intelligence and data protection and Associate to the Munich Center for Technology and Society (MCTS/TUM). Doctorate and Master's degree in Law from PUC-Rio, focusing on Digital Democracy, Internet of Things, Artificial Intelligence, Data Protection and Ethics. Affiliate at the Berkman Klein Center of Harvard University. President of the Instituto Nacional de Proteção de Dados in Brazil. Senior Global Fellow at the Konrad Adenauer Foundation's international cooperation program (European and International Cooperation team for Global Innovation Policy, Digitalization and Artificial Intelligence); Senior Fellow at Humboldt University of Berlin, at the Alexander von Humboldt Institute for Internet and Society (2017); Member of the Portuguese-German Association of Jurists (DLJV-Berlin) and of the German-Brazilian Association of Jurists (DBJV – Deutsch-Brasilianische Juristenvereinigung e.V.); Coordinator of the Instituto de Tecnologia e Sociedade do Rio de Janeiro (2017-2019); Associate Researcher and Consultant in Artificial Intelligence at Centro Latam Digital Center, in Mexico; Consultant in Digital Law at the University of Lisbon School of Law, associated with the Cyberspace Legal Research Center; Visiting Professor at the College of Law of Universidade Católica Portuguesa in the subjects of Internet of Things and Big Data & Law; Associate Researcher at the Law Schools Global League and member of the Global Network of Internet & Society Research Centers; Professor and consultant at Instituto CESAR in the city of Recife and for the Stricto Sensu Post-Graduate Program in Law – Master's and Doctorate Degrees – of Universidade Regional Integrada do Alto Uruguai e das Missões (URI) in the State of Rio Grande do Sul; Professor and Project Lead Researcher in the areas of E-democracy, Internet of Things and Disruptive Technologies at the FGV Technology and Society Center (2010-2017); Co-developer of the first comprehensive legislation on the Internet in Brazil: the Internet Civil Framework; Author of several books and articles in the area of Law & Technology, among which can be highlighted the "Cultura Digital" Collection, comprising the four works "Internet das Coisas", "Entre Dados e Robôs", "Democracia Conectada" and "Vida em Rede".

Renan Medeiros de Oliveira

Master's degree in Public Law and Bachelor of Laws degree from Universidade do Estado do Rio de Janeiro (UERJ); Specialization in Public Law from Pontifícia Universidade Católica de Minas Gerais (PUC Minas); Fellow of the Instituto Nacional de Proteção de Dados (INPD) and Academic Advisor to the Committee of Ethical Compliance in Research Involving Human Beings at Fundação Getulio Vargas (CEPH/FGV). Renan worked at the Center for Technology and Society (CTS) at FGV Direito Rio (2017-2018) and was a researcher at the Fundamental Rights Clinic of the UERJ Law School – UERJ Clinic of Law (2014-2019) and at the UERJ Economic Regulation Laboratory – UERJ Reg. (2018).

Tatiana Campello

tcampello@demarest.com.br

+55 21 3723 9851

Partner at Demarest responsible for the areas of Intellectual Property and Innovation, Data Privacy, Technology and Cybersecurity. Tatiana is recognized in her field of activities and participates in different institutions, such as: Vice President of Licensing Executives Society International (LESI -2018-2020); former President of LES Brasil (2014-2015) and current member of the association's board, President of the Lex Mundi Privacy group, Treasurer Director of the Brazilian Association of Intellectual Property (ABPI – 2018-2021); President of the Lex Mundi Intellectual Property Group (2015-2017), Member of the Steering Committee of the project "Effective implementation and regulation under the new Brazilian data protection law (LGPD)", of the Centre for Information Policy Leadership (CIPL) with Instituto Brasiliense de Direito Público (IDP), Specialist of the ABPI Dispute Resolution Chamber Regarding Domain Names (CASD-ND), Mediator of the ABPI Mediation Chamber, former Co-coordinator of the ABPI Dispute Resolution Committee (2014-2017) and for many years Co-coordinator of the ABPI Technology Transfer and Franchises Commission; Vice-Coordinator of the ABPI Copyright Commission. Tatiana also participates in the International Association of Privacy Professionals (IAPP) and the International Association of Trademarks (INTA), among others.

CONTENTS

INTRODUCTION: ENHANCING THE USE OF TECHNOLOGY

Artificial Intelligence (AI) is already a reality whose uses and benefits are apparent both for individuals and organizations. It is present in our daily lives and in our vocabulary and its countless utilities and applications become incrementally clearer. The use of chatbots, which help users with questions related to products and services, and virtual assistants, which perform tasks in response to commands and offer several types of assistance, are examples of the use of technology that benefit both the companies and the users.

According to a 2021 survey conducted by IBM[1], one third of the information technology (IT) professionals surveyed stated that their companies already use AI. The data also indicate that 74% of the companies are exploring or using the technology and that they have accelerated the implementation process due to the COVID-19 pandemic.

Despite this, the topic is not always discussed based on a clear concept of what AI really is, and ethical-legal aspects already consolidated in the local and international scenarios are also not considered in their entirety. There are ethical principles that underlie the use of this technology, which must be observed from the moment of creation to its practical application. That is the case, for example, of the principle of non-maleficence, which, in general terms, establishes that Artificial Intelligence should be used without causing foreseeable and intentional damage, as well as establishing that risks should be mapped and, as far as possible, avoided.

It is not uncommon that the debate on AI and even its application is not accompanied by a clear conceptualization, which is important to defining the applicable ethical-legal system, differentiating it from other technologies.. In general terms, we can say that Artificial Intelligence consists of algorithms trained by data to perform certain tasks autonomously – tasks that, until then, could only be performed by human beings. Hence, human intelligence and Artificial Intelligence are complementary. After the programming of the machine, however, AI is often able to independently perform tasks such as analyzing databases, identifying patterns and recognizing faces and objects. On the one hand, AI is, as many people presume it to be, complex. Choosing mathematical

[1]IBM WATSON. **Global AI Adoption Index 2021**. [s.l.]: IBM Watson, 2021, p. 3-4.

models and architectures of well-known artificial neural networks requires specialized technical knowledge. On the other hand, however, it is extremely practical, intelligible and applicable. Once created, its use can be simple.

Thus, this technology is very useful and provides countless advantages to the companies, such as reducing production time and costs, bringing clients closer and increasing productivity, aspects that, when combined, generate high economic impacts. In recent decades, we have witnessed an intense digital transformation, which has grown even further in recent years with the use of algorithms and their derivations. This has led several organizations and companies to dedicate themselves to business transformation by integrating new technologies into their operations.

This commercial and digital transformation has also been positive for attracting new consumers, who seek services and products adapted to the new reality. The levels of demand, publicity and public scrutiny to which companies are subject require processes that respect ethical tenets and legal rules. A company with processes adapted to local and international standards for the use of Artificial Intelligence also tends to attract more qualified professionals, creating a circle of generation of benefits for all parties involved. Companies improve their processes, products and services by applying technology in an ethically responsible manner; consumers have access to better quality products and services; professionals work in establishments that respect the existing rules in order to provide better services and seek companies that are in accordance with their level of qualification; finally, in this cycle, companies attract more consumers, increase sales figures and arouse the interest for collaboration from even more qualified professionals. In addition, there is also the possibility of establishing commercial partnerships, whether at the local or international level, with companies equally adapted to new technologies and concerned with ethical-legal issues related to their use.

In this regard, McKinsey Global Institute emphasizes that AI has great potential to contribute to economic activity around the world. It is estimated that by 2030 the technology will add 13 trillion dollars to economic production, with a 1.2% increase of the annual global GDP[2].

[2] MCKINSEY GLOBAL INSTITUTE. **Note from the AI frontier**: modeling the impact of AI on the world economy. Discussion paper, McKinsey&Company, 2018, p. 2-3.

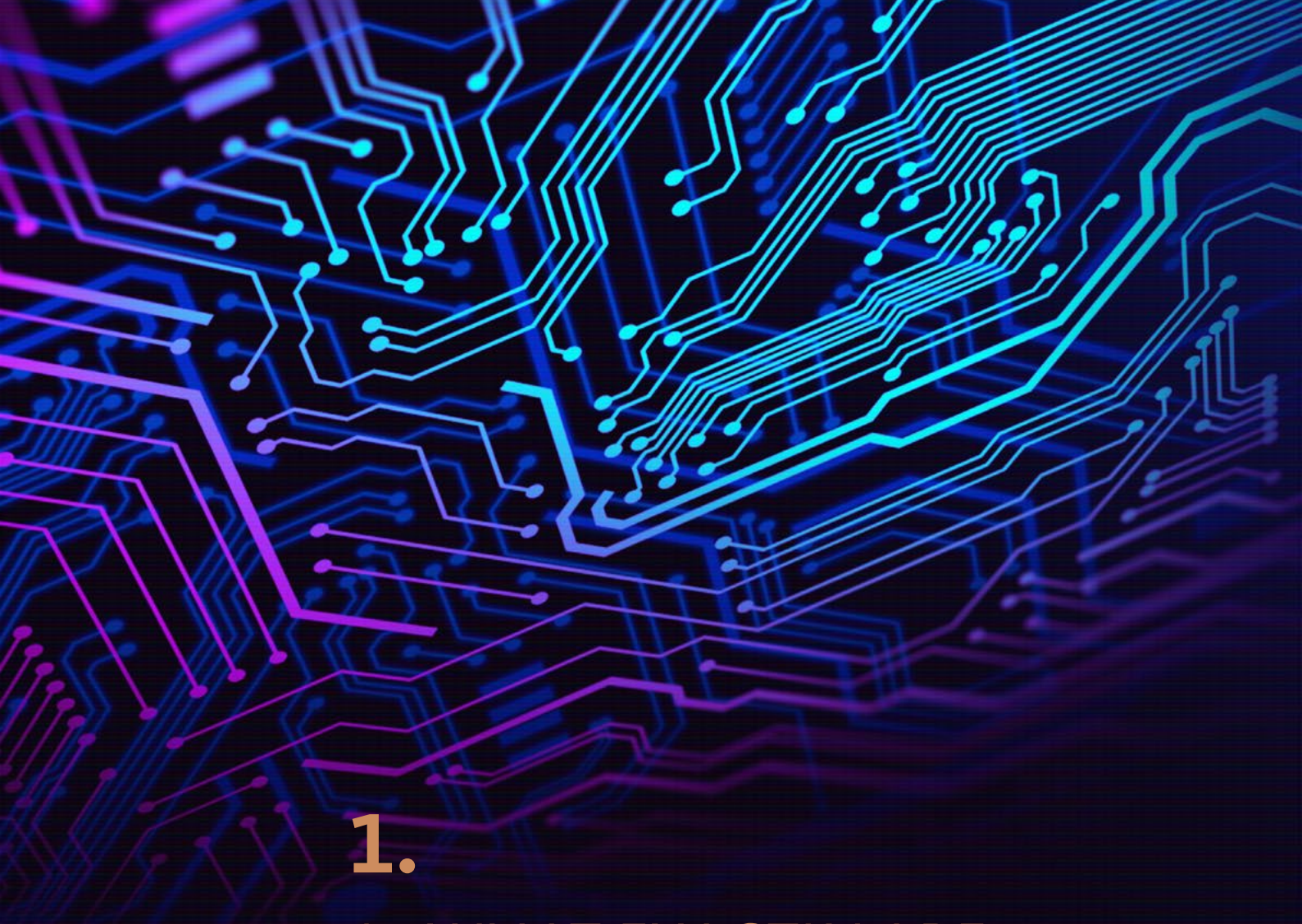
Despite all of these potential benefits, integrating Artificial Intelligence into a company's routine is not a trivial task and cannot be done in a careless or hasty manner. While the introduction of AI can bring major advantages, its careless use carries high risks. In other words, what is a great opportunity for growth can also be a threat and generate conflicts from the perspective of civil liability.

This leads us, once again, to the need for knowledge and practical observance of the ethical-legal standards that govern Artificial Intelligence in all processes and stages of application in a company. In general, it is necessary to take a preventive stance, attentive to possible risks and adopting measures to avoid them before they manifest. Accordingly, the ethical and responsible use of AI will ward off the main risks of its use and reinforce its potential.

As happened with the Brazilian General Data Protection Law (LGPD - Law no. 13,079/2018), which resulted in the mobilization of companies to adapt to the legal requirements related to the protection of personal data, Artificial Intelligence will soon be on the agenda of several organizations, which will seek its integration into their processes and adaptation to the legal regulations. Hence, companies that take notice of the technology's growing momentum, which is already apparent to the more attentive observer, will be able to define, in advance, measures to understand the ethical-legal contours and parameters of the technology and the ways in which it can be incorporated into their business. In a short time, these measures will constitute the grounds for a great competitive differential.

Within this context, the purpose of this Guide is to provide a comprehensive introduction to Artificial Intelligence, addressing its concept, the principles and rules involved and the consequences of its use. The approach used here is aimed at the practical application of the technology, focusing on serving as an aid for companies that are willing to implement it in their products. Thus, we seek to provide guidelines for the ethical and responsible use of AI, intensifying its potential applications and ensuring compliance with mandatory measures.

To achieve this purpose, we begin this Guide with a detailed explanation of Artificial Intelligence. In chapter 1, we address the concept of technology, its characteristics and modalities. Next, we present and explain the principles that must be observed for an ethical, responsible and reliable use of AI. Item 3 deals with AI compliance, commenting on the main local and international rules that dictate the use of the technology. Complementing this topic, in item 4 we discuss the relationship between Artificial Intelligence and data protection, which gives rise to the application of specific rules, including requirements of cybersecurity, and on the relationship between AI and Intellectual Property (IP). Following this, there are some recommendations and best practices to be followed by companies that already use or want to use AI in their processes. These encompass concrete measures that are intended to meet legal and ethical requirements and which reinforce the logic of risk mitigation. In item 6, we briefly present some success cases in the use of the technology that indicate its potential. This content is followed by a list of steps to be followed by companies: a brief summary focused on the practical and actionable steps that can be carried out by organizations. To conclude this Guide, we indicate some glossaries for consulting common concepts and expressions in the field of AI, in addition to the references used, which can also be consulted for further in-depth study.



1.

1. WHAT EXACTLY ARE WE TALKING ABOUT? UNDERSTANDING THE CONTOURS OF ARTIFICIAL INTELLIGENCE

Although it is possible to indicate more remote origins and significant groundwork related to Artificial Intelligence, it was mainly since the 1950s that the concept began to assume a format closer to how we conceive it today. There are two fundamental milestones that paved the way for the concept: the Turing Test and the conceptualization of John McCarthy.

Alan Turing, in an article published in 1950[3], described his experiment which became known as the *Turing Test or the imitation game*. In general terms, it refers to running some tests in which a human being and a computer should provide answers, and, in the sequence, another human being would analyze them and attempt to distinguish which were given by the machine. If the distinction is not possible, the machine is said to have passed the test.

The test was enthusiastically received by the scientific community, which began to explore the idea and experiments even further. A few years later, in 1956, a seminar lasting about two months was held at *Dartmouth College* in New Hampshire, in the United States, where scientists, who would become the most instrumental for AI, met to discuss various issues that were on the agenda at the time, such as the limits and possibilities of machines' capacity and their potential to perform tasks until then performed only by humans[4]. The pioneering use of the expression *Artificial Intelligence* was made during this seminar by John McCarthy.

Following these contributions, the topic became more present on the agenda of scientists, and debates, experiments and events were dedicated to testing and developing the capacity of machines until we reached the current state of evolution[5]. Even with countless studies and discussions, there is still no universal concept of Artificial Intelligence. However, it is possible to deduce the common idea between the different concepts, which allows us to clearly understand what we are talking about, when referring to AI.

In general terms, Artificial Intelligence is the term used to designate systems that have the capacity to learn and perform tasks previously only performed by humans. As a rule, this execution is preceded by the development and training of algorithms[6] carried out by a person, but there are already artificial intelligence systems that have the

[3] TURING, Alan. Computing Machinery and Intelligence. **Mind**, v. LIX, no. 236, pages 433-460, Oct. 1950.

[4] The main ideas guiding the organization of the event and which were in the spotlight of the authors' concerns are included in the invitation to the seminar, republished in 2006. Refer to MCCARTHY, John; MINSKY, Marvin L; ROCHESTER, Nathaniel; SHANNON, Claude E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. **AI Magazine**, volume. 27, no. 4, pages 12-14, 2006.

[5] In the present manual, it is out of context to present a detailed explanation on the modern background of Artificial Intelligence. For such purpose, refer, by all, to RUSSEL, Stuart J.; NORVIG, Peter. **Artificial Intelligence**: A Modern Approach. 4 ed. Hoboken: Pearson, 2021, page 5-30.

capacity to learn on their own as they carry out their activities. Models and algorithms have improved exponentially and performed increasingly complex tasks. There is an important AI feature that contributes to this development: the technology is based on the use of data – in general, a large volume of data, known as *big data*. These data feed the Artificial Intelligence, allowing for more accurate and informed decision-making. It is based on such data that different AI functions are performed, such as sorting data, generating forecasts, identifying and sorting images, making decisions, and many others[7]. These are things that, until the advent of AI, only humans have been able to perform.

Similarly, Luciano Floridi and Josh Cowls present a concept of AI that, despite having debatable aspects, allows a clear understanding of the theme and documents referring to technology:

> "The classic definition allows conceptualizing AI as an increasing resource of interactive agency, autonomous and often self-learning (...), which can handle tasks that would otherwise require human intelligence and intervention to be successfully performed. In short, AI is defined based on projected outcomes and actions, and, subsequently, we will treat AI as a reservoir of available intelligent agency."[8]

[6] The concept of algorithms also varies and its understanding has been analyzed by several scholars. An expression that has become widely known is that algorithms could be considered a black box, due to the difficulty in defining its functioning in a precise manner. On this topic, refer to PASQUALE, Frank. **The Black Box Society**: The Secret Algorithms That Control Money and Information. Cambridge (MA) and London: Harvard University Press, 2015.

[7] Similarly, Wolfgang Riem-Hoffman points out that "AI is a transversal technology aiming to train computers, by using big data, appropriate computing capacity and specific analysis and decision-making processes, to obtain accomplishments close to human capacity and which even exceed it, at least in some aspects". HOFFMANN-RIEM, Wolfgang. Inteligência artificial como oportunidade para a regulação jurídica. **RDU**, Porto Alegre, volume 16, no. 90, pages 11-12, Nov./Dec. 2019

[8] FLORIDI, Luciano; COWLS, Josh. A unified framework of five principles for AI in society. **Harvard Data Science Review**, no. 1.1, page 4, summer 2019.

The concepts of Nils Nilsson and Nick Bostrom also help with the more comprehensive understanding of what exactly is being referred to when talking about AI:

> "Artificial intelligence is that activity dedicated to making machines intelligent and intelligence is that quality that allows a certain entity to function properly and with viewing in its environment"[9]

> "Machines that are equal to humans in general intelligence – that it, having common sense and effective capacity to learn, reason and plan to face complex challenges on processing information in a wide range of natural and abstract domains"[10]

Due to the fast-paced technical evolution that AI is providing and will provide in the near future, it has even been said that its potential will surpass human intelligence. However, as we have already mentioned, there is an understanding that it is unfeasible to talk of replacing one intelligence with another. It is, in fact, a relationship of connection and complementarity. By performing even simple tasks, AI is already contributing to allow the human brain to focus on more difficult and complex issues that machines cannot yet perform, and even to develop machines with ever-greater capabilities. Thus, Artificial Intelligence will even enhance human capabilities and perform tasks that the human brain alone cannot – or that can carry out at a much lower speed. This may seem difficult to grasp, but we can think, for example, of a comparison between the human memory capacity, which is limited and tends to degrade over time, in relation to the storage capacity of a computer, whose data, if treated properly, will not degrade and storage is virtually unlimited. Similarly, AI is capable of processing a greater amount of information at an even greater speed. However, it is worth emphasizing once again that these feats are only possible, at least currently, as a result of human collaboration, which reinforces their complementary nature.

Today, AI shows a great deal of potential in specific applications that are already widely used in our daily lives. This Artificial Intelligence that we currently use is referred to as Weak or Narrow AI, due to its limited ability to

[9] NILSSON, Nils J. **The Quest for Artificial Intelligence**: A History of Ideas and Achievements. Cambridge: Cambridge University Press, 2010, page 13.

[10] BOSTROM, Nick. **Superintelligence**: Paths, Dangers, Strategies. Oxford: Oxford University Press, 2014, page 3.

perform functions other than those for which it was trained. There is, however, an expectation regarding what has been categorized as Strong or General AI, but we still do not know how or when this will become a reality.

"*Weak Artificial Intelligence (Weak AI or Narrow AI)* is a specific intelligence, focused on an activity or area and that performs only that for which it was taught. It is the classic example of the system programmed to play chess, which plays only chess, but does it with excellence. It is the most used form today and is more present in our daily lives. The fact that it is called weak, however, does not diminish its unprecedented utility to perform complex tasks, and to learn from the data and contexts it is given. This is the case, for example, of recommendation algorithms, chatbots and even autonomous cars.

In turn, *Strong Artificial Intelligence* or *General Artificial Intelligence*[11] is not limited to a specific issue, given that it possesses a more general and comprehensive intelligence, adapting to different contexts and situations. Due to its complexity, it has yet to be developed and will require greater attention due to the risks it can generate."

Below, we present an outline of the main learning techniques involved in training Artificial Intelligence algorithms, in order to understand how Machine Learning works.

[11] An in-depth discussion on the subject can be found in RUSSEL, Stuart J.; NORVIG, Peter. **Artificial Intelligence**: A Modern Approach. 4 ed. Hoboken: Pearson, 2021, pages 31 et seq. and 981 et seq.

Machine Learning

With this method, the algorithm is trained with a large amount of data in order to answer questions and solve problems. Among the most used algorithms in machine learning are “decision trees”, “K-means” and “deep learning” to solve tasks that can vary between “supervised learning”, “unsupervised learning” and “learning by reinforcement”. What makes Deep Learning attractive is its ability to solve all three tasks, as explained below.

Deep Learning[12]

This is a more sophisticated form of machine learning, which learns and performs more complex tasks. Based on neural networks, which are inspired by the functioning of the human brain, they use a large number of artificial neurons connected in multiple layers, hence the reference to their “deep” nature.

Its configuration can be carried out mainly through 3 different machine learning tasks:

1. Supervised learning

The most utilized technique today. It uses data labeled by humans for training and for testing the results to verify that the algorithm generates responses with the expected accuracy. With this technique it is possible, for example, to make forecasts, such as sales in a given scenario; analyze probabilities; sort images and other data.

2. Unsupervised Learning

Uses unlabeled data to explore possible results, seeking information that, a *priori*, is not known. Through, for example, grouping (according to client profile, behavior, etc.) and data associations, it generates useful results for the company, such as identifying patterns and problems.

3. Learning by reinforcement

Uses positive and negative reinforcements and uses its own mistakes and successes throughout the process to learn how to make a decision/choose the best path to follow. It can be used, for example, to recommend content and advertising: depending on whether a client consumes the recommended content and the products indicated, the algorithm learns more about the client’s preferences.

[12] HOFFMANN-RIEM, Wolfgang. Op. cit., page 12

Natural Language Processing (NLP): the area that is dedicated to linguistic processing, recognizing meanings in texts and speech, allowing communication with humans. This technique is widely used in Chatbots, translators and virtual assistants like Siri, Alexa, Google Home.

Computer vision: the name itself denotes its concept, which is to give vision to a machine so that it becomes able to see certain things. It is used to recognize images and videos and extract data, patterns and the like from them. A practical and widely used example is facial recognition.

Robotic process automation (RPA): as the name suggests, it refers to the automation of certain tasks and processes, which can involve different files and even systems. For example: filling in documents, reports and spreadsheet data; copying data between and among forms; detecting payment for subsequent issuance of invoice, followed by shipment of the product; and organizing registrations. It is useful for mechanical processes, which traditionally are done manually and in which there is practically no variation. It is an attractive alternative for starting the application of AI in a company, as it does not require major structural changes or an extremely advanced IT infrastructure.

Chatbots: the algorithms are programmed to, from their database, allow the machine to interact and converse with humans, either through voice or text. It is one of the easiest applications to understand and it can be used for customer service, FAQ and even to schedule services. It can be used in combination with other AI techniques, enhancing its effect. Along with NLP, for example, the chatbot becomes able to understand more clearly what the user means. With machine learning, the chatbot learns from conversations, improving its responses.

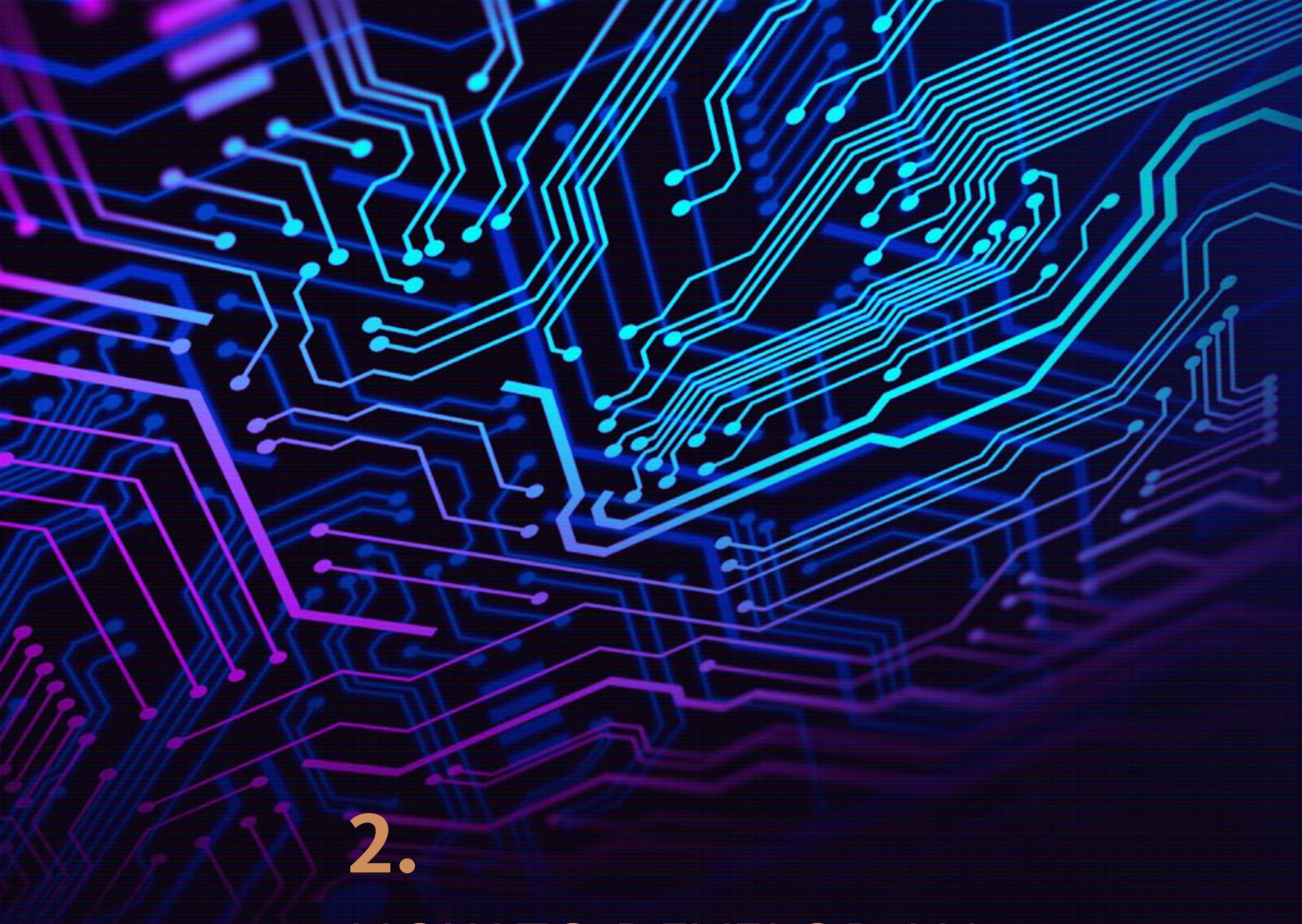
Recommendation algorithm: an algorithm that makes recommendations for content, products and services based on users' profiles. The algorithm needs to be fed with data, which will be categorized to identify patterns with cutouts, for example, of gender, age and social class. Based on this, it will analyze, depending on the case, consumption habits, consumed content and the like to make new recommendations. To keep up to date, the algorithm requires considerable data storage and processing capacity. This is what we see, for example, in movie, series and music streaming apps, as well as social networks recommending new content and stores that potentially will be of interest to the user. The use of this algorithm by a company allows a more precise reach of its target audience, but it must be used with caution – both during its preparation and feeding with data and when presenting recommendations to clients, as insistent contact, poorly directed or without clear consent in data collection can generate discomfort and alienate the client. The algorithm can be enhanced with the crossing of AI technologies, especially machine learning.

All of these techniques provide, as already mentioned, countless applications and bring several advantages to companies that use them, such as efficiency, due to: the low margin of error of the technologies; reduction of labor, especially in mechanical tasks, which can be done by a machine; standardization of processes and communication with clients; improvement of products and services; and through attracting new clients and more qualified

professionals. Among its applications, we mention only for illustrative purposes: the referred chatbots; the use of algorithms to recruit professionals and select candidates with a profile closest to that sought; employee and company performance analysis; recommendation of content, products and services to users and clients; value forecasting and market analysis; credit analysis; traffic management and indication of alternative routes; decision-making; applications in smart homes; assistance in medical diagnosis; image recognition and biometrics; verification of identity and veracity of documents; autonomous vehicles; prevention of process failures.

Note that these applications may demand and/or benefit from the use of other technologies that are not to be confused with AI, but which mutually enhance the potential. This is the case, for example, of drones, the Internet of Things and wearable devices. These are technologies that inevitably demand new and more complex processes, moreover because they considerably increase data production and touch on new privacy issues.

In any case, all Artificial Intelligence requires specific due care and must follow rules that are already circulated locally and internationally. As the benefits of AI grow, the associated risks also accumulate. Many consequences of such risks can be anticipated and avoided in advance, but those that are unforeseen still exist, and action must be taken to avoid and minimize the damage. Thus, the use of AI should be made within a specific ethical framework and legal regulations, points that will be explored in the topics further below.



2.

HOW TO DEVELOP AN ARTIFICIAL INTELLIGENCE BASED IN ETHICAL PRINCIPLES?

The various applications of Artificial Intelligence bring unprecedented ethical challenges to human beings. This is amplified by the fact that the technology is relatively recent and is evolving at a rapid pace, including its high potential to change the way we live, impacting all areas of human life. Some of its ethical risks involve, for example, the proliferation of social stigma, the perpetration of fraud and disrespect for individuals' rights. Consider, for example, the possibility of attributing to the machine the decision-making process that would have social and collective impacts, as would be the case with its use in public institutions. There are even consequences of current applications that are not yet foreseen, mainly due to the difficulty of understanding and controlling an algorithm in its entirety, as well as the existence of applications that we are not yet aware of, but which already raise moral concerns.

As highlighted by Nick Bostrom and Eliezer Yudkowsky[13], the ethical issues of Artificial Intelligence, especially strong AI, differ from those existing in relation to other non-cognitive technologies, because: 1) even if its configuration and training are done properly, the result may not be fully predictable; 2) checking the system security is a challenge, as it must be done in all possible contexts in which AI can be applied; and 3) the ethical concern must also be a concern of the engineering area.

There are already debates about the moral status of Artificial Intelligence itself. The cognitive capacity that the system may have (which allows autonomous decision-making, learning and development independent of human intervention), raises the question of whether these machines should be considered themselves as moral beings[14]. In addition, ethical concerns about Artificial Intelligence also focus on the fact that its various applications can be used for the benefit of society, companies and citizens, but can also be used for harm.

For these reasons, scholars on the theme, companies, international organizations and other important players in the field have indicated the need to observe ethical standards in the use of Artificial Intelligence. Today, some comprehensive ethical standards have already been developed and disseminated to create general limits for the use of AI. These are principles that have developed over the years and have contributed to delineating the limits and possibilities of applying AI and self-regulation, which are common in this field. Some companies, especially technology giants and global entities, already have solid Artificial

[13] BOSTROM, Nick; YUDKOWSKY, Eliezer. The ethics of artificial intelligence. In: FRANKISH, Keith; RAMSEY, William M. (eds.). **The Cambridge Handbook of Artificial Intelligence.** Cambridge (UK): Cambridge University Press, 2014, page 320.

[14] Refer to, for all, BOSTROM; YUDKOWSKY. Op. cit., page 321 et seq.; VERBEEK, P. **Moralizing Technology**: Understanding and Designing the Morality of Things. Chicago and London: The University of Chicago Press, 2011; MAGRANI, Eduardo. New perspectives on ethics and the laws of artificial intelligence. **Internet Policy Review,** volume 8, no. 3, pages 1-19, Sept. 2019.

Intelligence policies, which include a strong reference to principles. This is the case, for example, with Microsoft[15], Google[16] and the Organization for Economic Cooperation and Development (OCDE)[17].

Thus, there are principles and good ethical practices that are disseminated and accepted by the international community. These principles serve as the basis for the development of company policies, laws and guidelines from international organizations and for legislative discussion. In these policies and in laws already published on the subject matter, it is possible to notice the strong character of principles. There is actually a proliferation of principles, and each author and company treats them in a certain way. However, more in-depth analysis of these various principles allows us to identify that the multiple concepts all contain similar ideas. The concern, in general, is to make Artificial Intelligence responsible, transparent and, above all, trustworthy. Given the linearity of the logic underlying them, it is possible to group them into five general principles that should guide the use of Artificial Intelligence[18], which are presented below. It should be noted at the outset that behind each of these principles are the notions of promoting human values and mitigating harm, putting the human at the center of the technology's applications.

Before analyzing each of these ethical principles, it is important to highlight that they should be observed in all phases of AI use, from the formulation of ideas and project design, through systems architecture, programming and engineering, to its practical application and possible subsequent consequences, which leads us to the notion of ethics by design[19]. This requires, among other things, a multidisciplinary and diverse team, as will be discussed later, and whose training covers ethical issues, including for engineers, programmers and the like.

Finally, we add that the use of AI and the respect for principles should not be based on a passive stance of avoiding risks and biases. In other words, it is not enough just to program without biases or to believe that the security of the system itself guarantees that risks and damages will not occur[20].

Below, we will briefly explain what each of the principles consists of.

[15] The company emphasizes performing responsible use of AI and, for that, it bases itself on the following principles: fairness, inclusiveness, reliability & safety, transparency, privacy and security, accountability. Complete approach available at: <https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1%3aprimaryr6>. Accessed on June 13, 2021.

[16] Highlighting the belief in AI's potential to "empower people, broadly benefit current and future generations and work for the common good", the company presents 7 principles/objectives, namely: being socially beneficial, avoiding creating or reinforcing unfair prejudices, be built and tested for safety, be accountable for people to incorporate principles of privacy by design, maintain high standards of scientific excellence and be made available the for uses that are consistent with these principles. The complete approach is available at: <https://ai.google/principles/>. Accessed on June 13, 2021.

[17] We will address the organization's principles in item 6 of this Guide. The OECD's principled approach is available at: <https://www.oecd.org/going-digital/ai/principles/>. Accessed on June 13, 2021.

[18] The grouping presented in this Guide is based on the study by Luciano Floridi and Josh Cowls that identifies the core of the principles in synthesis, through a broad comparative analysis of the main ethical frameworks. Cf. FLORIDI; COWLS. Op. cit. In another comprehensive comparative analysis of the current AI and Ethics landscape, Anna Jobin et. al. indicate that these guiding principles of AI development should include: transparency, justice, non-maleficence, responsibility, privacy, beneficence, autonomy, trust, dignity, sustainability and solidarity. Jobin, A., et al. The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. 2019.

2.1. Beneficence

The principle of beneficence indicates, in general terms, that Artificial Intelligence must be used in a beneficial way for humanity. Thus, it must be focused on the well-being of human beings, empowering as many people as possible, respecting human dignity and also being attentive to the sustainability of the planet[21]. In summary, it can be translated into the idea of doing only good.

2.2. Non-maleficence

The principle of non-maleficence can be translated, in few words, into the idea that the use of AI should not cause harm. Thus, its misuse should be avoided, as would be the case of using it in an arms race. The development and application of AI technology should only occur within the limits of safety, and should prevent not only the violation of privacy, but also other fundamental rights. This principle also involves the assumption of the responsibility for working against possible technology risks[22].

2.3. Autonomy

The principle of autonomy seeks to emphasize that the Artificial Intelligence should be used in a way that it does not weaken or annihilate human autonomy, but that should, on the contrary, promote it. When using technology, we give the machine a certain power of decision, but this needs to be done in a balanced way, that is, balancing the power delegated to the machine and the power maintained by us. Furthermore, the power to delegate decisions and to withdraw that delegation must always be held by humans. Otherwise, there is a risk that AI autonomy will undermine human autonomy. Based on these ideas, Floridi and Cowls refer to a "target-autonomy" or "decide-delegate" model[23].

[19] Refer to, for all, KAUFMAN, Dora. Inteligência Artificial e os desafios éticos: a restrita aplicabilidade dos princípios gerais para nortear o ecossistema de IA. **PAULUS: Revista de Comunicação da FAPCOM**, São Paulo, v. 5, n. 9, p. 73-84, Jan./Jul. 2021.

[20] MAGRANI, Eduardo. Inteligência Artificial responsável. O que é isto? **Estadão**, March 13, 2021. Available at: <https://politica.estadao.com.br/blogs/fausto-macedo/inteligencia-artificial-responsavel-o-que-e-isto/>. Accessed on: June 03, 2021.

[21] FLORIDI; COWLS. Op. cit., page 6.

[22] FFLORIDI; COWLS. Op. cit., page 6.

[23] FLORIDI; COWLS. Op. cit., page 7: "human beings must retain the power to decide which decisions to make: exercise the freedom to choose when necessary and give it up in cases where compelling reasons, such as effectiveness, can overcome the loss of control over the decision-making. Any delegation must also remain void in principle (i.e., decision to decide again)."

2.4. Justice

The principle of justice is based on the idea that there is inequality of autonomy in society and, therefore, AI should be used to promote justice. It is a principle broadly addressed in different documents and varies according to the authors, but, in general, it indicates the need to eliminate any type of discrimination in the use and results of AI, in addition to promoting aspects such as prosperity, solidarity, justice, equity, diversity and equal access to benefits in society[24].

One of the consequences of the principle of justice, especially in terms of non-discrimination and absence of prejudice, is the non-existence of biases throughout the processes and, above all, in the results. Biases in AI can be defined as a biased and disproportionate analysis that disadvantages a person or a group. Take as an example the AI credit analysis, which, in some cases, has generated discriminatory results: credit denied for reasons of race and geographic location, to the extent that black people and people living in poorer neighborhoods had their requests denied, even if, initially, the AI was not intentionally configured to make a decision based on these factors.

As observed by Dora Kaufman[25], these errors may occur: before data collection, due to programmers' decisions; in the data itself, which occurs in the event that the data are not representative and proportional to the portion of the population to which it refers; still in the data itself, but reproducing prejudices that exist in society; and also due to labeling errors (prior to supervised learning) or in data generation.

In this way, the importance of the existence of diversity in the data and of adequate categorization and selection becomes clear. Otherwise, the decisions of the algorithm will reproduce the human biases already present, usually unconsciously, in the implementation and feeding of the technology.

As we have already highlighted above, overcoming biases is not possible if a certain factor is ignored: a positive action is necessary for the algorithm to understand that some factors should not be regarded in a negative sense in the decision-making. Once the AI system has been fed with data, it might be difficult to identify the source of the bias. Algorithm auditing, which will be covered in this Guide, can be fundamental in this process of recognition and overcoming of bias.

[24] FLORIDI; COWLS. Op. cit., page 7-8.

[25] KAUFMAN, Dora. Op. cit., page 79.

2.5. Explicability

The principle of explicability complements and reinforces the four principles observed so far[26] and it can be divided into two main ideas: intelligibility and accountability, as explained by Floridi and Cowls. Intelligibility is concerned with how the AI system works. Thus, it is necessary to have transparency and the possibility of understanding, interpreting and explaining how the decision-making process works. Accountability is concerned with who is responsible for the way the system works.

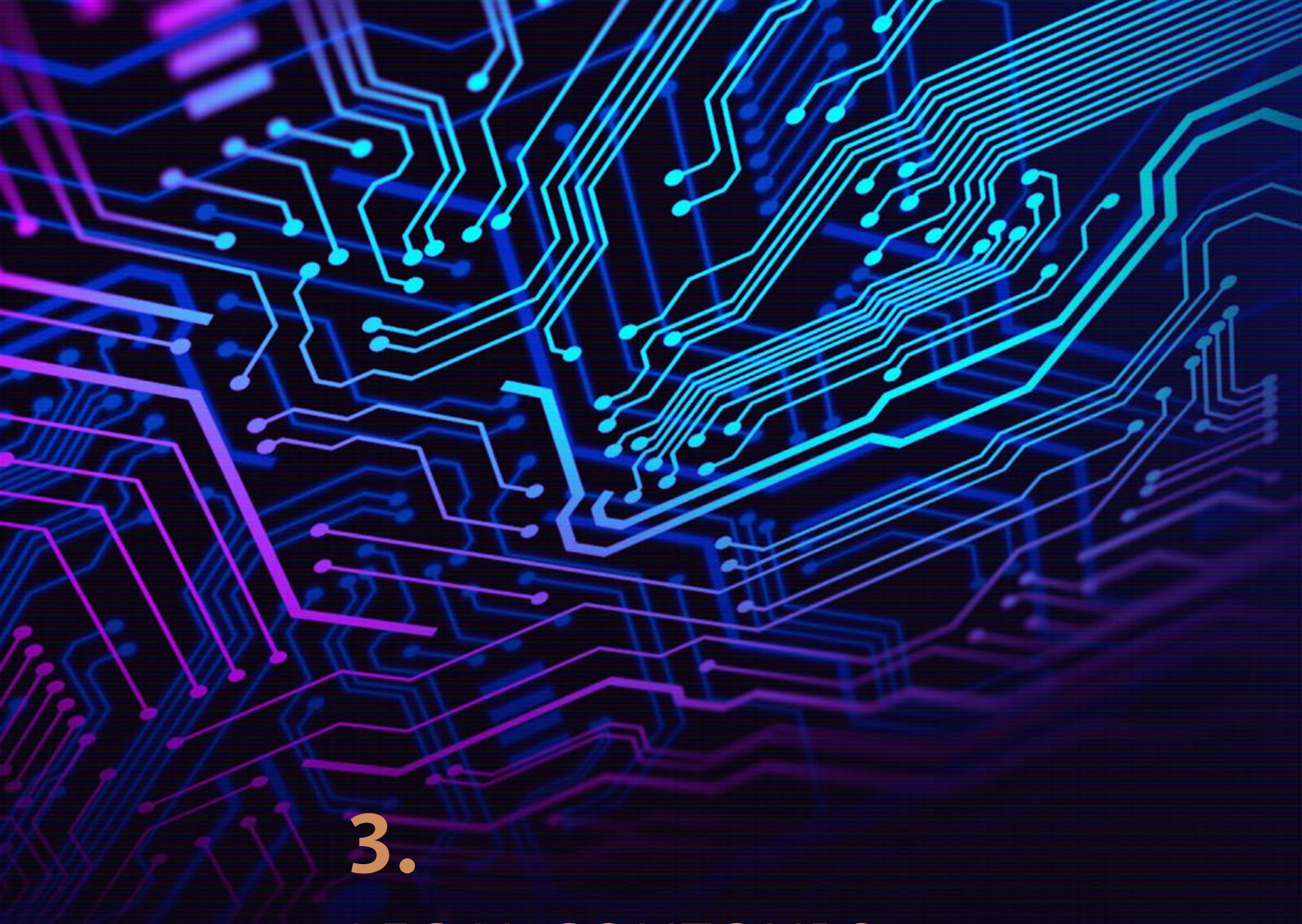
This principle of explicability is strongly related to the attempt that has intensified in recent years to reduce the black box aspect of algorithms, which is connected to the "lack of understanding of how the so-called data inputs generates the output data, how the system correlates the variables contained in the input data and the assigned weights (called 'parameters')"[27]. Thus, as Dora Kaufman observes, the purpose is to understand and transmit in a clear and accessible manner to the user how the system reached a certain result. It happens that accuracy and transparency are inversely proportional factors in AI systems, which still represents a challenge for programmers. This becomes even more real, according to the complexity of systems: processes based on deep learning, for example, become increasingly harder to explain.

In short, explicability, especially in terms of transparency, requires clear communication, through an accessible language to the consumer, and it is also recommended to use specific tools whenever possible[28]. Project traceability is also required, in order to be able to define what was done at each stage and throughout the entire process to explain the system and its decision-making to the user.

[26] FLORIDI, Luciano; COWLS, Josh. Op. cit., page 8: " for AI to be beneficial and non-maleficent, we must be able to understand the good or evil it is actually causing society and in what ways; for AI to promote rather than restrict human autonomy, our 'decision about who should decide' must be informed by knowledge of how AI would act on our behalf; and for AI to be fair, we must know who to hold responsible in the event of a serious negative outcome, which would in turn require a proper understanding of why that result arose."

[27] KAUFMAN, Dora. Op. cit., page 79.

[28] ALLIANZ. **Practical Guidance for AI.** [s.l.]: Allianz, [s.d.], page 14.



3.

LEGAL CONTOURS OF USING ARTIFICIAL INTELLIGENCE: AI COMPLIANCE

As already observed, it is increasingly clear that AI will significantly change the way we live. The forecasts point to great impacts in all areas of life, with emphasis on the social, cultural and economic aspects. Technology can impact lifestyle, private choices and political decisions[29]. Thus, ethical principles that delimit both the use of Artificial Intelligence and the legal aspects become even more important. In Europe, for example, a trend towards the regulation of technology has been observed and, as happened with the LGPD, which was inspired by the European Union's General Data Protection Regulation (GDPR), we believe that the treatment of AI in Brazil will follow a similar path.

It was mainly from the 2000s onwards that saw stronger investment in this technology[30], causing it to develop at a faster pace. Over time, the potential of AI and its commercialization became increasingly real, but the discussion on regulation of AI only started to advance relatively recently.

However, regulating technology is not an easy task, especially one such as Artificial Intelligence, which is evolving at a rapid pace, resulting in an equally rapid obsolescence and insufficiency of laws, and that will have even more profound impacts on society and the economy. While seeking to protect rights and values, it is necessary to leave room for technological development. However, it is unfeasible to presuppose a total dichotomy between innovation and legal treatment or the idea that the legislation would create barriers to economic progress. It is necessary and it is possible to find a balance. Many times, the adoption of similar principles and rules – whether from a state regulation or a self-regulation of industries and companies – enhances the attraction of investments, as it ensures a suitable environment for business development, protects Intellectual Property and individual rights.

In light of this, what has in fact been observed is a greater rate of the proliferation of principles, the self-regulation of large technology companies and the establishment of policies by international entities, aspects addressed in the preceding item. We can also observe an increase of debates, proposal and discussion of bills, public forums, initiatives of working groups and research laboratories and other measures aimed at developing an efficient state regulation. The debate is more advanced, especially in developed countries,

[29] On this subject, refer, for example, to LATZER, Michael et. Al. The Economics of Algorithmic Selection of the Internet. **Working Paper – Media Change & Innovation Division**, Oct. 2014. Available at: <https://mediachange.ch/media//pdf/publications/Economics_of_algorithmic_selection_WP.pdf>. Accessed on: July, 07, 2021.

[30] PARENTONI, Leonardo Netto; VALENTINI, Rômulo Soares; ALVES, Tárik César Oliveira e. Panorama da regulação da inteligência artificial no Brasil: com ênfase no PLS n. 5.051/2019. **Revista Eletrônica do Curso de Direito da UFSM**, Santa Maria, volume 15, no. 2, pages 1-29, May/Aug. 2020.

but there are still countries that find themselves in a scenario of insufficient legislation on the subject. This is the case of Brazil, which does not yet have a specific law addressing the use of Artificial Intelligence, but the discussion has been already been initiated and is expected still to expand, within the scope of discussions in society, the legislative and executive branches, the public sphere and among the entities and authorities interested in the discussion.

Note that the discussion here is carried out considering *specific* rules on Artificial Intelligence and, in some cases, more comprehensive rules related to digital law. Nonetheless, this does not exclude the need to observe other rules also applicable, such as those of Intellectual Property, competition law, civil liability, consumer law, among other local and international legal rules. Due to the innumerous possibilities of using Artificial Intelligence, many other rules are mandatory depending on the case and the area of operation, as seen in the insurance, medical, pharmaceutical and financial areas.

Therefore, we present below a general overview of the Brazilian and international scenarios for the regulation of Artificial Intelligence. As we shall see, a common point of the rules is to seek to ensure enforcement of ethical principles, guarantee respect for the rights of users, delimit the possibilities of using the technology and minimize risks. Note, however, that although some countries do not have laws on the subject and others have more general laws, this does not prevent the use of AI.

On the one hand, the law itself establishes general principles and rules that already present the direction and delimitation of what must be done and what is prohibited in the use of the technology. This is the case, for example, of the constitutional principles, which prohibit discrimination and ensure the protection of honor and image, and also other laws that address the issue, such as the LGPD and the Brazilian Internet Civil Framework. On the other hand, although the trend has been to regulate AI, this does not prevent companies from acting proactively to create their own rules and good practices to align their operations with the best standards of the market and compliance.

After all, regulation, especially of a technology, does not operate solely in a top-down movement, in which the State creates laws to be followed by companies. Companies themselves can be proactive in this context, creating

guidelines and rules for their self-regulation. Thus, respect for legal rules should take place through the design, codes and architectures, so that the law acts increasingly as a meta-regulation, which "creates a normative-evaluative framework" within which technology can operate[31]. This is the importance, among other things, of the aforementioned ethics by design, as well as privacy by design and security by design:ethics by design, as well as privacy by design and security by design:

> "We no longer run the risk of technology completely overriding law, but neither do we ignore its power of influence. The decision-making role of technology is accepted, but through the design, a space for legal regulation and of ethical-normative bases within which it can operate is guaranteed."[32]

3.1. The Brazilian scenario

As mentioned, we still do not have in Brazil a specific law addressing Artificial Intelligence and even the bills of law on the subject are recent, proposed from 2019 onwards[33]. Among these, Bill PL 21/2020, proposed in the House of Representatives, has gained some prominence in the national debate. The Bill "establishes foundations, principles and guidelines for the development and application of Artificial Intelligence in Brazil". Throughout its articles, the Bill deals, among other things, with conceptualizations of important issues for Artificial Intelligence; introduces fundamentals for the use of AI, listing aspects that should be promoted with such use; presents the principles to be observed for responsible use of AI, which approach the aspects mentioned in item 2 of this Guide and also principles already established in the area of data protection. The other bills already proposed so far follow similar paths, despite having important differences.

Mateus Fornasier and Norberto Knebel[34] note that such bills corroborate the trend towards social auditability of algorithms (accountability), indicating that algorithms must be justified in light of open and verifiable principles. The authors also state that the bills, in a way, reinforce the trend of self-regulation by the companies, as was seen with the LGPD, which, in its Article 50, provides

[31] MAGRANI, Eduardo; OLIVEIRA, Renan Medeiros de. Desafios regulatórios da Internet das Coisas: reflexões sobre a LGPD, o consentimento e o direito à explicação. In: CRESPO, M. X. F. (Coord.). **Compliance no direito digital**. São Paulo: Thomson Reuters Brasil, 2020, page 80. This point has been further explored in MAGRANI, Eduardo. **Entre dados e robôs**: ética e privacidade na era da hiperconectividade. Porto Alegre: Arquipélago Editorial, 2019, pages 247 et seq.

[32] MAGRANI; OLIVEIRA. Op. cit., page 80-81.

[33] In the Federal Senate, we can list the following Bills: PL No. 5051/2019, PL No. 5.691/2019, PL No. 872/2021. In the House of Representatives, we can point out PL No. 21/2020, PL No. 1969/2021 (attached to the first) and PL No. 240/2020.

[34] FORNASIER, Mateus de Oliveira; KNEBEL, Norberto. Inteligência artificial: desafios e riscos ético-jurídicos. **Direito & Paz**, São Paulo, v. XIV, n. 43, p. 221, jul.-dez. 2020.

that controllers and operators "may formulate rules of good practices and governance" in line with the principles and rules of law. Thus, in the context of AI regulation, "developers and operators are inserted in this same context of regulatory entity, at the internal level, both upon creating the systems and in their activities, enshrines the reality of regulation by design, that is, that the developers themselves are also responsible for the preservation of public order principles"[35].

Legislative activity in the country generated a reaction from the Executive Branch, which also adopted some initiatives with the aim of regulating Artificial Intelligence. This was the case, for example, of the public consultation of the Ministry of Science, Technology and Innovation (MCTIC) on the Brazilian strategy for Artificial Intelligence[36] and the call to fund research centers for allocation in AI[37]. In 2021, the Brazilian Strategy for Artificial Intelligence (EBIA) (Ordinance GM No. 4617/2021, modified by Ordinance MCT No. 4979/2021) was published. However, in addition to the restricted scope of the EBIA, which provides guidance only to the Federal Government, its laconic forecasts do not resolve several doubts and issues that exist in the country. Although it is an important measure and presents concerns such as the elaboration of ethical principles and standards, the promotion of investments in research and development in the field and provides, for example, that AI should benefit people and the planet, the Strategy was received with criticism by AI scholars. In this regard, it should be observed that the document arrived late, presents a level of maturity below expectations, especially in comparison with the plans of other countries, and attributes few roles to the government itself in the sense of innovating in AI[38]. The document also lacks concreteness, as it does not contain clear responsibilities or budgetary considerations.

The legal vacuum and legislative uncertainties, however, are not reasons enough to prevent the use of the technology in the country. Companies and organizations often create their own policies delimiting the parameters of AI (self-regulation), which is often based on internationally adopted rules and practices. In this context, a good practice is usually the adoption of stricter rules and more demanding parameters to avoid potential problems.

[35] Ibidem, p. 222.

[36] Refer to <http://participa.br/profile/estrategia-brasileira-de-inteligencia-artificial>. Accessed on: July 15, 2021.

[37] For a brief history on AI regulation in Brazil, and analysis of PLs, refer to PARENTONI; VALENTINI; ALVES. Op. cit., page 1-29.

[38] PAIVA, Fernando. A EBIA é uma estratégia sem estratégia, diz Eduardo Magrani. **Mobiletime**, April 15, 2021. Available at: <https://www.mobiletime.com.br/noticias/15/04/2021/a-ebia-e-uma-estrategia-sem-estrategia-critica-eduardo-magrani/>. Accessed on: July 22, 2021. In a similar sense, refer to NUNES, Ana Carolina. Estratégia Brasileira para IA é um passo importante, mas já chega precisando de atualização. **Época Negócios**, May 28, 2021. Available at: <https://epocanegocios.globo.com/Tecnologia/noticia/2021/05/estrategia-brasileira-para-ia-e-um-passo-importante-mas-ja-chega-precisando-de-atualizacao.html>. Accessed on: July 22, 2021 and LEMOS, Ronaldo. Estratégia de IA brasileira é patética. **Folha de S. Paulo**, April 11, 2021. Available at: <https://www1.folha.uol.com.br/colunas/ronaldolemos/2021/04/estrategia-de-ia-brasileira-e-patetica.shtml>. Accessed on: July 22, 2021.

A case that exemplifies the use and self-regulation of the technology is that of the Brazilian Judiciary Branch. A survey carried out by the Center for Innovation, Administration and Research of the Judiciary Branch (CIAPJ) of the Fundação Getulio Vargas (FGV) indicates that half of the courts in the country have an AI project under development or even implemented[39], which is done seeking, among other things, greater efficiency, productivity and procedural speed. Thus, there are tools designed, for example, to categorize and search for case law, transcribe voice to text, identify whether a lawsuit is linked to any precedent, and semi-automatically generate procedural instruments. Aware of this movement, the National Council of Justice (CNJ) issued Resolution No. 332/2020, which provides for ethics, transparency and governance in the production and use of Artificial Intelligence in the Judiciary Branch, and Ordinance No. 271/2020, which regulates the use of Artificial Intelligence within the scope of the Judiciary Branch.

Besides the specific regulation of AI – non-existent in Brazil –, there are some other laws applicable, directly and indirectly, to those who use the technology, with a focus on the LGPD and Internet Civil Framework. Because of their importance to AI, we will address these laws in a separate item.

3.2. The International scenario

Internationally, the regulatory debate on Artificial Intelligence is in a more advanced stage. It is possible to identify the issuance of more robust national plans as well as specific laws on AI, which address issues such as accountability and robotics. Even so, regulation is fragmented and there are also countries at a less advanced stage, for which it is necessary to outline more clearly the implications of AI, the limits and possibilities of action in society and in the economy, among other factors. In any case, the analysis of existing standards allows us to identify certain constants, such as the importance given to the principles that guide the use of an ethical, responsible and reliable Artificial Intelligence; the focus on humans and on protecting the planet; the need for algorithms to be explainable and transparent; technical and security robustness; human autonomy; and diversity, justice and equality. The debate on accountability in the field of AI has also made advances.

[39] SALOMÃO, Luís Felipe (Coord.). Inteligência Artificial: tecnologia aplicada à gestão dos conflitos no âmbito do poder judiciário brasileiro. [s.l.]: FGV Conhecimento, 2020.

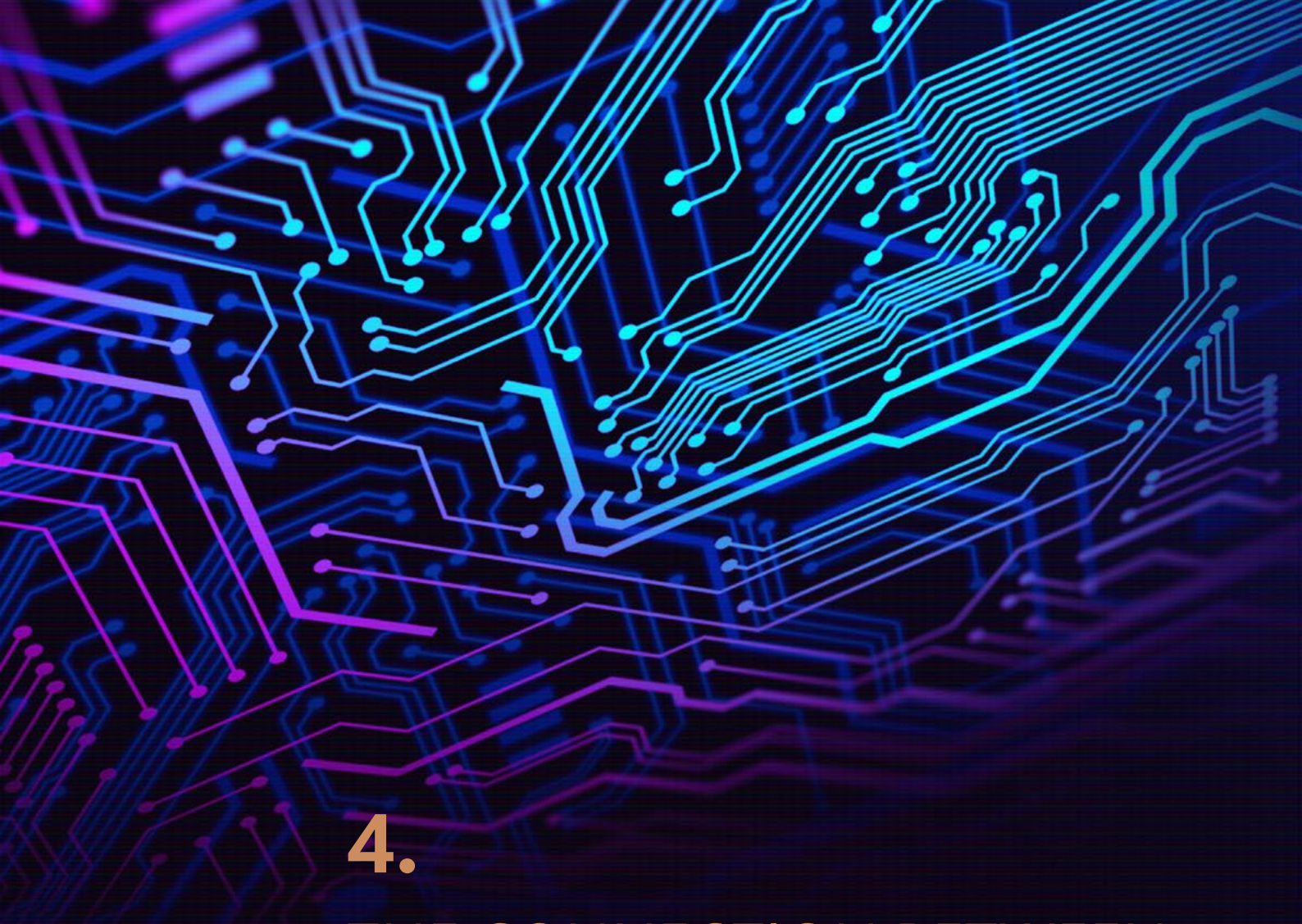
Recently, in 2021, Europe released a proposal for a general regulation on AI with a risk-based approach. This is the Artificial Intelligence Act (AI Act), which sets out specific rules for the various systems and applications of the technology. The proposal is the outcome of an extensive process of debate and publication of guidelines by the European Commission. Previously, in April 2019, the Commission's Independent High-Level Expert Group on Artificial Intelligence published the "Ethical Guidelines for Reliable Artificial Intelligence". The document recognizes the potential of AI, but also pays attention to the risks, seeking to address how to deal with certain problems. Emphasizing that a reliable AI must be legal, ethical and robust, the document addresses issues such as accountability, human autonomy and systems oversight, as well as technical robustness. It also provides principles similar to those already covered in this Guide. The following year, the Commission published the White Paper *"On Artificial Intelligence – A European approach to excellence and trust"*, giving continuance to the guidelines to be followed and pointing out future regulatory actions. Other measures, such as public consultations, were adopted until reaching the aforementioned regulation proposal of 2021.

In the United States, the Algorithmic Accountability Act, of 2019, defines which systems that involve personal data or make automated decisions are considered to be of high risk, presenting requisites to be followed in these cases. The Commercial Facial Recognition Privacy Act, also of 2019, prohibits the processing of facial recognition data, except in the event that the organization presents documentation on the capabilities and limitations of its technology and there is express affirmative consent from the end user preceded by notification about the reasonable uses for this type of data. In turn, the Self Drive Act, of 2017, establishes rules for testing and implementing automated cars, in addition to other provisions on the subject. Local governments have also acted to regulate AI. The city of New York, for example, has approved the New York City Automated Decision Systems Task Force – ADS Task Force, aimed at government agencies that use algorithms in decision-making processes. In general, the purpose is to assess whether the

algorithms are consistent with the purposes of making the city more just and equitable. The State of Washington is discussing a bill aimed at eliminating biases in automated decision-making processes.

In the context of adopting national plans, China should be highlighted, which has an ambitious AI strategic development plan. Seeking to become the world's leading country in implementation of AI technology by 2030, the New Generation Artificial Intelligence Plan (AIDP) was launched in 2017 and is focused on the areas of international competition, economic growth and social governance[40]. Among other countries that also have plans, we can mention Australia, with the Artificial Intelligence Ethics Framework, of 2019, and Germany, with The German Strategy for Artificial Intelligence, of 2018.

[40] On the Chinese approach to AI, refer, for all, ROBERTS, H. et al. The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. **AI & Soc**, volume 36, pages 59-77, 2021.



4.

THE CONNECTION BETWEEN ARTIFICIAL INTELLIGENCE AND OTHER AREAS OF LAW: DATA PROTECTION, CYBERSECURITY AND INTELLECTUAL PROPERTY

Artificial Intelligence is closely connected to other technologies, besides being based on large-volume data usage. One of the initial steps in programming and defining how an AI tool works is to determine what data it will rely on to work and generate the expected results. This is because AI uses a large amount of data – personal and even sensitive data[41], oftentimes – and generates new data. This is not a simple task, but it is essential for the proper performance of the technology and, accordingly, for a company's own performance, since the lack of definition or disorganization of such data can produce negative outcomes, either by delaying processes or generating undesired legal effects, or even increasing the risks triggered by the systems themselves. Therefore, the use of AI requires compliance with other legal rules, such as those related to data protection and cybersecurity. Such compliance is paramount and is related to the idea of *resilience or robustness* of the AI, that is, companies must have "technical robustness and compliance of their AI and its agility in all platforms and resistance against malicious agents"[42].

In addition, due to the creative, copyright and also industrial nature of AI, the discussion about Intellectual Property in this field becomes relevant. There are questions to which the IP regulatory framework already provides clear answers. Others, however, are still in the field of debate. In this regard, the European Parliament, for example, published a resolution in October 2020 to regulate the relationship between IP and AI.

Thus, we shall analyze here the main rules on data protection, cybersecurity and Intellectual Property that must be followed by those who apply or intend to apply Artificial Intelligence technologies. We emphasize that, beyond the cold observance of the law, companies tend to benefit even more when a culture of respect for data and cybersecurity is created. In other words, all the chains of a company must be guided by a common logic of respect and attention to data protection standards and principles, such as collecting only what is necessary for the proposed purposes, adopting appropriate storage techniques, respecting security measures and the like. This point will be further explored in item 5 of this Guide.

[41] According to Brazilian Law, personal data is all "information related to an identified or identifiable natural person" (Article 6, I, LGPD). Sensitive personal data, on the other hand, is a kind of personal data that refers to "racial or ethnic origin, religious conviction, political opinion, affiliation to a union or organization of a religious, philosophical or political nature, data relating to health or sexual life, genetic data or biometric data, when linked to an individual" (Article 6, II, LGPD).

[42] KPMG. **The shape of AI governance to come**. [s.l.]: KPGM International, 2021, page 3.

4.1. Brief analysis of the regulatory scenario

Within the context of protection of personal data and the internet, two laws gain relevance. First, the General Law for the Protection of Personal Data (LGPD) (Law No. 13.709/2018)[43], which regulates the processing of personal data, establishing principles to be followed, rights of data subjects, rules for processing[44] (such as for collection, use and storage), responsibility of the actors involved in this process, sanctions, and even encourages the formulation of good practices and governance on the subject. Thus, the Law specifies the ways to protect the fundamental rights of freedom and privacy and the free development of the individual's personality in the context of the use of personal data. Second, there is the Internet Civil Framework (MCI) (Law No. 12.965/2014), which is aimed at the use of the internet, setting out principles, users' rights, aspects of registration and personal data, liability for damages, among other topics. In addition, there are other specific rules to be observed, depending on the case and area of activity of each company, as already mentioned.

In the international scenario, we highlight the General Data Protection Regulation (EU GDPR 2016/679), which regulates the processing of personal data in the context of the European Union. Considering that the LGPD was drew strong inspiration from this regulation, the laws have great similarities. The rules of the GDPR must also be observed by those who, even in Brazil, process the personal data of data subjects from the European Union. Other international laws may also apply to Brazilian companies, depending on the origin of the data, the residence and citizenship of the data subjects and the commercial relationships established. This is the case, for example, of the California Consumer Privacy Rights Act, which will become effective as of July 2023, and of the Stop Hacks and Improve Electronic Data Security Act (New York's SHIELD Act). It is also worth mentioning the approval, in China, of the Personal Information Protection Law (PIPL), which bears great similarities with the GDPR and the LGPD.

[43] To aid in understanding the LGPD, Demarest has released a booklet explaining the main concepts and measures that should be adopted to protect personal data. DEMAREST. **Lei Geral de Proteção de Dados Pessoais no Brasil**. [s.l.]: Demarest, [s.d.]. Available at: <http://online.fliphtml5.com/hbkzh/kxbw/#p=1>. Accessed on: May 03, 2020.

[44] Processing is "every transaction with personal data such as those concerning the collection, production, reception, classification, use, access, reproduction, transmission, distribution, processing, archiving, storage, deletion, assessment or control of information, modification, communication, transfer, dissemination or extraction" (Article 6, X, LGPD).

4.2. Personal data protection

The use of AI applications, especially in light of the benefits they can generate for a company, can cause a desire for data collection in increasing amounts, which tends to be beneficial for machine learning and deep learning techniques, for example. However, careless collection can have negative consequences for the company, including the application of sanctions. Therefore, it is essential, in relation to any data collection, to pay attention to the rules on data protection set out in the LGPD. The Law provides for the limitation of the processing to the minimum necessary for the accomplishment of its purposes, with the scope of pertinent data, proportional and not excessive in relation to the purposes of the data processing.

To this is added the idea that processing, including collection and use of the data, must comply with the legal rules, such as the requirement of a specific purpose duly informed to the data subject and that such processing must take place in a manner appropriate to the proposed purposes, among other rules that are established by the LGPD.

It should also be noted that data cannot be stored indefinitely and without a purpose for its storage. There are specific rules determining the adequacy of the storage, the possibility of accessing the data by the data subject and, above all, a specific period for this storage. After the intended use, the data must, as a rule, be deleted. An alternative to this is irreversible anonymization, which removes the personal character of the data.

Similarly to when we refer to the adoption of ethics by design, the perspective of privacy by design is also strongly recommended, in regard to the protection of data. In general terms, it refers to observing and respecting privacy at all stages of the system, from its inception. There are seven principles guiding the use of this technique, summarized below:

> They are: 1) *proactive, not reactive; prevent, not remedy*: this indicates that the technique anticipates potential problems and avoids them before they happen, that is, the purpose is to prevent risks from materializing; 2) *privacy as the default setting:* the standard should be to provide as much privacy protection as possible, such that if users do nothing, their privacy will not be at risk. In this context, privacy by default gains significance. Under this model, as soon as a consumer purchases a product or service, standard privacy settings apply and should be as restrictive as possible. (...). 3) *privacy embedded in the design* and architecture of systems and business practices. It is an essential component, not a complement; 4) *full functionality – positive sum, not zero sum:* this indicates that legitimate interests and goals will be accommodated so that everyone wins. (...); 5) *end-to-end security – protection in the complete lifecycle*, that is, the protection of privacy is present from the beginning to the end of the cycle in which the data are involved, starting, as already seen, from the design and creation of the data. systems and devices; 6) *visibility and transparency – keeping open* the possibility of verification of the components and operations performed, as well as the objectives and promises made. This implies the need, for example, to document the steps taken; 7) *respect for user privacy – maintain user-centricity* throughout the data planning and processing process. The focus should be on the users and on protecting their privacy by providing them with strong default settings, adequate information and easy-to-access adjustment options.[45]

The LGPD also affirmed the right to explanation for the field of personal data protection. This right applies to those who use Artificial Intelligence applications that use personal data. The Law provides that, "[the] data subject has the right to request the review of decisions taken solely on the basis of automated processing of personal data that affect their interests, including decisions aimed to define their personal, professional, consumer and credit profile or aspects of their personality." Although the need for the review to be carried out by a human person has been removed from the Law, the doctrine has advocated a review under these terms[46]. Human supervision is even an aspect that is recurrently highlighted as necessary for an ethically and legally responsible use of AI. This provision of the LGPD, which is similar to what is stated in Article 22 of the GDPR, also reinforces the need to keep records of processes, data usage, AI activities and algorithms in general.

[45] The presentation of the principles, based on the works of important authors in the field, was made in MAGRANI; OLIVEIRA. Op. cit., pages 78-79.

[46] MAGRANI, Eduardo; OLIVEIRA, Renan Medeiros de. Op. cit., pages 81-85.

4.3. Cybersecurity

With the concept of cybersecurity, the aim is to adopt technical and administrative measures that ensure the security and integrity of systems (software) and equipment (hardware), networks and digital infrastructure, as well as data protection[47] in the cyberspace, safe storage and technical robustness. In brief, the aim is to have an adequate and safe environment for the processing of personal data and for the operation of AI systems, free from leaks and undue access, with the prevention of incidents such as accidental data destruction, programming errors, among others. In accordance with information security rules, it is necessary to ensure data confidentiality, integrity and availability. For this purpose, it is necessary to anticipate possible incidents, such as hacker invasion. This must be an ongoing concern accompanied by constant updating of security standards, which is reinforced by the fact that even techniques considered highly secure are, at some point, contested, as is the case with biometrics[48].

In this respect, there are AI tools that can be used to detect threats and prevent their materialization, seeking solutions capable of overcoming a particular problem. AI solutions can be used, for example, to verify the identity of who is accessing a particular account or system and to check for unusual and/or inappropriate usage and access patterns. Appropriate hardware and software licenses can be expensive, but there are also open-source options available.

Reinforcing the concern related to the subject and the need for concrete actions, the National Cyber Security Strategy[49] was approved in Brazil, in 2020, through Decree No. 10,222/2020, guiding the actions to be adopted by the Federal Government until 2023.

[47] In relation to personal data, the LGPD expressly establishes these requirements, as provided for in Article 6, VII and Articles 46 et seq.

[48] RODRIGUES, Renato. Falsificação de impressões digitais – difícil, mas não impossível. Kaspersky Daily, Apr. 2020. Available at: <https://www.kaspersky.com.br/blog/sas2020-fingerprint-cloning/14974/>. Accessed on: June 28, 2021.

[49] On this subject in Brazil, refer to HUREL, Louise Marie. Cibersegurança no Brasil: uma análise da estratégia nacional. **Instituto Igarapé – Artigo Estratégico,** volume 54, pages 1-39, April 2021.

4.4. Intellectual Property

With the advances and the increase in the use of Artificial Intelligence, a subject that has gained prominence is its connection with intellectual property (IP). In this relationship, there are points of contact and tension, which are intensified by the strong industrial and/or artistic character of AI. On the one hand, IP provides legal tools to ensure protection for intellectual creations, such as for patenting technologies and for copyright. On the other hand, there are difficulties arising from the very functioning of AI: it is not always possible to clearly determine who is the creator of an AI technology, especially the more complex ones, which involve numerous codes and algorithms. Traceability of intellectual property rights becomes a difficult task in these scenarios.

In the Brazilian regulatory context, the following laws gain relevance in the field of the connection between artificial intelligence and IP: Law No. 9,610/1998, known as the Copyright Law; Law No. 10,695/2003, which deals with violations of copyright; Law No. 9,609/1998, known as the Software Law, which deals with the protection of software intellectual property; Law No. 9,279/1996, referred to as the Industrial Property Law; and Law No. 11,484/2007, which establishes protective provisions for the topography of integrated circuits (chips).

In the international context, it is worth highlighting the role of the European Parliament[50], which, in October 2020, published a resolution on "intellectual property rights related to the development of technologies related to artificial intelligence" (P9_TA(2020)0277).

With technological development, new discussions have emerged, which are also accompanied by reflections on the extent to which IP laws will need to be changed. This is the case, for example, with the debate on attributing authorship to creations made by an AI system. Under the current Brazilian Copyright Law, an author is only the "individual who creates a literary,

[50] For a discussion of AI's connection with IP in the context of the European Union, refer to MORENO, Guillerme Palao. A União Europeia dá seus Primeiros Passos na Regulamentação da Relação entre Inteligência Artificial e Propriedade Intelectual. **Rede de Direito Digital, Intelectual & Sociedade**, Curitiba, volume 1 no. 1, pages 45-68, 2021.

artistic or scientific work" (Article 11, *caput*, of Law no. 9.610/1998). Thus, in the current legislation there is no possibility of attributing authorship to the machine. However, going beyond the current regulation, there is a reflection on the theme in the academic sphere[51], which may, in the future, lead to changes in the law and new regulations as the technology advances. We already have concrete examples of this issue. In 2016, from the analysis and identification of patterns in paintings by the painter Rembrandt, an algorithm created a new painting that became known as *The Next Rembrandt*[52]. Another example is the science fiction short film 'Sunspring', whose screenplay was written by AI system[53] from the analysis of numerous film scripts of the same category.

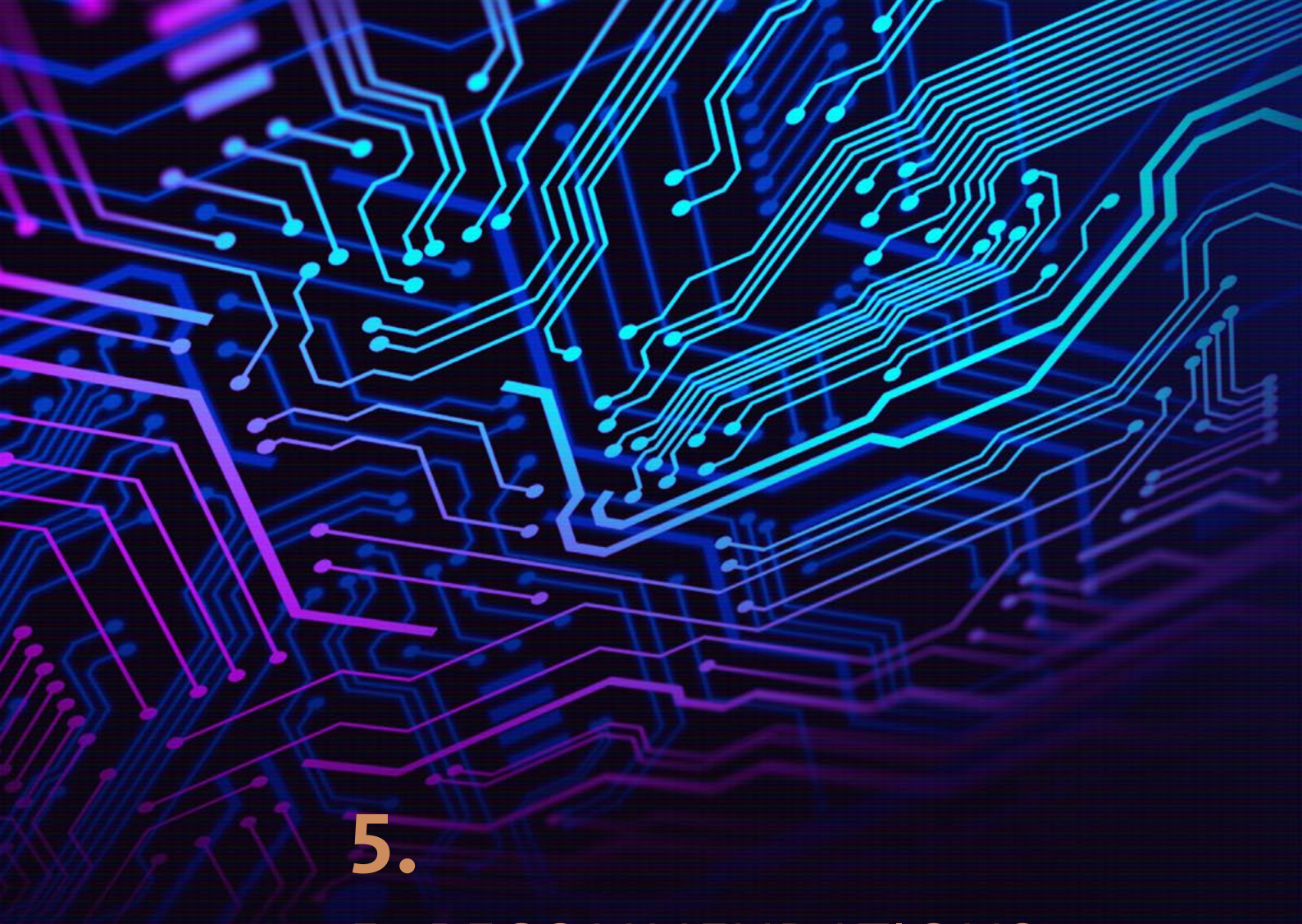
This discussion delves into machine learning and deep learning systems which, after being taught, continue to learn on their own new ways to find the solution to a problem and to give the answers sought. The more complex the technology, as in the case of deep learning and neural networks, the greater its distance from its creator and the more difficult it is to understand how the machine reached a certain result, making it difficult to assign the referred rights. The level of human interaction with the system is also often identified as an important factor in this equation.

Finally, AI has also contributed to the identification and processing of intellectual property rights more quickly than human beings, reducing costs and optimizing time. This is the case, for example, of two YouTube systems: the Copyright Match Tool, which identifies matches or possible matches of videos on the platform, to be analyzed by those who request removal by copyright; and the ContentID, which analyzes the videos posted on the platform and identifies if there is use of any content of a copyright owner, in which case the owner will have certain options, including blocking the viewing of the video.

[51] On this subject, refer to KIM, Daria. 'AI-Generated Inventions': Time to Get the Record Straight? **GRUR International,** volume 69, no. 5, pages 443-456, 2020; WIPO. WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI). Second Session. **DRAFT ISSUES PAPER ON INTELLECTUAL PROPERTY POLICY AND ARTIFICIAL INTELLIGENCE** (WIPO/IP/AI/2/GE/20/1). [s.l.], Dec. 2019. Available at: <https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_ge_20/wipo_ip_ai_2_ge_20_1.pdf>. Accessed on: August 10, 2021; WORLD ECONOMIC FORUM. **Artificial Intelligence Collides with Patent Law**. White Paper. [s.l.], Apr. 2018. Available at: <http://www3.weforum.org/docs/WEF_48540_WP_End_of_Innovation_Protecting_Patent_Law.pdf>. Accessed on: August 10, 2021.

[52] Refer to <https://www.nextrembrandt.com/>. Accessed on: August 10, 2021.

[53] Refer to <https://www.youtube.com/watch?v=LY7x2Ihqjmc&ab_channel=ArsTechnica>. Accessed on: August 10, 2021.



5.

5. RECOMMENDATIONS AND GOOD PRACTICES: CONCRETE MEASURES FOR IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE

As observed throughout the previous chapters, Artificial Intelligence is already a reality present in a large number of organizations as well as in people's daily lives. Many companies already use AI in their procedures and products or are exploring its potential. The absence of specific regulation on the technology in Brazil is not enough reason to prevent its use. Based on constitutional principles, other laws and especially the framework of principles internationally widespread on AI, the regulatory debates carried out and operation of companies already advanced in this process, it is possible to infer certain concrete measures to be taken for an ethical, responsible, and reliable implementation of AI. Thus, we shall present, in this item, recommendations and good practices to be observed along this path. These are measures that are related to ethical principles and whose consideration is necessary in all stages of a company's digital transformation, from the beginning of planning AI incorporation to its results and uses in the company's products.

5.1. A culture of data and ethical, responsible, and reliable performance of operationsperformance of operations

First, it is essential to develop a *culture* focused on Artificial Intelligence, covering issues such as data protection and cybersecurity. We do not ignore that a cultural change in the company is not a simple task. However, organizations have in their favor the fact that the digital transformation has been taking place for some years and people in general are already able to embrace and integrate technology into their daily lives, although in simpler ways. Therefore, this change in the internal culture comes in the midst of the flow of external transformations and amounts to adapting our way of living, working, generating products, among other things.

For Patrícia Prado, there are three fundamental points to be observed in this cultural change[54]. First, the change must come from the company's leaders, who need to "understand what this culture change means and participate in it"[55]. In other words, it is necessary that the concern about data

[54] PRADO, Patrícia. A inteligência dos dados como ativo estratégico para as empresas na era digital. In: LAURIA, Marco A. (Coord.). **Como implementar inteligência artificial** na sua empresa. [s.l.]: I2AI, 2021, pages 9-14.

[55] Ibidem, page 12.

and cybersecurity and the cultural change are noticeable in the actions, plans and statements of the company's directors that will gradually transmit and strengthen this culture to the other employees. Second, Prado highlights the need to "choose metrics based on key business objectives". Accordingly, the company's objectives must be well defined, as well as the operating strategy to achieve them and the measures for their constant evaluation. Finally, the author points out that it is necessary to build a "reliable and transparent database", which includes joint action between the business sectors and the technology sector. We have addressed this point above, by mentioning the need to define and organize the data already during preparation of the AI. Here, we point out the need for such care to be observed throughout the use of AI, even impacting the form of relationship between members of the company.

This culture encompasses not only aspects directly linked to data, but also requires respect for ethical principles on Artificial Intelligence. These principles, mentioned in item 2, must be respected at all stages of the AI. Thus, they need to guide the process of planning objectives and strategies, so that issues such as damage prevention are already considered in this project. As we will see, the principles are not as abstract as they might appear at first reading. In fact, they relate directly to good practices that will be addressed here, and which seek to ensure factors such as transparency, accountability, responsibility, and assurance of the review of decisions.

5.2. Internal mapping and strategy definition

In order to develop a culture in the sense mentioned in the previous item and to implement AI projects, it is essential that an internal mapping is prepared. Before applying changes and including new technologies, it is essential to analyze what the intended objectives are, which strategies will be adopted to achieve them, what tools and infrastructure are available, what level of qualification and training of the team is already part of the company, and, mainly, what problems are faced throughout the processes. AI solutions must be well targeted to specific issues and come with cross-sector support. Otherwise, they will not deliver the expected benefits.

In other words, it is necessary to identify where technology can be used to improve the company's processes, which area requires changes and can benefit most from the transformation, especially in a scenario where there are limited resources. This also involves analysis of already existing results and sincere conversation with all the employees, who are involved in the day-to-day activities and are aware of the details of the company, and who will thus provide more accurate insights, pointing out needs for improvement and bottlenecks.

From there, the strategy can be designed, which will certainly involve the implementation of some changes, such as hiring and/or training of personnel, enhancing computer infrastructure, purchasing software licenses and purchasing equipment. In general, there will be a redefinition of the business model, which will be based, even partially, on AI. The products will be distinguished and further improved.

One thing to be considered while still in strategic planning is the need to dedicate time to the execution of pilot projects, in order to analyze whether the technology is well adapted, whether it is generating the expected results and which aspects should be improved before the actual execution. It is a project to be thought out and applied with caution, mapping and with the mitigation of risks in mind.

Finally, we emphasize that it is desirable that this internal mapping be redone with some regularity, in order to identify flaws and gaps.

5.3. Chart of Professionals /inclusiveness

In the sequence of the mapping and definition of the company's strategy, it will be necessary to focus on the adequate body of professionals. A team responsible for AI must be organized in the company, in charge of its planning, monitoring and execution. This team must involve, among others, engineers, data scientists, product manager, marketing analyst, person in charge of personal data processing (Data Protection Officer – DPO), and a legal team with interdisciplinary knowledge.

An available alternative is the contracting of other companies to perform one, some or all of the steps of implementing and executing projects, which will not require a full AI team in the company. This can be useful for small companies, which cannot yet afford this type of investment; for those that want to take it slow and test AI solutions before incurring huge financial expenses, hiring personnel and changing their infrastructure; or also for large companies that, despite the high number of employees, might not have expertise in one or more areas essential for AI. For all companies, regardless of their size, hiring an external professional responsible for the organization's compliance is a positive step, as the person will be someone with a posture that is not biased by the company's procedures and will have greater capacity and freedom to point out deficits and recommend improvements.

A company tends to benefit from a body of professionals with diverse training and backgrounds[56], which also includes aspects of diversity, such as race and gender. This will allow everyone to contribute from different perspectives, which will prevent risks – as a wider array of people will be analyzing the possible consequences of a project – and will increase the potential for using the technology. Thus, although the formation of this diverse and specific group may demand some investment, it will prove to be a great competitive advantage. The diversity will also demonstrate its benefits upon formulation and setting of algorithms, allowing the identification of biases from the perspective of gender, race, age, nationality, among others. Thus, the company takes an important step to ensure respect for the principle of justice.

[56] In a similar sense, Eduardo Magrani points out: "To achieve the ultimate goal of developing responsible AI, organizations must also be committed to diversity in the workplace. This means that the development of responsible AI permeates the debate and scrutiny of diverse social groups and qualified professionals such as engineers and computer scientists, but also experts in ethics, social scientists, workflow and operations managers, consultants and lawyers. One of the most effective strategies for preventing potential risks associated with AI is hiring employees with diverse skills and experience." MAGRANI, Eduardo. Inteligência Artificial responsável. Op. cit.

5.4. Personnel training

Even with the hiring of specialized personnel, it is also important to invest in the qualification and training of current employees, so that everyone is integrated with the new technologies used in the company. It will be necessary to present and teach how to use the system, provide training on ethical and legal aspects, among others.

5.5. Creation of comprehensive privacy and data policy

The creation of a privacy policy aimed at the user that makes clear all forms of processing of personal data is already a requirement of the LGPD. In the context of Artificial Intelligence, this will continue to be a duty of companies, but such documents will have to be improved in order to also make clear the procedures involving technology. This will meet a requirement for transparency and will reinforce reliability, in addition to serving to standardize the company's behavior.

Thus, in addition to listing all the processing of personal data carried out by the organization in its wide variety of operating contexts, it will be necessary to explain the uses of AI and the way in which the data relate to it. The internal mapping and definition of objectives and strategies will be fundamental at this time, as they will serve as a complementary guide to the policy. The document should contain a presentation of possible risks and their levels, followed by actions to mitigate them. Issues such as mechanisms for systems oversight, planning in cases of incidents and methods for compliance with ethical principles and legal rules should also be included. The document must be easily accessible by the data subjects/user of the technology and present simple and easy-to-understand language, in addition to allowing the data subjects to effectively exercise their rights related to data protection.

5.6. Impact analysis and risk assessment

As already mentioned, the use of AI applications requires constant attention to identify possible risks and act in order to mitigate them. It is necessary to always adopt a preventive stance. There are different parameters for identifying risk levels and the requirements for each. Allianz, for example, identifies three levels of risk in AI applications and in the use of personal data, categorized according to the difficulties in overcoming them and their probability of causing harm[57] : 1) *Low risk:* there may be some risks and inconveniences, but overcoming them tends to be simple and the occurrence of errors is quite low; 2) *Medium risk*: even though the probability of incidence is low, there are significant chances of inconveniences, whose overcoming brings some difficulties; 3) *High risk*: the inconveniences and consequences are more serious, with major difficulty to overcome them, and they might even be irreversible. The probability of damage is high.

The risks and impacts of AI vary both by the technology itself and context. Therefore, determining the level of risk and human involvement must be done on a case-by- case basis. For this, Allianz[58] recommends considering the following aspects:

[57] Definition of risks is made in ALLIANZ. Op. cit., pages 8-9.

[58] ALLIANZ. Op. cit., page 11.

Data Category	The use of sensitive data automatically leads to a medium level of risk, a trend that is often followed by data from social networks as well. For example, because of the amount of sensitive data and the impact of decisions, a health insurance company is classified as high risk.
Type of client	The use of personal data from vulnerable groups also carries a medium risk;
Impact on the client	Legal consequences, impacts of financial aspects and long-term decisions must be taken into account. Thus, an automated decision generates an average risk, which can be reduced by the possibility of explaining, challenging and even reviewing the decision. Financial issues, when problematic, increase risk, as do irreversible long-term decisions.
Field of application	Issues such as advertising without profiling and automating internal processes are often low risk, but it is possible that a combination of several low risk applications creates a high risk scenario.

Mitigation of these risks involves aspects already mentioned, such as applying technologies that have an adequate level of development and maturity; ensuring that technical and cybersecurity requirements are respected and systems are updated; and anticipating and preventing the occurrence of biases and discrimination, which involves comprehensive training on the system and feeding it with varied data.

As stated, there are other parameters to identify, categorize and regulate AI applications according to the actual risks. The proposed regulation of Artificial Intelligence in the European Union - the AI Act mentioned in item 3.2. of this manual - establishes a risk-based approach, but without creating unnecessary restrictions. The idea of forecasts, in general, is to deal with situations where there is justified concern or that such concern can be reasonably anticipated in the near future, so that the AI is reliable, safe, respects legal rules and generates benefits for European citizens.[59] Therefore, the proposal seeks to: deal with the specific risks of AI applications; define which systems are high risk and which requisites must be met in these cases; propose the conducting of a compliance assessment prior to putting a high-risk AI into service or on the market; and also to propose a governance structure at the European and national level.

In this way, the proposal presents four different risk levels, ranging from low to unacceptable risk. Each one demands the observance of specific rules, and requirements are expressly provided for in the proposal. There are even some exceptions to the general rules of risk classification and restriction and/or prohibition of use. Presented below is the outline[60] for each of the risk levels:

Unacceptable risk	Particularly harmful AI applications and prohibited for violating values of the European Union, such as systems for subliminal manipulation of individuals and systems for real time distance biometric identification at public sites for security purposes.
High risk	Systems that pose a high risk to health, safety and fundamental rights of people. They are subject to predictable mandatory requisites to ensure safety and respect of fundamental rights throughout the system lifecycle;
Limited risk	Systems subject to minimal transparency obligations, as in the case of chatbots;
Low risk	Applications with free use in the European market;

[59] MAGRANI, Eduardo. Considerations on the Proposed Regulation for Artificial Intelligence (Artificial Intelligence Act – AIA) in the European Union. 2021. In press.

[60] Op. cit.

5.7. Algorithm auditing

In general, algorithm auditing aims to "evaluate the consistency of the models in relation to the principles and standards in force, focusing on the review of source codes and the impacts of the 'outputs' of the algorithms (predictions indicated by the models)"[61]. In other words, the auditing of algorithms seeks to analyze compliance with legal rules, such as data protection and ethical standards, added to the adequacy of cybersecurity techniques and requirements. The focus is on the practical aspect, that is, if and how the rules and principles are applied in the systems and processes. In addition, the algorithm itself is also analyzed, how it was programmed and which code was used[62], in addition to the results obtained, in order to ensure that they are not discriminatory. In short, this is an important governance and compliance mechanism, aimed at analyzing AI behavior.

With this broad objective in mind, some examples of specific aspects to be verified and some audit tasks are: "to support the decision-making, visualize and monitor the results; inform users of why a decision was made and how to challenge it; alleviate human suffering by anticipating and mitigating damage; allocate responsibilities; and balance conflicts of interest"[63].

The audit can be carried out by a government agency, by an outsourced professional hired for this function, or it can be a task specifically assigned in multilateral organizations[64]. It is important that it be carried out independently from the other tasks of the company.

Note, however, that auditing is not and should not be seen as the single and most important mechanism for ethical and legal compliance. It is a fundamental mechanism that is strongly related to the idea of transparency, but which should be one among several measures and good practices adopted by the company to ensure ethical, responsible and reliable use of AI. There are even some issues and obstacles to be observed in an audit, such as the difficulty of interpreting an algorithm, the different behavior of the algorithm in different contexts[65] and its mutation in contexts of machine learning and deep learning.

Thus, for ethics-based auditing to be viable and effective, Jacob Mokander and Luciano Floridi point out that it must: (i) be a continuous process; (ii) be

part of the sociotechnical system as a whole, and not something isolated; (iii) be seen as a dialectical process, and not a static and preconceived process, ensuring that the appropriate questions and expected answers are given according to that context; (iv) be aligned with organizational policies and incentives; and, reinforcing the importance of ethics by design (v) "interpretability and robustness must be incorporated into systems from the beginning. Ethics-based auditing supports this purpose by providing active feedback to the continuous process of *(re)design*"[66].

In the Brazilian data protection scenario, the provisions of the LGPD, focusing on the principle of transparency, the provisions on accountability and the right to review automated decisions, all reinforce the idea of auditing. There is an express provision attributing competence to the Brazilian National Data Protection Authority (ANPD) to "perform an audit to verify discriminatory aspects in the automated processing of personal data" (Article 20, § 2) in the event that the controller does not provide "clear and adequate information regarding the criteria and procedures used for the automated decision" (Article 20, § 1) and generic provision in Article 55-J, XVI.

[61] KAUFMAN, Dora. Op. cit., page 80.

[62] To preserve the trade secret, Ana Frazão suggests "that at least the main aspects and the logic of algorithmic decision – and especially the decision-making criteria – need to be clarified. Accordingly, it would even be possible to preserve the company's secret, since the code would not be fully revealed, but rather the most prominent aspects of the algorithmic decision, which would be converted from mathematical language to natural language". FRAZÃO, Ana. Transparência de algoritmos x segredo de empresa. Jota, June 09, 2021. Available at: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/transparencia-de-algoritmos-x-segredo-de-empresa-09062021>. Accessed on: July 20, 2021.

[63] KAUFMAN, Dora. Op. cit., page 80.

[64] MOKANDER, Jacob; FLORIDI, Luciano. Ethics-Based Auditing to Develop Trustworthy AI. **Minds and Machines**, Online first, 2021, page 2. Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3788841>. Accessed on: July 30, 2021

[65] KAUFMAN, Dora. Op. cit., page 80 and MOKANDER, Jacob; FLORIDI, Luciano. Op. cit., pages 3-5.

[66] MOKANDER, Jacob; FLORIDI, Luciano. Ethics-Based Auditing to Develop Trustworthy AI. **Minds and Machines**, Online first, 2021, page 3. Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3788841>. Accessed on: July 30, 2021.

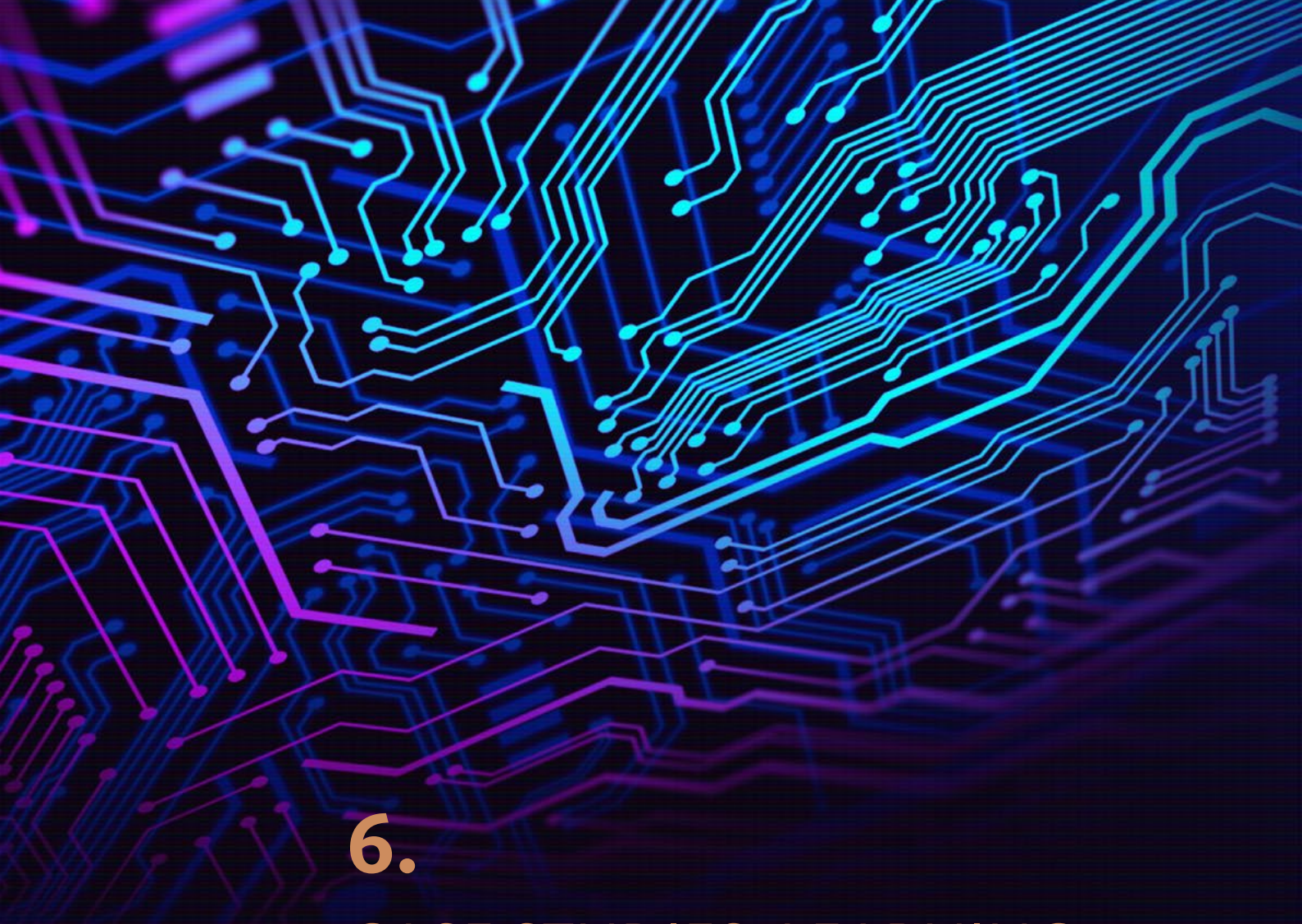
5.8. Report preparation

Finally, it is recommended that reports are prepared detailing how the systems were developed and applied, which processes were implemented and what results were obtained. The preparation and maintenance of up-to-date documents and reports is beneficial for any potential audit and for meeting the accountability requirements. This will also provide clearer control and analysis of processes and systems. Thus, in line with the internal mapping, continuous control of systems and "retrospective impact assessments, carried out in the form of self-control and/or third-party control"[67], contribute to maintaining security and respect for ethical principles and legal rules. As observed by Wolfgang Hoffmann-Riem:

> "To enable this control, obligations of documenting the software and its modifications are offered as support, as well as, in the case of learning systems, the training programs. Obligations to mark the data used, as well as to prepare application records and of the use of training programs, as well as obligations to prepare reports and provide information, can also make sense."[68]

[67] HOFFMANN-RIEM, Wolfgang. Op. cit., page 24.

[68] HOFFMANN-RIEM, Wolfgang. Op. cit., page 24.



6.

CASE STUDIES: LEARNING FROM REAL LIFE EXPERIENCES

In this item, we briefly mention some successful examples of principled approaches and concrete uses of the technology that forms the scope of this manual. The objective is to point out, on the one hand, that there is room for success with the use of AI, while on the other, to ascertain what these concrete experiences can teach us. First of all, we return to an idea already mentioned briefly: that international organizations have also played an important role in this context.

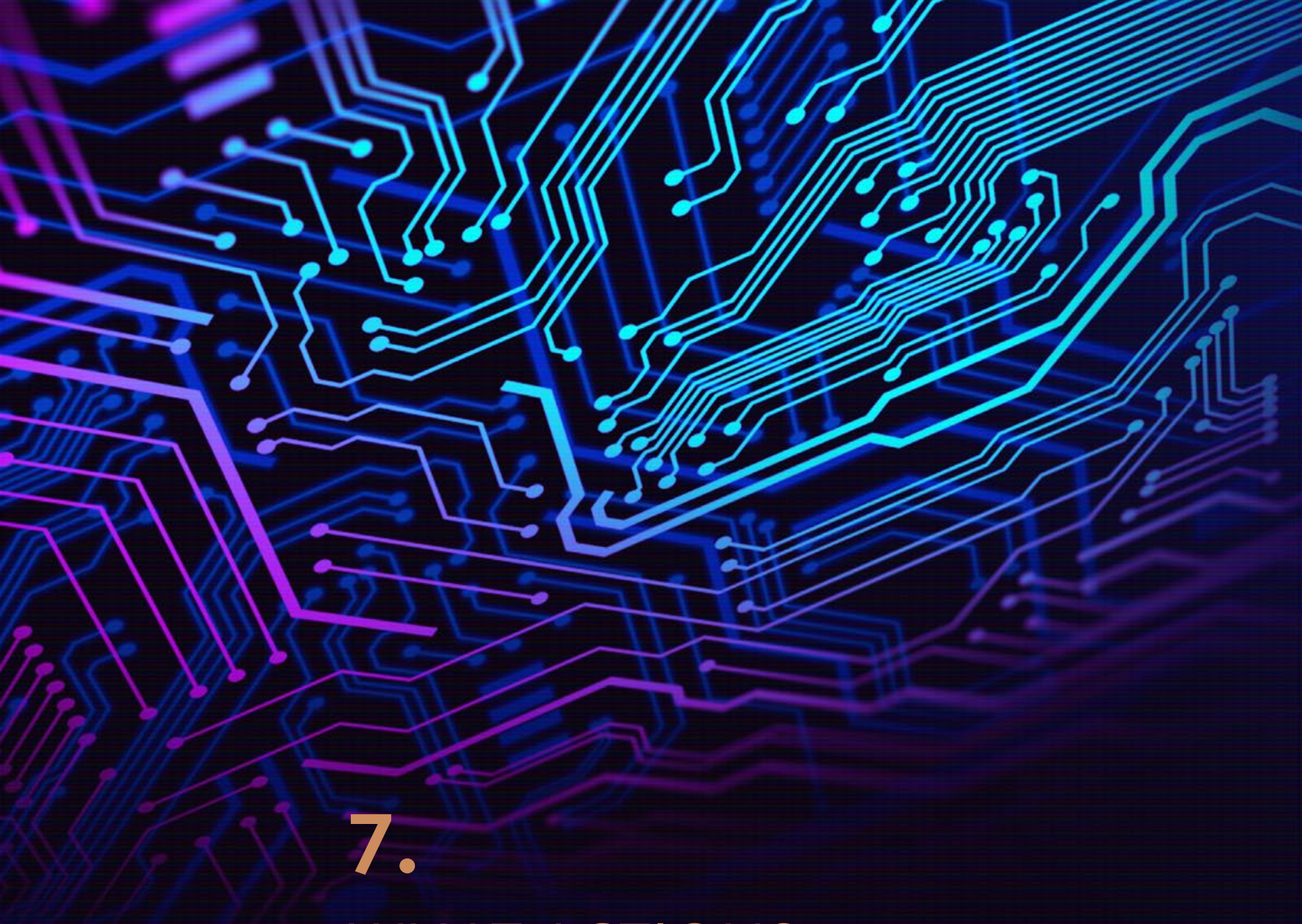
The OECD, for example, brings a principled approach to the technology with the aim of promoting an AI that is "innovative and reliable and respectful of human rights and democratic values". For the organization, responsible use of reliable AI involves observing 5 complementary principles, which can be summarized in the following ideas[69]: i) be beneficial to people and the planet, stimulating sustainable development and well-being; ii) respect for the Rule of Law and adoption of safeguards to guarantee a fair and equitable society; iii) transparency and explanation; iv) robustness and security; v) accountability. These principles have been incorporated by an increasing number of countries and have served as the basis for the elaboration of principled approaches by countries and organizations, such as the G20. This shows us the importance of a unified perspective of principles and the potential of its existence.

A concrete example of positive application was observed through Amazon's Alexa. The famous virtual assistant was the target of concern from parents and associations dedicated to the rights of children who noticed a cold and protocol-based communication in minors. Through a politeness feature strategy, the company programmed the virtual assistant to encourage children to say things like "please" and "thank you"[70]. This case illustrates the attention to the principle of beneficence, focusing the use of AI to be aimed at the welfare and benefits to humanity, in addition to exemplifying the possibility of improvement and revision of technology as it is used. It also demonstrates how a use of ethics by design can work, as these aspects had to be considered at the time of designing the technology.

[69] A complete presentation of the principles adopted by the organization is available at: <https://www.oecd.org/going-digital/ai/principles/>. Accessed on: June 13, 2021.

[70] Refer to AMAZON ALEXA TO reward kids who say: 'Please'. **BBC News,** April 25, 2018. Available at: <https://www.bbc.com/news/technology-43897516>. Accessed on: July 07, 2021.

Finally, it is worth mentioning briefly that there are also negative examples of the use of AI. Image recognition with discriminatory results; the unexpected behavior of chatbots, who end up learning prejudiced speech; inappropriate content recommendations; and the generation of deepfakes and the influence of this category of information on AI systems are some of the negative results to which companies should already pay attention to adopt preventive action now.



7.

WHAT ACTIONS COMPANIES CAN TAKE TODAY

Even without specific comprehensive regulation on AI, ethical principles, international regulations and even national laws on related topics indicate paths that must be followed by those who already wish to adopt the technology. With its potential already quite clear, Artificial Intelligence will bring many benefits to organizations, and those that start preparing now will have a great competitive advantage over other companies, in addition to already taking action to avoid risks.

Bearing this in mind, we have prepared below, in a simplified way, the main steps that should be taken proactively by companies that have encountered Artificial Intelligence, are considering integrating it into their processes and products and do not yet know which path to follow.

1. Develop a data culture and pay attention to the ethical principles of AI at all stages of its application, from planning to its actual use and analysis of results. As mentioned, several principles have emerged to delimit and guide the possibilities for using AI. An analysis of these principles indicates a consensus of ideas to make Artificial Intelligence accountable, transparent and reliable. These ideas are translated into the principles summarized briefly in the table below:

Beneficence	AI must be used for the benefit of Mankind, focusing on the well-being of people, respecting human dignity and paying attention to the planet's sustainability.
Non-maleficence	Use of AI should not cause harm, and it is necessary to anticipate and prevent risks and adopt safety techniques.
Autonomy	AI may not weaken or eliminate human authority. Users must be assured the autonomy.
Justice	AI should promote justice, exclude any prejudice, discrimination and biases and should promote aspects such as solidarity, equity and diversity.
Explicability	AI requires that its use be inteligible (user should be able to undestand how the system works) and there must be accountability (responsibility must be assigned for the system's operation).

2. Map the company's personnel, resources and rules to identify gaps and areas where investment should be increased and processing should be enhanced. It is recommended to hire specialized and diverse personnel, in addition to training the entire workforce of the company;

3. Create policies for the use of AI, data protection, security and other aspects, in order to make clear the general technical, behavioral and security standards to be observed by the company as a whole. It may be necessary to redefine or adapt the planning and business model;

4. Create an AI governance in the company to be observed at all stages involving the technology and its projects, such as strategy, implementation and monitoring. It may be appropriate to create "separate governance committees and councils to deal with the exclusive risks and complexities associated with AI and data"[71];

5. Encourage studies and projects aimed at delivering AI solutions through the creation of a framework for this purpose, which should be guided by clear guidelines be supplied with appropriate tools and investment[72];

6. Invest in risk analysis, creating a responsible structure that is capable of identifying risks in advance and acting to mitigate and avoid them, which may involve the very use of algorithms to identify and manage risks. This includes a joint review with cybersecurity teams and requires attention to aspects of ethical integrity and technical robustness;

7. Prepare and maintain documents and reports on the creation of algorithms, their application and their results. This favors continuous analyses, allows for the modification of systems that produce unwanted results, helps to prevent risks and proves to be useful in carrying out a potential audit.

[71] KPMG. Op. cit., page 12.

[72] KPMG. Op. cit., page 12.

GLOSSARIES

Although we are dedicated to detaching ourselves from the technicality of the language, the study and use of Artificial Intelligence is accompanied by frequent and important technical words and expressions for understanding the discussions on the topic. There is, in fact, specific vocabulary that should be known to those involved in AI. This has been taken into account throughout this Manual, in which we seek to present the concepts used, which is also something observed by other authors and organizations. Therefore, it is not uncommon to find glossaries presenting the main concepts used in the area. Below, we indicate some of these documents for consultation:

Glossary of the European Council. Available at: <https://www.coe.int/en/web/artificial-intelligence/glossary>. Accessed on: June 20, 2021.

Definition of AI (and related issues) prepared by the European Commission's Independent High-Level Expert Group on Artificial Intelligence. The definition follows the "Ethical Guidelines for Reliable Artificial Intelligence". Available at: <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60669>. Accessed on: July 20, 2021.

Glossary of I2AI (International Association of Artificial Intelligence). Available at: <https://www.i2ai.org/content/glossary/>. Accessed on: July 30, 2021.

RANSCHAERT, Erik R.; MOROZOV, Sergey; ALGRA, Paul R. (Eds.). **Artificial Intelligence in Medical Imaging**: Opportunities, Applications and Risks. Cham: Springer, 2019, pages 347-363[73]. Available at: <https://link.springer.com/content/pdf/bbm%3A978-3-319-94878-2%2F1.pdf>. Accessed on: June 18, 2021.

[73] Although the book is focused on the use of AI in medical imaging, the concepts, in general, are neither specific nor exclusive to medicine.

REFERENCES

AMAZON ALEXA TO reward kids who say: 'Please'. **BBC News**, Apr. 25, 2018. Available at: <https://www.bbc.com/news/technology-43897516>. Accessed on: July 07, 2021.

BOSTROM, Nick. **Superintelligence**: Paths, Dangers, Strategies. Oxford: Oxford University Press, 2014.

_____; YUDKOWSKY, Eliezer. The ethics of artificial intelligence. In: FRANKISH, Keith; RAMSEY, William M. (eds.). **The Cambridge Handbook of Artificial Intelligence**. Cambridge (UK): Cambridge University Press, 2014, pages 316-334.

DEMAREST. **Lei Geral de Proteção de Dados Pessoais no Brasil**. [s.l.]: Demarest, [s.d.]. Available at: <http://online.fliphtml5.com/hbkzh/kxbw/#p=1>. Accessed on: May 03, 2020.

FLORIDI, Luciano; COWLS, Josh. A unified framework of five principles for AI in society. **Harvard Data Science Review**, no. 1.1, summer 2019.

FORNASIER, Mateus de Oliveira; KNEBEL, Norberto. Inteligência artificial: desafios e riscos ético-jurídicos. **Direito & Paz**, São Paulo, volume XIV, no. 43, pages 207-228, Jul.-Dec. 2020, pages 207-228.

FRAZÃO, Ana. Transparência de algoritmos x segredo de empresa. **Jota**, June 09, 2021. Available at: <https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/transparencia-de-algoritmos-x-segredo-de-empresa-09062021>. Accessed on: July 20, 2021.

HOFFMANN-RIEM, Wolfgang. Inteligência artificial como oportunidade para a regulação jurídica. **RDU**, Porto Alegre, volume 16, no. 90, pages 11-38, Nov./Dec. 2019.

HUREL, Louise Marie. Cibersegurança no Brasil: uma análise da estratégia nacional. **Instituto Igarapé – Artigo Estratégico**, volume 54, pages 1-39, Apr. 2021.

I2AI. **Inteligência Artificial**: Guia Rápido. [s.l.]: I2AI A Connected AI World, [s.d.].

IBM WATSON. **Global AI Adoption Index 2021**. [s.l.]: IBM Watson, 2021.

Jobin, Anna, et al. The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399. 2019 .

KAUFMAN, Dora. Inteligência Artificial e os desafios éticos: a restrita aplicabilidade dos princípios gerais para nortear o ecossistema de IA. **PAULUS: Revista de Comunicação da FAPCOM**, São Paulo, volume 5, no. 9, pages 73-84, Jan./Jul. 2021.

KIM, Daria. 'AI-Generated Inventions': Time to Get the Record Straight? **GRUR International**, volume 69, no. 5, pages 443-456, 2020.

KPMG. **The shape of AI governance to come**. [s.l.]: KPGM International, 2021.

LATZER, Michael et. Al. The Economics of Algorithmic Selection of the Internet. **Working Paper – Media Change & Innovation Division**, Oct. 2014. Available at: <https://mediachange.ch/media//pdf/publications/Economics_of_algorithmic_selection_WP.pdf>. Accessed on: July 07, 2021.

LAURIA, Marco A. (Coord.). **Como implementar inteligência artificial na sua empresa**. [s.l.]: I2AI, 2021.

LEMOS, Ronaldo. Estratégia de IA brasileira é patética. **Folha de S. Paulo**, April 11, 2021. Available at: <https://www1.folha.uol.com.br/colunas/ronaldolemos/2021/04/estrategia-de-ia-brasileira-e-patetica.shtml>. Accessed on: July 22, 2021.

MAGRANI, Eduardo. Considerações sobre a Proposta de Regulamento para a Inteligência Artificial (Artificial Intelligence Act – AIA) na União Europeia. 2021. Awaiting publication.

_____. **Entre dados e robôs**: ética e privacidade na era da hiperconectividade. Porto Alegre: Arquipélago Editorial, 2019.

_____. Inteligência Artificial responsável. O que é isto? **Estadão**, March 13, 2021. Available at: <https://politica.estadao.com.br/blogs/fausto-macedo/inteligencia-artificial-responsavel-o-que-e-isto/>. Accessed on: June 03, 2021.

_____. New perspectives on ethics and the laws of artificial intelligence. **Internet Policy Review**, volume 8, no. 3, pages 1-19, Sept. 2019.

_____; OLIVEIRA, Renan Medeiros de. Desafios regulatórios da Internet das Coisas: reflexões sobre a LGPD, o consentimento e o direito à explicação. In: CRESPO, M. X. F. (Coord.). C**ompliance no direito digital**. São Paulo: Thomson Reuters Brasil, 2020, pages 65-91.

MCCARTHY, John; MINSKY, Marvin L; ROCHESTER, Nathaniel; SHANNON, Claude E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. **AI Magazine**, volume 27, no. 4, pages 12-14, 2006.

MCKINSEY GLOBAL INSTITUTE. **Note from the AI frontier**: modeling the impact of AI on the world economy. Discussion paper, McKinsey&Company, 2018.

MOKANDER, Jacob; FLORIDI, Luciano. Ethics-Based Auditing to Develop Trustworthy AI. **Minds and Machines**, Online first, 2021. Available at: ‹https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3788841›. Accessed on: July 30, 2021.

MORENO, Guillerme Palao. A União Europeia dá seus Primeiros Passos na Regulamentação da Relação entre Inteligência Artificial e Propriedade Intelectual. **Rede de Direito Digital, Intelectual & Sociedade**, Curitiba, volume 1 no. 1, pages 45-68, 2021.

NILSSON, Nils J. **The Quest for Artificial Intelligence:** A History of Ideas and Achievements. Cambridge: Cambridge University Press, 2010.

NUNES, Ana Carolina. Estratégia Brasileira para IA é um passo importante, mas já chega precisando de atualização. **Época Negócios**, May 28, 2021. Available at: ‹https://epocanegocios.globo.com/Tecnologia/noticia/2021/05/estrategia-brasileira-para-ia-e-um-passo-importante-mas-ja-chega-precisando-de-atualizacao.html›. Accessed on: July 22, 2021.

PAIVA, Fernando. A EBIA é uma estratégia sem estratégia, diz Eduardo Magrani. **Mobiletime**, April 15, 2021. Available at: ‹https://www.mobiletime.com.br/noticias/15/04/2021/a-ebia-e-uma-estrategia-sem-estrategia-critica-eduardo-magrani/›. Accessed on: July 22, 2021.

PARENTONI, Leonardo Netto; VALENTINI, Rômulo Soares; ALVES, Tárik César Oliveira e. Panorama da regulação da inteligência artificial no Brasil: com ênfase no PLS n. 5.051/2019. **Revista Eletrônica do Curso de Direito da UFSM**, Santa Maria, volume 15, no. 2, pages 1-29, May/Aug. 2020.

PASQUALE, Frank. **The Black Box Society**: The Secret Algorithms That Control Money and Information. Cambridge (MA) and London: Harvard University Press, 2015.

ROBERTS, H. et al. The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. **AI & Soc**, volume 36, pages 59-77, 2021.

RODRIGUES, Renato. Falsificação de impressões digitais – difícil, mas não impossível. **Kaspersky Daily**, Apr. 2020. Available at: ‹https://www.kaspersky.com.br/blog/sas2020-fingerprint-cloning/14974/›. Accessed on: June 28, 2021.

ROSA, João Luís Garcia. **Fundamentos da inteligência artificial**. Rio de Janeiro: LTC, 2011.

RUSSEL, Stuart J.; NORVIG, Peter. **Artificial Intelligence**: A Modern Approach. 4 ed. Hoboken: Pearson, 2021.

SALOMÃO, Luís Felipe (Coord.). **Inteligência Artificial:** tecnologia aplicada à gestão dos conflitos no âmbito do poder judiciário brasileiro. [s.l.]: FGV Conhecimento, 2020.

TURING, Alan. Computing Machinery and Intelligence. **Mind**, volume LIX, no. 236, page 433-460, Oct. 1950.

UOL. Carro autônomo da Uber teve 37 acidentes antes de matar pessoa nos EUA. **Uol**, November 08, 2019. Available at: ‹https://www.uol.com.br/carros/noticias/redacao/2019/11/08/carro-autonomo-da-uber-teve-37-acidentes-antes-de-matar-pessoa-em-2018.htm›. Accessed on: June 14, 2021.

VERBEEK, P. **Moralizing Technology**: Understanding and Designing the Morality of Things. Chicago and London: The University of Chicago Press, 2011.

WIPO. WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI). Second Session. **DRAFT ISSUES PAPER ON INTELLECTUAL PROPERTY POLICY AND ARTIFICIAL INTELLIGENCE** (WIPO/IP/AI/2/GE/20/1). [s.l.], Dec. 2019. Available at: ‹https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_ge_20/wipo_ip_ai_2_ge_20_1.pdf›. Accessed on: August 10, 2021.

WORLD ECONOMIC FORUM. **Artificial Intelligence Collides with Patent Law**. White Paper. [s.l.], Apr. 2018. Available at: ‹http://www3.weforum.org/docs/WEF_48540_WP_End_of_Innovation_Protecting_Patent_Law.pdf›. Accessed on: August 10, 2021.

Artificial Intelligence (AI) is a topic of growing importance in Brazil and indeed in the international context. Its various benefits, uses and applications are part of the daily lives of individuals, and companies are also observing this movement. Many of them have already incorporated, to some extent, technology in their processes and want to deepen this use even further. Others are observing the technological changes and, realizing that in the near future AI will inevitably be part of their processes, companies are already seeking to adopt it now in products and procedures, in anticipation of market changes. Chatbots, recommendation algorithms, support in the recruitment of professionals, prevention of anomalies and errors, forecasting of sales and prices and aid in medical diagnosis are just some of the innumerable applications possible. However, this topic is not always discussed based on a clear concept of what AI is, and ethical-legal aspects already established in the local and international scenes are also not considered in their entirety. Therefore, the purpose of this Manual is to provide, in accessible language, a comprehensive presentation of Artificial Intelligence, dealing with its concept, the principles and rules involved and the consequences of its use. The approach used here is aimed at the practical application of technology, focusing on serving as an aid for companies that intend to implement it in their products. Thus, we seek to provide guidelines for an ethical and responsible use of AI based on the best practices adopted in the market, intensifying its potential applications and ensuring compliance with mandatory measures.

SÃO PAULO
RIO DE JANEIRO
BRASÍLIA
NEW YORK

demarest.com.br