

# Artificial Intelligence: fairness, justice and consequences

By Diogo Cortiz <sup>1</sup>

Artificial Intelligence (AI) is infiltrating every aspect of our lives. Everyday situations, such as movie recommendations on streaming services, or purchase cancellations on credit cards through anti-fraud mechanisms, are examples of decision-making that is probably arising from AI systems. However, as the development and adoption of this type of technology for delivering quicker and safer services grows, there is increased concern regarding possible side effects.

A central theme in discussions around the world, AI has drawn the interest of experts from a wide range of disciplines. While economists are concerned about impacts on productivity and unemployment; psychologists and neuroscientists are investigating possible consequences on cognition and behavior; lawyers are concerned about changes in the judicial system; and sociologists, about social transformations. In international politics, State leaders have declared it to be one of the most valuable assets for social, political, and economic development.

In recent years, there has been a race among nations in terms of plans, projects and incentives to accelerate AI expansion within their borders. All developed and some developing countries have created strategic plans to establish their leading role in this field. In its public document *Next Generation Artificial Intelligence Development Plan*<sup>2</sup>, China has set the goal of becoming the global leader in AI by 2030.

Given the equipment, techniques and computational models currently available, it is not likely that machines will achieve singularity<sup>3</sup> in the near future – perhaps not even in the 21<sup>st</sup> century. Nevertheless, it is important to pose a philosophical discussion about the risks of AI for humanity. Even though it can be beneficial to society, assisting, for example, in the optimization of processes or in the discovery of new drugs and diagnoses, it also brings unexpected situations, since it is an emerging technology. If the necessary safeguards are set in place, it will be possible to maximize its capabilities and minimize its potential risks.

<sup>1</sup> Researcher at the Web Technologies Study Center (Ceweb.br), of the Brazilian Network Information Center (NIC.br), and professor at Pontifical Catholic University of São Paulo (PUC-SP), with a PhD and master's degree in intelligence technologies and digital design, a PhD Fellowship from Université Paris 1 – Sorbonne, and an MBA in international economics from the University of São Paulo (USP). He did a postdoctoral internship at the Artificial Intelligence and Creative Technologies Laboratory at the University of Salamanca and is a visiting researcher in the cognitive sciences laboratory of Queen Mary University of London.

<sup>2</sup> Available at: <http://fi.china-embassy.org/eng/kxjs/P020171025789108009001.pdf>.

<sup>3</sup> The theory that machines may develop to the point of surpassing human intelligence. However, this perspective has been rejected by various prominent researchers in the field of AI.

## /Internet Sectoral Overview

---



### **Diogo Cortiz**

Researcher at  
the Ceweb.br  
and professor at  
PUC-SP.

At present, the technologies used in AI can be divided into the following main categories: supervised learning; unsupervised learning; and reinforcement learning. Supervised and unsupervised learning are the most common, and they depend on large volumes of data in order for machines to learn. In the first case, the data needs to have been previously “labeled,” in reference to what it represents. When a set of data on credit card transactions includes a field that specifies whether fraud has occurred or not, for example, there is supervised learning, and it is possible to train an anti-fraud system to classify future purchases. When the same dataset does not have that field, there is unsupervised learning, and the development of an anti-fraud system is not possible.

A well-trained AI system can recognize patterns that are not very visible even to experts. However, the source and quality of the data used in this system influence how it will behave. In this context, in 2016, the investigative journalism network ProPublica published an article<sup>4</sup> that posed a new challenge: Data can have discriminatory biases, and AI is able to leverage them. With the provocative title “Machine Bias,” the article showed that one of the systems used the most in the United States to assess the risk of criminal recidivism had biased results according to the skin color of the person analyzed, classifying black people as two times more likely to reoffend than white people. Through learning the patterns of the sentences issued by judges, the system started deliberating in a similar manner. This case leads to reflection as to whether it is acceptable that a system perpetuates undesirable behavior, even though it has learned from a history of human decisions.

Used in much research, as well as by the technical and scientific community, the terms “discriminatory algorithm” or “discrimination in algorithms” spark disagreement among researchers. Some argue that algorithms can truly “be prejudiced,” whereas others find this hypothesis absurd. The problem lies in the fact that the word “algorithm” is used to refer to two distinct things (illustrated in Figure 1, as “Training algorithm” and “Trained model”) in the development process of an AI project, although both are, in fact, algorithms.

As well as depending on data, AI requires training algorithms. There are innumerable algorithms – decision trees and neural networks, among others – that are recommended, depending on the situation and objectives of the project. Algorithms can be compared to a “cake recipe” that uses training data as “ingredients” to produce something, such as a fraud detection model<sup>5</sup>. Therefore, there are two algorithms in an AI project: a training algorithm and a trained model (which will, in fact, be used in production). A training algorithm is a set of rules that does not express an opinion or contain bias of any kind. However, its output, which is the model, can manifest biased behavior due to the data used in the training. Figure 1 illustrates this process:

<sup>4</sup> Available at: [www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing](http://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing)

<sup>5</sup> Although the model is also an algorithm, I decided to name it as such for didactic purposes. Apart from this, it is a term used frequently in the AI community.

Figure 1- TRAINING PROCESS



Source: Prepared by the author.

Let us consider two sets of data. Even though the same training algorithm is used (a decision tree, for example), it will result in two models with different behaviors. If in one of the cases the dataset used represents a profile of a reality where black people have less access to the health system than white people, this pattern will be reflected in the model.

Therefore, since an AI system reflects the patterns of the training data, the database used in the learning process must be taken into account. One possible strategy would be to analyze and understand the data before training, but this is not a simple task. The algorithms work with an enormous amount of data, making it exceedingly difficult for a person to do the analysis alone, or even detect subtle patterns. Nonetheless, the data cannot be discarded, since ultimately it is needed so that the AI system can learn. Neither is it always possible to perform a prior analysis of the patterns that will be learned by the model. Perhaps the most important aspect will be to examine its behavior after training and determine whether it represents the values desired to be reproduced by a system.

In this sense, it is necessary to define new assessment metrics so that AI models take into consideration, not only technical requirements, but also social aspects. As a rule, systems are developed by scientists and engineers based on performance metrics, such as accuracy. A model with an acceptable success rate according to these criteria is suitable for implementation in a production system, regardless of manifesting bias in one direction or another.

However, many undesirable cases, as discussed in the ProPublica article, demonstrate the importance of intensified work to define additional metrics for AI models. This is not simple, since the technical complexity of the models constitutes a challenge. However, it is necessary in today's context, as shown by current discussions on AI governance. Organizations in the area have argued for heterogeneous and interdisciplinary teams to lead AI projects, formed by people capable of establishing metrics to ensure that the system respects human rights and behaves in accordance with the social and cultural values of the region.

The geographic issue is also a relevant point in AI projects. Most of the tools used today in the Global South were developed by companies in the North, which generally fail to take into consideration the specific cultural aspects of our reality. The most popular sets of data focus on the United States and Western Europe, and data from other places, when it exists, tends to be neglected by the architects of the systems, who are unfamiliar with the cultures of those places.

(...) it is necessary to define new assessment metrics so that AI models take into consideration, not only technical requirements, but also social aspects.

When a system goes live, the results are immediate and large-scale. If it is not adequately tested and approved, the consequences may be problematic for certain social groups.

It is quite common, for example, for image recognition systems to classify the photo of a woman in a white dress as a wedding ceremony, but they will rarely do the same in the case of a bride wearing a sari in an Indian celebration. Major companies are seeking alternatives to address this issue, such as the Crowdsourcing project by Google, which provides a website and application so that people can add regional content and thereby expand its database.

It is clear, therefore, that the complexity of designing AI projects extends beyond technical aspects. When a system goes live, the results are immediate and large-scale. If it is not adequately tested and approved, the consequences may be problematic for certain social groups. This was demonstrated in a study published in the journal *Science*<sup>6</sup>, in 2019, on how a system for allocating preventive treatment for patients with chronic diseases in the United States manifests discriminatory behavior toward black people. While carrying out the research, the authors found that, within the context of the study, a black patient would need to be sicker than a white patient to be entitled to receive the same care. This is because the system allocates treatment based on the prediction of future care costs for each patient, i.e., it does not consider the individual's medical condition, but just the financial variable. Based on the historical data of the U.S. health system, to which black people have limited access, the system learned that the cost is higher for white patients, on average, and that they should therefore receive preventive treatment. The authors of the study argue that to build a more equal system, the project design should prioritize a trained model to identify patients according to medical and not financial parameters.

Regarding the criteria for such a system, the term “fairness” has become popular in the academic community. Published in 2019 by the Web Technologies Study Center (Ceweb.br)<sup>7</sup>, the *Mapping Principles of Artificial Intelligence*<sup>8</sup> showed that the concept of fairness is always present in principles documents released by leading technology companies, governments and international organizations, which denotes concern about preventing AI from perpetuating injustice. However, working with the idea of fairness is not a simple task. Reflections on the concept of justice are not restricted to the field of AI, but are a constant in the human sciences. In the technical and scientific field, there is concern about generating broad discussions to incorporate the topic and make it a relevant aspect of present and future AI projects.

The subject has been evolving through events organized by the academic community. The Association for Computing Machinery, an important international association in the area of computation, launched the FAccT Network<sup>9</sup>, a network of conferences that address the topics of fairness, accountability and transparency in technology. Another initiative was the FATES (*Fairness, Accountability, Transparency, Ethics and Society on the*

<sup>6</sup> Available at: [science.sciencemag.org/content/366/6464/447](https://www.science.org/doi/10.1126/science.1234567)

<sup>7</sup> Ceweb.br, a department of NIC.br, aims at enabling participation of the Brazilian community in the global development of the Web and assist in the formulation of public policies. More information can be found at: <https://ceweb.br/about-ceweb-br/>

<sup>8</sup> Available at: [ceweb.br/publicacao/mapeamento-de-principios-de-inteligencia-artificial/](https://ceweb.br/publicacao/mapeamento-de-principios-de-inteligencia-artificial/)

<sup>9</sup> Find out more: <https://facctconference.org/network/>

<sup>10</sup> Find out more: [fates.isti.cnr.it/](https://fates.isti.cnr.it/)

Web) workshop<sup>10</sup>, organized by Ceweb.br and held in conjunction with the prestigious The Web Conference 2020<sup>11</sup> to discuss the social impacts of AI.

It is worth investigating whether the search to define a single criterion for fairness is a mistake, an attempt to simplify something that is complex and plural. Let us look at the hypothetical situation narrated by Amartya Sen in the book *The Idea of Justice*. In relation to three children, you must decide which of them will receive a flute: Bob, a poor child who has no toys; Carla, who has always cleaned the flute; or Anne, the only one who knows how to play it. There are justifications for each choice. Someone with utilitarian beliefs would argue that the flute should go to Anne, whereas an egalitarian would suggest that Bob is the worthiest. In turn, for a libertarian, Carla would deserve the toy. The three solutions, which are totally different, are based on rational arguments and supported by philosophical frameworks. Which is the fairest?

Reflections about what constitutes fairness do not end in the realm of social discussion and are far from being fully developed in the field of AI. Initial efforts have been made<sup>12</sup>, but the path is long. At present, it is urgent to discuss fairness and transparency, when AI applications are found in various aspects of the decisions made in our midst, but the criteria still lack analysis and development of techniques for implementation in models. If it is difficult to define what is fair in the social sphere, how can this be established at the technical level?

One strategy would be to strive to fight that which is unfair. Part of the AI community understands unfairness to mean any harm or benefit that could be caused by an irrelevant characteristic. Despite seeming simple, this definition is complex, since deciding that a characteristic is irrelevant depends on the situation and context of the application.

In a selection process, for example, the candidate's gender could be considered an irrelevant characteristic. There is the well-known case of Amazon's recruitment system, which lowered the score of resumes submitted by women, in violation of fairness. This does not mean, however, that gender should be discarded in training all AI models, since it can be an important and sensitive aspect, such as in a diagnostic system in the health sector. This also applies to other sociodemographic characteristics, such as income, race, ethnicity and religion. For this reason, it is essential to have heterogeneous and interdisciplinary teams discuss the issue.

There is no ready-made formula for dealing with concepts of ethics, fairness and justice in AI, even though they are highly relevant concerns. Technologies evolve rapidly, and the discussion of social impacts and necessary strategies for making AI models less unfair is in the beginning stages. In this context of rapid transformation, this article sought to point out possible paths, with an understanding that, at present, there are more questions than answers. The challenge has been set.

<sup>11</sup> Find out more: [www2020.thewebconf.org/](http://www2020.thewebconf.org/)

<sup>12</sup> As a recent example, the publication *Responsible AI: A Global Policy Framework*, from 2019, is the result of work by more than 50 researchers in the field. The book devotes a chapter to fairness and non-discrimination, addressing issues related to the topic in different contexts and dimensions.

<sup>13</sup> Find out more: [www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G](http://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G)

Technologies evolve rapidly, and the discussion of social impacts and necessary strategies for making AI models less unfair is in the beginning stages.

# Interview I



**Nnenna  
Nwakanma**

Chief Web  
Advocate at  
World Wide  
Web Foundation.

***I.S.O.\_ In a context where technologies are increasingly applied in different sectors of society, is Artificial Intelligence (AI) more likely to reduce or exacerbate existing inequalities? What are possible future scenarios?***

**N.N.\_** In looking back at what was called technology then, and at what is seen as technology now, we can see a pattern. We can only build on existing social and economic infrastructure. Populations that have access to land, education, information, and basic human rights are those that are faring better in access to digital opportunities.

While in the developed countries a high percentage of the population may have access to broadband connectivity, in many developing nations, the divide is widening. Around the world, only about half the global population is connected to the Internet – the other half is left behind. Research by the Alliance for Affordable Internet (A4AI)<sup>14</sup> and its partners has shown that people that are yet to be connected are mostly poor, live in rural area, and mainly women.

AI runs on data, efficient information transmission and processing, and reliable connectivity. The possible scenario I see is that, as long as it makes sense economically, we will keep developing solutions for the available “markets,” for the “digitally capable” and the “data literates,”. It is possible that at some point, the gap will become so wide that building bridges may become so capital-intensive that we may just decide it is no longer worth the effort to connect the last billion of people.

In sum, the way we are currently running the technology and innovation space can only exacerbate existing inequalities.

***I.S.O.\_ How can Artificial Intelligence be used as a tool to safeguard human rights? Does the use of AI systems pose possible risks or adverse impacts to human rights, and, if so, how could these be mitigated?***

**N.N.\_** I have seen diverse uses of AI in citizen services and, in general, in the commercial and economic sectors. Robotics has come in handy in health services, education and digital governance. In the context of the new coronavirus, a lot of governments have deployed track-and-trace technological solutions that are powered by AI; in times of crisis, when life and death of individuals can be determined by five-minute latency, the use of AI can save lives.

However algorithmic intelligence, based on data, habits, personal preferences and use of interfaces, is still limited. In the case of the abovementioned pandemic, issues regarding transparency in data governance, respect for individual privacy, and use and abuse of personal data, already present huge challenges.

There are also concerns about algorithmic manipulation for political and election purposes, especially on social media platforms. The role of AI in

<sup>14</sup> See: [www.A4AI.org](http://www.A4AI.org)

viral misinformation is one that we need to explore more, as human lives and businesses have actually been harmed.

In mitigating adverse effects resulting from the use of AI, I am strongly in favour of the multi-stakeholder, multi-sector approach. I cannot think of any technology problem that can be solved by one stakeholder group alone.

Led by Tim Berners-Lee, the inventor of the Web, the Contract for the Web<sup>15</sup>, presents a set of principles. I emphasize the sixth, which challenges us to develop technologies that support the best in humanity and challenge the worst. This means that in our global multi-stakeholder digital cooperation, we must keep our eyes open to what the dangers are. The one thing we must understand is that dangers in technology evolve on a daily basis – what constituted a major danger yesterday is not the same as that of today, and certainly the danger of tomorrow will be different.

While governments play the governance role, companies must develop technology in a way that respects human rights, and civil society must play its role as a watchdog. And these roles may shift from time to time. That is why we need continued dialogue, collaboration, research and cooperation.

***I.S.O.\_ Given that the datasets that feed AI systems may be imprecise, biased or incomplete, what are the possible implications of using AI technologies for policymaking? Which population groups are most likely to be affected by this situation? How can these issues be addressed?***

***N.N.\_*** I don't see policymaking as a one-off action. I prefer to look at the policy process, the end-to-end cycle, instead of one action in particular. AI technology can be used at different stages in the policy process: to follow up on early trends; to identify and track emerging problems; to set agendas; to consult about and consider policy options; to support decision-making and implementing decisions; and to monitor and evaluate implementation.

We have the opportunity to use AI technology throughout the process in order to reduce delays, allow faster action, maximize budgets, increase accountability, mine data, allow fast and efficient reporting, and most importantly, recognize and provide alerts about emerging trends.

Which brings us to the core question: Whose data is available? Whose data is being well-collected, well-managed and well-analysed? If my data is missing, I will be affected.

What type of data is being prioritized? If the datasets that are the most meaningful to me are not included on databases used by AI, it only follows that my issues will not even make it to the table for any decisions. If the data that speaks to me is missing, I will be affected.

I need to be engaged to have a voice in the process. What citizen space is there for inclusive policy engagement? Who gets to come to the table? Is my voice heard? As a woman? As a person in the rural community? As an unconnected person? As a person from a minority group? As a person living with a disability? Do I have enough agency – social, digital, economic, and political – to be at the table? If the answer is “No,” I will lose out.

Building just and equal societies is a difficult task. There are no perfect models. The original vision of Tim Berner-Lee and other founders regarding

"In mitigating adverse effects resulting from the use of AI, I am strongly in favour of the multi-stakeholder, multi-sector approach. I cannot think of any technology problem that can be solved by one stakeholder group alone."

<sup>15</sup> See: [www.contractfortheweb.org](http://www.contractfortheweb.org)

the Web was that giving technology to the world meant that humanity would make positive use of it. It is, in his own words, #ForEveryone. This principle should therefore be our most important guide when we design, develop, deploy and implement technologies.

## Article II



### **Luiz Fernando Martins Castro**

Partner in the firm Martins Castro Monteiro Advogados, and board member of the CGI.br.

## Use of Artificial Intelligence in the judicial realm: myths and reality

By Luiz Fernando Martins Castro <sup>16</sup>

**A**lthough it is a branch of computational science that has been known about and discussed for more than 60 years, Artificial Intelligence (AI) is undoubtedly the subject of the moment due to its wide reach and ability to cause important transformations in society. Due to the evolution of processors and software, AI use has been growing, with an immediate impact on repetitive and less complex activities, directly affecting citizens and workers.

The judicial world – considered herein as all those working in the field of law, including lawyers and judicial services – is aware of the possibilities of using AI in its daily activities. There have been many conjectures that AI solutions will help resolve the notorious problem of slowness in the Judiciary. However, they could also eliminate jobs, culminating in “robot-judges” who would replace flesh-and-blood magistrates for issuing sentences and resolving conflicts.

For judges presiding over Brazil’s courts, AI represents an opportunity to relieve the courts and permit more agile application of justice. Such a prospect always conjures up the image of a machine rendering decisions in the place of professionals, but this is a mistaken notion. Changes will in fact occur, but there are many misconceptions as to how AI can be used in providing judicial services.

The first studies on AI in the 1950s explored topics such as problem-solving and symbolic methods. In the following decade, the U.S. Department

<sup>16</sup> Lawyer and civil engineer, PhD in law and informatics from the University of Montpellier (France), partner in the firm Martins Castro Monteiro Advogados, and board member of the Brazilian Internet Steering Committee (CGI.br).



of Defense started training computers to imitate human thinking. By formalizing mental processes through the use of models based on decision trees, these studies paved the way for intelligent automation.

Since then, due to increased computational capacity – in terms of processing speed and memory – and lower costs, AI has acquired exceptional potential, particularly through the use of machine learning. Furthermore, AI can evolve through progressive learning algorithms, allowing the data itself to reprogram the system (deep learning). This new processing standard makes it possible to analyze growing volumes of data in ever greater depth – a task of such complexity that it would be difficult for human beings to perform.

Therefore, AI has great potential for use in various segments of human activity, such as industry, health, transportation and the judicial system. In relation to the last sector, it should be asked: How should AI techniques be incorporated into the provision of judicial services? And how can its use be guided by the principles of justice, understood here as the translation of social values concerning that which is correct, fair and adequate?

## Judicial decisions rendered by machines

The idea that machines can issue judicial decisions has existed for decades – it is not, therefore, something new. The big question is: How to introduce the premises and knowledge that are part of the theoretical training of judges into AI systems?

Among agents in the Judiciary, magistrates are responsible for making decisions that are just and in accordance with the law. It is their duty to resolve real conflicts brought before them and consistently observe the ethical values and legal principles underlying their jurisdictional activities, such as transparency, legality, the right to be heard, and full defense. Their work must be based, not only on legal doctrine, but also on a range of knowledge, consolidated over centuries of experience, which studies and considers the judicial phenomenon according to different aspects – referred to as “legal sciences.”

At the end of the 1980s, at the University of Montpellier, in France, I was able to learn about initiatives to build decision-supporting systems, considered the initial stage of AI in the judicial realm. The greatest challenges and limitations at the time stemmed from the difficulty of processing natural language and semantic networks, in addition to limited capacity for storing and processing data.

The systems arose from decision trees, which presupposed previous knowledge of the legal rules applicable to a specific topic and identification of all the factual hypotheses liable to occur in the real world. When confronted with the actual facts of the case under analysis, the system was supposed to indicate the most appropriate decision according to the rules it was taught.

These programs were used occasionally and were adopted for topics with clearly defined rules and hypotheses. Even today, however, this type of scenario

How should AI techniques be incorporated into the provision of judicial services? And how can its use be guided by the principles of justice, understood here as the translation of social values concerning that which is correct, fair and adequate?

In the first batch, the system processed almost 70,000 cases in 15 days – work that would have taken a team of 11 employees 18 months to complete – with an accuracy rate exceeding 96%.

is associated with the use of AI systems by the judicial system. When taking into account the evolution in capability to process information and the tools available for dealing with the processing of natural language, the subject takes on a new form.

### Uses of AI in the judicial realm

Six decades after the first studies on the use of Artificial Intelligence, it is possible to consider achieving the potential of AI use in the judicial realm. In recent years, we saw the emergence of numerous lawtechs, companies that adopt software and information technologies to boost the productivity of judicial services and that are more concrete practices than the aforementioned “robot-judge” dream.

A recent survey conducted by experts at the University of Oxford indicated that 80% of the interviewees used lawtech products for document management and 43% for document automation/matter workflow<sup>17</sup>. According to the study, the use of systems that incorporate AI is still in the early stages. Of the interviewees who used these tools, 27% did so for document research, 16% for due diligence, and 12% for seeking proof and evidence.

There is a wide range of possibilities for using AI in the Judiciary, but the experiences implemented have been few and limited in terms of the resources they offer. This is becoming more sophisticated through the use of machine learning, wherein the software algorithm, when analyzing judicial decisions, identifies the facts, rules and language applied in them, recognizing the way in which such decisions are built. However, the use of this information also depends on the interpretation made by humans.

There are already examples in Brazil, such as Elis software, from the Justice Court of Pernambuco, which analyses the first stage of tax execution procedures, carries out screening, identifies the parties, verifies whether the statute of limitations has elapsed (i.e., loss of right to action due to expiry of term), and prepares the citation of the debtor. In the first batch, the system processed almost 70,000 cases in 15 days – work that would have taken a team of 11 employees 18 months to complete – with an accuracy rate exceeding 96%<sup>18</sup>.

The Federal Supreme Court also uses the Victor system, developed in partnership with the University of Brasília (UnB). This tool analyzes whether an appeal sent to that court deals with a topic already excluded from the court's assessment due to noncompliance with the requirement of general repercussion, i.e., because it does not represent a collective benefit, whether in legal, economic, political or social terms. In practice, a task that used to take, on average, 44 minutes of an employee's time is now done in only five seconds<sup>19</sup>. According to the court, the system only takes care of 23 out of 1,065 possible matters or theses, but there are already other topics under training. It is reasonable to predict tremendous growth in its use and gains in productivity.

<sup>17</sup> The study *Lawtech Adoption and Training: Findings from a Survey of Solicitors in England and Wales* is available at: [www.law.ox.ac.uk/sites/files/oxlaw/oxford\\_lawtech\\_adoption\\_and\\_training\\_survey\\_report\\_18\\_march\\_2.pdf](http://www.law.ox.ac.uk/sites/files/oxlaw/oxford_lawtech_adoption_and_training_survey_report_18_march_2.pdf)

<sup>18</sup> Find out more: <https://blog.juriscorrespondente.com.br/inteligencia-artificial-no-tipe-agiliza-processos-de-execucao-fiscal>

<sup>19</sup> Find out more: <http://www.stf.jus.br/portal/cms/verNoticiaDetalhe.asp?idConteudo=380038>

On the part of law firms, lawtech activities that use AI are predominantly based on its use to search large volumes of documents, through gathering procedural information and decisions handed down by court and even by judge. Analyzing this significant amount of data enables lawyers and parties to foresee decisions and manage risks. This is referred to as “predictive justice,” which permits choosing the most appropriate argument and obtaining the best proofs to support a lawsuit.

The aforementioned examples, as well as many others in use and under development, attest to enormous gains in productivity, especially in relation to repetitive tasks or those that involve consulting immense databases. Merely bureaucratic activities or those that require much time and attention on the part of experts can be effectively delegated to machines that will accurately and correctly perform them at any time of day. Apart from these applications, other possible uses of AI by the judicial system include:

- Carrying out notary public routines that do not need to be examined by a judge, such as citations, subpoenas and attachments, which eliminates the time when a case is waiting to be processed by a clerk;
- Locating and freezing assets of debtors, known to be the most critical point in enforcing judicial decisions;
- Detecting procedural fraud (flawed distribution, vexatious litigants, embezzlement of assets by debtors, etc.);
- Providing support to the decisions of judges based on the use of knowledge databases proposing solutions backed by precedents;
- Rendering decisions in predictable cases that can be resolved through simple validation of premises established by law;
- Controlling the enforcement of criminal penalties and conditions for reduction of prison sentences;
- In electoral cases, for registering candidates and the accountability of candidates, where the system is more capable of detecting fraud than an analyst;
- In social security cases, for analyzing the concession and review of benefits.

## Ethical parameters for using AI in the judicial system

Important investigations must be done regarding the use of AI in the judicial system. There is no point to a system being able to expedite the outcome of lawsuits, if it does not ensure full defense, the right to be heard, freedom of judgment for judges, and a system with lower courts and appeal courts, i.e., basic principles of procedural law, controlled and made effective by the judge responsible for the case. This gives rise to concerns about ensuring that AI systems will incorporate these premises.

In this regard, important guidelines were adopted by the European Commission for the Efficiency of Justice, in October 2018, in its approval

Merely bureaucratic activities or those that require much time and attention on the part of experts can be effectively delegated to machines that will accurately and correctly perform them at any time of day.

of the Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment<sup>20</sup>. The document stipulates that the use of such tools by judicial systems must respect:

- The fundamental rights of citizens;
- The principle of non-discrimination, specifically preventing the development or intensification of any discrimination toward individuals or groups of individuals;
- The principle of quality and security, so that judicial decisions use certified sources and that the models adopted be conceived in a multidisciplinary manner, in a secure technological environment;
- The principles of transparency, impartiality and fairness: making data processing methods accessible and understandable, authorizing external audits.

Even though AI is heralded as a breakthrough for dealing with the slowness of the judicial system, it is clear that celerity is not the only premise for good provision of judicial services. AI systems need to be built carefully, and the participation of multidisciplinary teams composed of trained professionals is essential, so that the tools can incorporate the aforementioned principles to the utmost, in addition to the knowledge consolidated over centuries of studies and through the practice of legal sciences. Above all, even if magistrates come to depend on the assistance of such systems, they must never delegate their constitutional duties to machines.

## Interview II

***I.S.O.\_ Artificial Intelligence (AI) solutions are being increasingly used for making decisions in different sectors of society. What are the main benefits from using them in the judicial system?***

***I.F.\_*** First of all, we need to differentiate between AI and automation. The Judiciary Branch already uses many tools associated with automation, such as citations and subpoenas that occur automatically in electronic systems, control over deadlines, etc. Until recently, these tasks were performed by humans and are now done by software, in predetermined situations, without any “intelligent” activity on the part of these tools. Even so, this transfer of tasks to machines frees civil servants to focus on areas where human work is essential.

AI tools have started to be seen as one of the ways to deal with the large number of pending lawsuits in Brazil. The use of AI software, adopted with the necessary precautions, could generate an unprecedented gain in efficiency in the judicial system, in addition to greater consistency in decisions, thereby increasing legal certainty.

<sup>20</sup> Find out more: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>.

**I.S.O.\_ How have algorithms been used by the Brazilian judicial system for making decisions? Is there any coordination at the national level, or do these initiatives vary by court or state?**

**I.F.\_** There are both national and local initiatives for implementing AI solutions in the judicial system. National coordination is carried out by the National Council of Justice (CNJ), which published a handbook on the subject in 2019, entitled *Artificial Intelligence in the Brazilian Judicial System (Inteligência Artificial no Poder Judiciário brasileiro)*. In it, the CNJ outlined the main tools in use and under development, particularly the following:

- Mass case screening (*Triagem de Casos de Grande Massa*), for classifying initial petitions according to pre-established areas (energy, banks, airlines, etc.), thereby facilitating the work of judges' offices. Developed by the Justice Court of Rondônia (TJRO), the system is in the approval stage.
- Intelligent procedural movement (*Movimento Processual Inteligente*), for suggesting, in accordance with the Unified Procedural Tables of the CNJ, which movement judges should adopt in their rulings (legal aid, a procedural ruling or an injunction). To this end, the system uses existing decisions in the Electronic Judicial Process (PJe) as a database. The model, which was also developed by TJRO, is ready to be used in the PJe.
- Magistrate text generator (*Gerador de Texto Magistrado*), which uses a probabilistic model to produce automated suggestions for texts based on similar material in the Court's database. Created by TJRO, it has already been approved for production.
- Elis, an AI-supported solution designed to automate the activities of tax collection proceedings, such as the initial screening stage. The software reduces manual and repetitive activities, thereby expediting lawsuits and reducing costs and congestion rates. Developed by the Justice Court of Pernambuco (TJPE), some of its functionalities have already been implemented.
- Corpus927, developed by the Brazilian National Judicial School for Formation and Development (Enfam) to consolidate in one place the jurisprudence of the Federal Supreme Court (STF) and Superior Court of Justice (STJ), in addition to binding decisions, statements and guidance in relation to Article 927 of the Brazilian Code of Civil Procedure. The project has already been implemented.
- Victor, multifunctional software used by the STF to analyze the general repercussions of extraordinary appeals filed with the Court. To this end, the system "reads" all the appeals and organizes and highlights the main documents of lawsuits, thereby reducing analysis time to a matter of seconds.

**I.S.O.\_ What are the main risks associated with the use of algorithms for making decisions in the judicial realm? What mechanisms could mitigate them?**

**I.F.\_** The use of algorithms has risks that are not evident, particularly arising from the following factors: the use of incomplete databases; the possibility of promoting discrimination, even unintentionally, which occurs when the



**Dr. Isabela Ferrari**

Federal judge, academic coordinator of the New Law Institute, with a master's degree and PhD candidate from the State University of Rio de Janeiro (UERJ), and visiting researcher at Harvard Law School.

"Algorithms designed to classify actions, trained with a database that is not rich enough, can yield improper results, leading to loss of time and efficiency in cases."

databases used to train them reflect biases present in society; and, primarily, opacity in their manner of operation, resulting from certain machine learning techniques that hinder quick identification and correction of problems.

In relation to the first factor, it is worth noting that "an algorithm is as good as the data that feeds it." Inadequate data management can, therefore, lead to imperfect decisions or suggestions. Algorithms designed to classify actions, trained with a database that is not rich enough, can yield improper results, leading to loss of time and efficiency in cases.

The consequences associated with the second type of risk are more serious. It is worth recalling, due to its seriousness and symbolism, the case of Eric Loomis. He was caught in the act of stealing a vehicle, and a pre-trial detention warrant was issued based on the results of recidivism risk analysis software. When he requested access to the reasons that he was found to be a hazard, his request was denied by the lower court and Supreme Court of Wisconsin (USA), which stated that it was a proprietary software and, therefore, it was not necessary to reveal the code. Loomis' defense team appealed to the U.S. Supreme Court, which refused to review the decision, arguing that since it involved novel technology, it was premature to take a position on the issue. The situation is even more troubling in that the software was a private tool leased by the state to assist in the risk analysis of prisoners. Based on an audit of its results, the investigative journalism network ProPublica concluded that the software was racially biased.

The example discussed above is directly associated with the third risk of using algorithms supported by AI techniques for making decisions: the characteristic opacity of such decision-making processes. In these hypotheses, computers "write" the programming itself based on a series of correlations. Factors such as lack of articulation of decision rules, enormous amounts of data used in their training, and the fact that source codes only contain the learning rules of algorithms, and not their form of operation, explain the difficulty of understanding how these algorithms actually work. This may result in any of the other risks remaining hidden and delay their discovery, making opacity a meta risk.

***I.S.O.\_ Several software used for decision-making by the Judiciary are developed by private companies. What are the possible implications of this fact in relation to governments and individuals? How can this be circumvented?***

***I.F.\_*** There are difficulties, all of which can be circumvented through the adoption of good practices. The keywords should be: understandability and security. The first refers to the need for public authorities to understand the fundamental aspects of the operation of these algorithms. This includes knowing, for example, the data used to train them, the AI techniques

employed – which may be more or less opaque – and the strategies adopted by developers to impart controllability to the operation of the software. Once public authorities start setting requirements in this regard, the market will naturally begin to operate on the basis of these premises. And since access to source codes is not necessary to achieve this, there will be no issues regarding intellectual property.

As for security, those responsible for contracts or agreements must be concerned about preventing other uses of public data to which companies gain access, even if indirectly. This type of concern – essential in an economy where data is the “new oil” – is currently almost non-existent. These agreements also need to be easily accessible by the public.

***I.S.O.\_ Is there a consensus as to which types of decisions can be delegated to Artificial Intelligence solutions? Will appeals against decisions made on the basis of algorithms be subject to review by the Judiciary?***

***I.F.\_*** I consider the use of technology in the judicial system, including AI tools, to be productive. There is clearly a trend toward the adoption of online Courts in countries with different legal systems, driven by distinct ideological schools of thought. There appears to be a consensus that the use of technology can help us deal with the demand for adjudication in a way that is more efficient, more accessible and less litigious than traditional channels.

These initiatives generally start with a stage that automates activities, after which, AI is inserted to assist human agents – for example, classifying lawsuits, indicating jurisprudence, and even drafting decisions based on the databases used in training the algorithms. Applications in the field of criminal law are more controversial, but the underlying responsibility of judges to examine decisions (which did not occur in the Loomis case) is usually emphasized by the advocates of these measures.

Looking ahead, judicial leaders around the world are already discussing what types of decisions can be definitively delegated to algorithms, without human intervention or often without any possibility of appeal. The response of the majority indicates that adopting this approach would be appropriate to deal with pecuniary issues of small value, where maintaining access to the judicial system would not be justified in economic terms, and where there would be no increased risk to the values held dear by society, such as life and freedom. It is possible to imagine, however, that this would be only the first of a number of future steps, on a path of change accelerated even further by the new coronavirus pandemic. More than ever, it is important for civil society to be attentive and demand a voice in the discussions related to these transformations before it is too late.

"More than ever, it is important for civil society to be attentive and demand a voice in the discussions related to these transformations before it is too late."

## Domain Report

### The dynamics of registration of domains in Brazil and the world

The Regional Center for Studies on the Development of the Information Society (Cetic.br) carries out monthly monitoring of the number of domain names registered in the 15 largest country code top-level domains (ccTLD) in the world. Combined, they exceed 100 million registrations.

In May 2020, domains registered under .tk (Tokelau) reached 31.31 million, followed by China (.cn), Germany (.de) and the United Kingdom (.uk), with 17 million, 16.43 million and 9.46 million registrations, respectively<sup>21</sup>. Brazil had 4.18 million registrations under .br, occupying 10<sup>th</sup> place on the list. With 2.88 million registrations, Canada (.ca) ranked 15<sup>th</sup>, as can be seen in Table 1.

**Table 1 – REGISTRATION OF DOMAIN NAMES IN THE WORLD – MAY 2020**

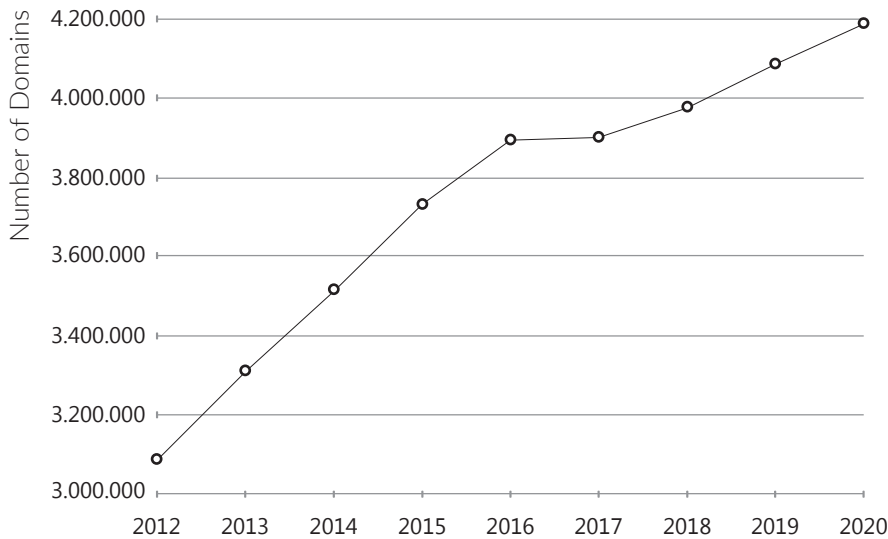
Position	ccTLD	Domains	Source
1	Tokelau (.tk)	31.316.500	research.domaintools.com/statistics/tld-counts
2	China (.cn)	17.002.449	research.domaintools.com/statistics/tld-counts
3	Germany (.de)	16.438.239	www.denic.de
4	United Kingdom (.uk)	9.468.306	www.nominet.uk/news/reports-statistics/uk-register-statistics-2020
5	Netherlands (.nl)	5.960.565	www.sidn.nl
6	Gabon (.ga)	5.396.252	research.domaintools.com/statistics/tld-counts
7	Central Africa (.cf)	5.195.239	research.domaintools.com/statistics/tld-counts
8	Russia (.ru)	4.971.445	www.cctld.ru
9	Mali (.ml)	4.586.703	research.domaintools.com/statistics/tld-counts
<b>10</b>	<b>Brazil (.br)</b>	<b>4.189.546</b>	<b>registro.br/estatisticas.html</b>
11	European Union (.eu)	3.556.502	research.domaintools.com/statistics/tld-counts
12	France (.fr)	3.506.685	www.afnic.fr/en/resources/statistics/detailed-data-on-domain-names
13	Italy (.it)	3.296.543	www.nic.it
14	Australia (.au)	3.170.464	www.auda.org.au
15	Canada (.ca)	2.886.768	www.cira.ca

<sup>21</sup> Table 1 shows the ccTLD domain count according to indicated sources. The numbers shown correspond to the registry published by each country. For countries that do not have or do not publish updated official statistics provided by the domain name registration authority, the total was obtained from "https://research.domaintools.com/statistics/tld-counts/". It is important to note that there is variation between the reference period, although it is always the most up-to-date for each country. In addition, the comparative performance analysis of domain names should take into account the different models of ccTLD registry management. Thus, analysis of the ranking of total registrations under the ccTLDs should consider the diversity of existing business models.



Graph 1 shows the performance of .br since 2012.

**Graph 1 – TOTAL NUMBER OF DOMAIN REGISTRATIONS PER YEAR FOR .BR – 2012 to 2020\***



\*Data in reference to May 2020.

Source: Registro.br

In May 2020, the five generic top-level domains (gTLD) totaled more than 181 million registrations. With 147.75 million registrations, the .com stood out, as shown in Table 2.

**Table 2 – MAIN GTLDS – MAY/2020**

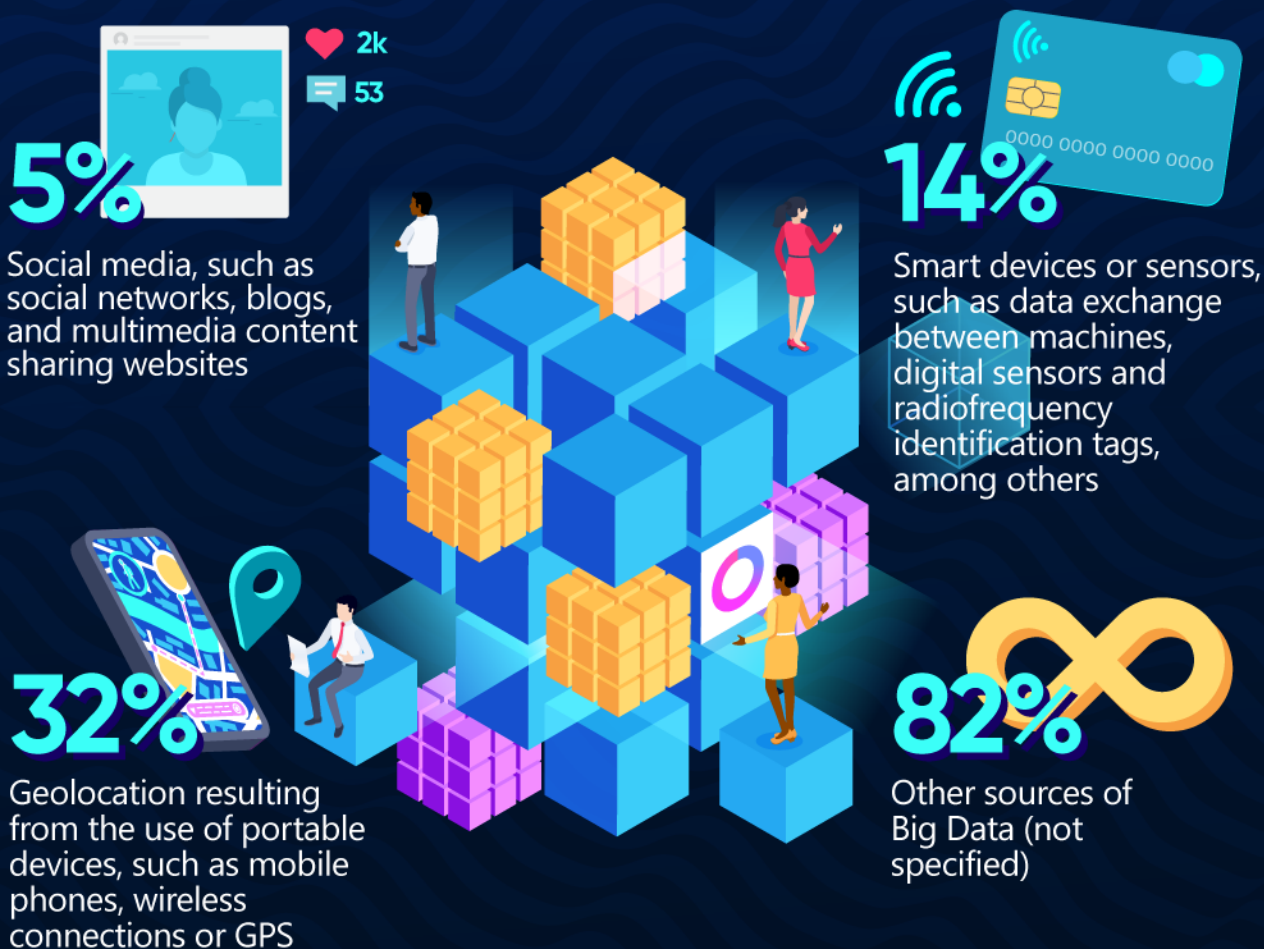
Position	gTLD	Domains
<b>1</b>	.com	147.753.231
<b>2</b>	.net	13.254.709
<b>3</b>	.org	10.146.646
<b>4</b>	.icu	6.215.213
<b>5</b>	.info	4.496.035

Source: DomainTools.com. Retrieved from: <<http://research.domaintools.com/statistics/tld-counts/>>.

# Analysis of Big Data PERFORMED BY THE JUDICIARY<sup>22</sup>

Artificial Intelligence solutions base their learning on the analysis of large databases.

Among the organizations in the **Judiciary Branch** that had areas or departments dedicated to information technology<sup>23</sup>, **26% performed Big Data analyses**. The organizations of the **Judiciary Branch** that performed Big Data analyses used the following data sources<sup>24</sup>:



<sup>22</sup> Based on data from the ICT Electronic Government 2019 survey, by Cetic.br/NIC.br. The reference period for the indicators presented is 12 months prior to the survey. See more at: <https://cetic.br/en/pesquisa/governo-eletronico/indicadores/>

<sup>23</sup> According to the ICT Electronic Government 2019 survey, 98% of the federal and state government organizations of the Judiciary Branch reported having an information technology area or department.

<sup>24</sup> It should be noted that government organizations can use more than one data source to perform Big Data analyses.

# /Credits

## TEXT

### ARTICLE I

Diogo Cortiz  
(Ceweb.br/NIC.br)

### ARTICLE II

Luiz Fernando Martins Castro  
(Martins Castro Monteiro  
Advogados and CGI.br)

### DOMAIN REPORT

José Márcio Martins Júnior  
(Cetic.br/NIC.br)

## GRAPHIC DESIGN AND PUBLISHING

Giuliano Galvez  
(Comunicação/NIC.br)

Klezer Uehara  
(Comunicação/NIC.br)

Maricy Rabelo  
(Comunicação/NIC.br)

## TRANSLATION INTO ENGLISH

Grant Borowik, Lorna Simons,  
Luana Guedes and Luísa Caliri  
(Prioridade Consultoria Ltda.)

## PROOFREADING AND PORTUGUESE REVISION

Mariana Tavares

## EDITORIAL COORDINATION

Alexandre Barbosa  
(Cetic.br/NIC.br)

Tatiana Jereissati  
(Cetic.br/NIC.br)

Javiera F. M. Macaya  
(Cetic.br/NIC.br)

## ACKNOWLEDGMENTS

Diogo Cortiz  
(Ceweb.br/NIC.br)

Nnenna Nwakanma  
(World Wide Web Foundation)

Luiz Fernando Martins Castro  
(Martins Castro Monteiro  
Advogados and CGI.br)

Isabela Ferrari  
(Tribunal Regional Federal – TRF2)



United Nations  
Educational, Scientific and  
Cultural Organization

cetic.br

Regional Center for Studies on the  
Development of the Information  
Society under the auspices of UNESCO

nic.br

Brazilian Network  
Information Center

25 years  
cgi.br

Brazilian Internet  
Steering Committee

CREATIVE COMMONS  
Attribution NonCommercial  
NoDerivs  
(CC BY-NC-ND)



# STRIVING FOR A BETTER INTERNET IN BRAZIL

CGI.BR, MODEL OF MULTISTAKEHOLDER GOVERNANCE

[www.cgi.br](http://www.cgi.br)

[nic.br](http://nic.br) [cgi.br](http://cgi.br)

