Internet Sectoral Overview

Artificial Intelligence in perspective

By Carolina Bigonha[1]

#TechforGood

As the Fourth Industrial Revolution gains momentum, new technologies are shaping the future of production, mobility, trade, information and entertainment, financial systems, health and well-being, education, consumption, the economy and many other areas. Combined with other emerging technologies – such as the Internet of Things, blockchain, advanced robotics and new materials – artificial intelligence (AI) is already a major driver of the digital transformation of organizations, governments, industries and our lives. Smart agriculture, self-driven and connected electric vehicles, natural disaster response solutions, smart cities, and climate prediction and modeling for ocean and forest management are only a few examples of direct applications of AI that can have a positive impact on worldwide challenges.

When combined with traditional statistical methods and calculation of indicators, AI techniques can also contribute to current data generation processes in areas such as accuracy, coverage, cost and speed. In addition to making analysis and detection of patterns more sophisticated, thereby broadening understanding and the capacity for decision-making, the inclusion of data and realities that previously could not be efficiently accessed, collected or comprehended directly contributes to the mission of the 2030 Agenda of the United Nations to "leave no one behind." Examples of AI for measurement purposes include: new opportunities in processing and analyzing satellite images, such as mapping of schools in isolated regions, detection of deforestation in real time, and understanding of urban flows; use of digital media to predict epidemics and social trends, and to predict sanitation and consumption patterns to improve the supply of drinking water and sanitation; understanding of energy consumption patterns and optimization; and many other possibilities.

1   An entrepreneur, data scientist and product expert, Carolina is the co-founder of Hekima, a technology company focused on artificial intelligence solutions. She is a Global Shaper, an initiative created by the World Economic Forum, and co-founder of *Em Perspectiva* (In Perspective), an initiative aimed at promoting a positive, inclusive, ethical and responsible narrative for artificial intelligence in Brazil. [carolina.bigonha@hekima.com]

As leaders and scientists around the world build products and projects and design solutions with the potential to directly affect how people live and interact, there is an important discussion taking place on the best use of artificial intelligence for society. Making these technologies and applications more ethical and inclusive requires proactive collaboration between data scientists, civil society, policymakers, governments, the private sector, investors and experts. The Fourth Industrial Revolution will also bring major changes to the labor market: New formats and types of jobs will emerge, and a significant portion of current jobs will drastically diminish.

This article conceptualizes artificial intelligence and highlights the main opportunities and necessary precautions so that this revolution can result in a positive transformation of society.

What is Artificial Intelligence?

Artificial intelligence is a field of study that began in the 1950s. Its main objective is studying and building systems capable of exhibiting behaviors normally associated with people, such as learning and solving problems. Some lines of study in artificial intelligence are more focused on reproducing how people reason, whereas others concentrate on understanding and simulating behavior.

Previously restricted to large research centers, artificial intelligence technologies are nowadays inserted in the market, in products that people consume, and in various aspects of their lives. The growing popularity of these technologies is directly linked to the abundance and lower cost of processing infrastructure; advances in algorithms; greater data availability ; provision of these technologies in open source formats; and higher connectivity in today's world.

Increased processing capacity unlocks a key aspect: speed in decision-making – a bottleneck for any type of automation. The more advanced and accessible data processing and storage technologies become, the more powerful artificial intelligence systems are, simply because they are able to make more complex decisions in a timely manner. Imagine the computational power required for a self-driven vehicle to decide, in milliseconds, whether it should turn right or left because of an obstacle in its path, or to determine its current speed or identify other vehicles on the road, to give just a few examples.

If algorithms are the engine, then data is certainly the fuel of this technological revolution. What is striking nowadays is the growing volume of data produced and the information available. This diversity includes different types of information – about people, organizations, governments, transactions, behavior, events – and even different methods for data collection: from the Web, images, sound, light, movement, videos, acceleration, gravity and temperature sensors, and many others. In this abundance resides all the potential of artificial intelligence, especially in one of its highly used subsets of techniques, called machine learning.

If algorithms are the engine, then data is certainly the fuel of this technological revolution.

Machine learning

As a subset of artificial intelligence, machine learning techniques are also not so recent. They were created in the 1980s. They are programs capable of learning to perform tasks, not on the basis of explicit instructions, as in traditional programming, but through experience. The higher the quantity, quality and diversity of data – experiences – available, the more complex the tasks learned and executed by these algorithms can be.

There are many types of machine learning systems and different ways of categorizing them: whether there is a need for human interference – supervised, unsupervised, semi-supervised learning algorithms, reinforcement learning; whether the algorithm can learn in real time – batch or online learning; whether it understands patterns from training data and creates a predictive model, or simply compares new data with known data – instance-based or model-based learning; and even whether it is statistical or neural learning – such as linear regression or deep learning.

Tools for implementing machine learning systems are increasingly available, which contributes to the popularity of AI. Over the course of time, major technology providers, and even open source communities, have provided a rich range of libraries and useful tools for free and open building of these systems, such as TensorFlow[1], PyTorch[2] and scikit-learn[3], in addition to a wide range of available commercial platforms. This facilitated access to tools has disseminated the creation of AI solutions; therefore, in a couple of hours and with few resources, systems can now be created that recognize whether a photo contains a person or a dog, for example.

What makes machine learning, and consequently artificial intelligence, so interesting is the diversity of challenges that this set of techniques is able to address. They are highly powerful techniques that operate very well in complex scenarios where contexts change very dynamically, or automation requires long, detailed (and, at times, impractical) lists of rules, or there is a large volume of data that cannot be treated with traditional programming.

What makes machine learning, and consequently artificial intelligence, so interesting is the diversity of challenges that this set of techniques is able to address.

[2] To find out more, go to <https://www.tensorflow.org/>
[3] To find out more, go to <https://pytorch.org/>
[4] To find out more, go to <http://scikit-learn.org/stable/>

Human-centered Artificial Intelligence

In a couple of hours and with few resources, systems can now be created that recognize whether a photo contains a person or a dog, for example.

Building a good artificial intelligence system involves much more than the technology itself. Modeling the challenge, understanding which type of approach is more or less appropriate for each situation, observing and carefully assessing the raw material – the data – and making the best decision given the restrictions are only a few of the considerations for building artificial intelligence solutions.
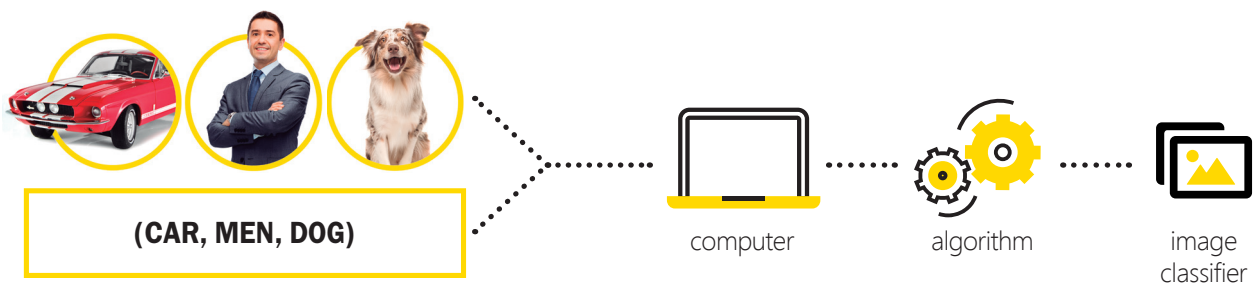
The insertion of algorithms into important decision-making processes, such as personnel recruitment, credit approval assessments and medical diagnoses, among others, leads to reflection on the quality of such decisions. These questions include challenges related to transparency, inclusion, privacy and even accountability. Who is to blame for a bad decision by an algorithm? What causes one model to choose one person or another?

The fundamental point of this new automated decision-making paradigm is: Technology itself is only one of the various active elements. Fei-Fei Li (2018), a computer science professor at Stanford who is a co-founder of AI4All and a board member of Google AI, recently wrote an article pointing out that , despite its name, there is nothing "artificial" about AI, since it is made by humans, intended to behave like humans, and affects humans. So if the goal is for this technology to play a positive role in society, the conversation must also be guided by human concerns. In this context, this author proposes the concept of human-centered artificial intelligence.
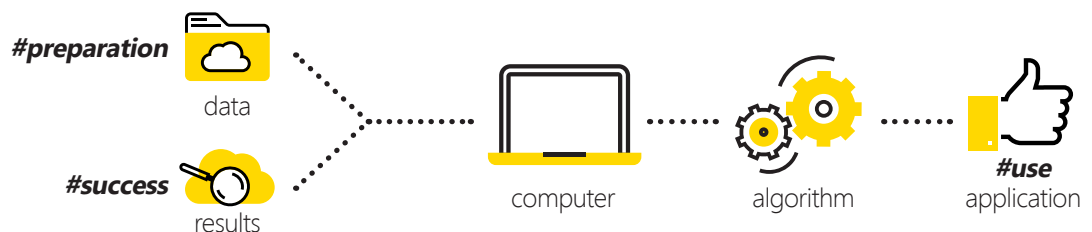
In the creation of machine learning models – one of the most common ways to implement AI – people directly intervene at three points: data selection and preparation; solution design and definition of success; and intended use of the system. For each of these points of interference, there is a number of precautions that need to be taken, so that the AI built will be safe, fair and inclusive.

SIMPLIFIED SCHEME OF A MACHINE LEARNING SYSTEM FOR IMAGE CLASSIFICATION
Data and results serve as input; an algorithm is created for a specific use.

(CAR, MEN, DOG) → computer → algorithm → image classifier

THERE ARE THREE IMPORTANT POINTS OF HUMAN INTERVENTION IN THE PROCESS

#preparation data
#success results → computer → algorithm → *#use* application

1. Data selection and preparation

The first point deals with the definition and treatment of the data set used as training for creation of the model.

The challenge of bias. When training a machine learning model, it is very important for the data set to be representative and complete in relation to the application objective. Artificial intelligence systems have a direct relationship with the data they receive as input, and there can be major differences in the quality of outcomes, depending on who designed the solution and the attributes considered by the model.

A recent study by Buolamwini and Gebru (2018) indicated that there are substantial differences in facial recognition accuracy, depending on race and gender. The authors pointed out that AI systems have a degree of intelligence that is permitted by the data, showing that one of the bases of facial recognition widely used by various algorithms contains approximately 75% images of men and more than 80% images of white people. Since the algorithm has few examples of black women, its accuracy in recognizing them is much lower.

Social inequality, historical baggage and even cultural and geographic differences can give rise to challenges involving representativeness and lead to the reproduction of stereotypes. Algorithms are a reflection of human behavior, and certain biases, even when not explicitly expressed, can influence a system's behavior. A common example is implicit associations between female names and family, and between male names and career, as shown in a study by Caliskan, Bryson and Narayanan (2017).

If not carefully observed, bias of training bases can result in biased machine learning models. When it comes to people, biases related to sex, race and other sociodemographic aspects can occur, as in the case of an application for predicting future criminals, created in the United States and notoriously biased in relation to black people[5].

The Facets[6] tool exemplifies the investigation process that needs to be conducted by data scientists to build effective machine learning systems. The creators of the tool stress that good professionals must go beyond algorithm training. They must be detectives, in order to better understand their models and the databases that feed them, and even monitor the results in production.

Privacy as a priority. Artificial intelligence algorithms, by definition, require a large volume of data as input. The constant search for more customization and higher quality in the models' results makes companies increasingly concerned about collecting and storing any and all information that could be useful and add more value to their services and platforms. Most of the time, this value is derived from companies having greater knowledge of consumers, who often provide their personal data on different digital platforms.

Personal data includes information that indicates digital identity, content and behavior on social networks, browsing history, such as cookies or geolocation information, and even information resulting from analysis, inference and segmentation. The sophistication of AI techniques and the increasing accuracy with which individuals can be analyzed and even identified has given rise to various discussions of the topic. The Cambridge Analytica scandal[7] is an example of the power of this data.

Social inequality, historical baggage and even cultural and geographic differences can give rise to challenges involving representativeness and lead to the reproduction of stereotypes. Algorithms are a reflection of human behavior, and certain biases, even when not explicitly expressed, can influence a system's behavior.

[5] To find out more, go to
<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

[6] To find out more, go to
<https://pair-code.github.io/what-if-tool/>

[7] To find out more, go to
<https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm>

> Users, not companies, should have control over use of data, as any data solution that is created must allow for informed consent that is clear enough so users truly understand how, when and for what purposes their data will be used, in a granular way, allowing for different levels of consent, and dynamic, so that authorization of use can be revoked by users.

Laws recently went into effect in Europe that regulate the responsible use of third-party data – called the General Data Protection Regulation (GDPR)[8]. Brazil is also structuring a similar set of laws[9]. As pointed out in a report from the World Economic Forum (2017), some users are willing to provide elements of their personal data in exchange for added value in services and platforms. Consent is the main aspect to be considered during a data solution project, in addition to being the focal point of laws in force in Europe and under development in Brazil.

Users, not companies, should have control over use of data, as any data solution that is created must allow for informed consent that is clear enough so users truly understand how, when and for what purposes their data will be used, in a granular way, allowing for different levels of consent, and dynamic, so that authorization of use can be revoked by users.

Also, from a technical point of view, there are ways to ensure adding value by customization through the use of data without raising issues of infringement of privacy. One such technique that can be used is federated learning, which consists of an algorithm that learns locally, without data leaving the user's device.

2. Definition of success and designing the solution

The second point covers the selection of metrics to assess performance of models and defining what constitutes a good decision by a system.

Defining success. Another important point of human influence in the building of machine learning models is the definition of success metrics, i.e., the metrics by which the performance of models will be evaluated. Accuracy is not always the best indicator of the performance of a specific algorithm, particularly in databases that have bias. It is important to choose a metric that considers the human and social factors involved in the process.

Consider, for example, a model that approves credit and predicts a person's rating, given the likelihood they will repay the loan. How is success of the system defined? More profit? More people receiving loans? A higher number of payers? A study by Wattenberg, Viégas and Hardt (2016) showed that there are various ways to treat the same problem and that applying different metrics can help reduce inequality when using artificial intelligence.

Attacks. Like any other technology, artificial intelligence models are also subject to attacks. Adversarial attacks are minor modifications to instances, done in such a way as to not hinder perception by a human being, but sufficient to mislead the algorithm. An example of an adversarial attack would be making small modifications to images in order to confuse AI algorithms and force an outcome of interest, such as a small sticker on a stop sign that can lead to a reading error by the image classifier, for example, a speed limit sign.

---

[8] General Data Protection Regulation is a regulation in EU law on data protection and privacy applicable to all individual citizens of the European Union and European Economic Area. It also regulates the export of personal data outside the EU and EEA.

[9] To find out more, go to <http://www.camara.gov.br/proposicoesWeb/fichadetramitacao?idProposi cao=548066>

EXAMPLE OF AN ADVERSARIAL ATTACK ON AN IMAGE: A SMALL YELLOW SQUARE CONFUSES AN ALGORITHM FOR INTERPRETING TRAFFIC SIGNS

When putting a machine learning system into production, data scientists must consider the sensitivity of automated decision-making. They must also apply safety and protection practices for potential side effects of the technology and, in appropriate cases, create test conditions in simulated environments and, after integration of the model, monitor its operation.

Trust and transparency. How do you know whether you can trust an automatic decision? Is the algorithm doing what you expect it to? And even if the result was what you expected, how do you verify if the system was correct in the evaluation process?

Trust is a key factor in the interaction between people and machines and, most of the time, reflection involves understanding what takes place behind automation. In simpler algorithms, such as decision trees, it is easy to understand the operational rationale of the model, since it entails a sequence of rules. However, for more sophisticated algorithms, such as deep learning, the model can be quite complex, making it difficult to understand the decision made by the algorithm.

Through techniques such as LIME[10] and SHAP, it is possible to develop explanations for classifications, even for algorithms that are originally black box. This helps the decision-maker to understand the motivation of the algorithm behind a given classification.

Increasing the transparency of algorithms, apart from helping understand and validate the results of AI predictions, can lead to the discovery of new facts. After achieving success in a deep learning application for diagnosing diabetic retinopathy[12], Google researchers[13] decided to open the black box of the algorithm. Applying attention techniques to understand the rationale used by the predictive algorithm to make the diagnosis, the team observed patterns linked to blood vessels

---

[10]   Local Interpretable Model – Agnostic Explanations. To find out more, go to <https://arxiv.org/pdf/1602.04938.pdf>

[11]   Shapley Additive explanations. To find out more, go to <https://arxiv.org/pdf/1705.07874.pdf>

[12]   To find out more, go to <https://ai.googleblog.com/2016/11/deep-learning-for-detection-of-diabetic.html>

[13]   To find out more, go to <https://ai.googleblog.com/2018/02/assessing-cardiovascular-risk-factors.html>

and identified that eye imaging could also predict other indications of cardiovascular disease. This is a new way of making scientific discoveries and can help scientists generate more targeted hypotheses and spur a wide range of future studies.

The implications of moral decision-making by machines are also part of the discussion. When critical decisions are made tangible in algorithms and programming, moral dilemmas quickly arise. An example of this challenge is the different decision-making situations of self-driving vehicles: When faced with an impossible decision, where one or more lives are at risk, what is the best technical decision? The Moral Machine experiment (Awad, 2017) is one of the studies that were conducted on the subject, which attempts to understand the human perspective in relation to moral decisions made by machines.

A widely used strategy is to build solutions so that the AI system is an informed decision-making tool, not the decision-maker itself. In artificial intelligence applications for health, explaining how the algorithm makes predictions gives physicians the necessary information and confidence to make the best decision. Even though technology has made the process faster and more accurate, the medical specialist still has the final word.

3. Intended use

Finally, the intended use of artificial intelligence systems is another question that needs to be posed. The same technology can have applications that are beneficial or not, depending on the intention of the person operating it. Image recognition, for example, is a technique that can help people with a visual impairment understand the content of images, or farmers diagnose the quality of planted land, or physicians perform complex diagnoses. However, image recognition has also been used in experiments to infer the sexual orientation of a person based on facial features[14].

Ensuring that the solutions created by people using artificial intelligence will be ethical, beneficial and safe for society is a real challenge, which involves the creation of codes of ethics; balancing forces, such as markets, technological infrastructure, rules and laws; and the participation of multiple stakeholders.

[14]   To find out more, go to <https://www.nytimes.com/2017/10/09/science/stanford-sexual-orientation-study.html>

The future we want with artificial intelligence

The more opportunities for artificial intelligence multiply, the greater the need to engage industry leaders, researchers, experts, governing bodies and civil society in reflecting and taking action in favor of solutions that benefit society as a whole. Ethical principles and technical standards will help ensure that the design and development of such technologies are guided by the concern for their effect on people. However, there are also other necessary perspectives and initiatives.

In light of the impact of automation on the labor market, and also to ensure the active participation of civil society in the advancement of these technologies, equally intense efforts in education will be important, particularly in relation to STEM (science, technology, engineering and mathematics), as well as training in intrinsically human skills, such as creativity, resilience, flexibility and solving complex problems, among others. Therefore, governments, companies and universities will play a fundamental role in this process.

Artificial intelligence is here to stay and to transform markets, society and people's lives. As diversified as the audience is that will be impacted by this technology, so should the people involved in its construction be, with different experiences, paths and training. Our future with AI must be the result of a connection between technology and human sciences, technology and specialties, and technology and society.

As diversified as the audience is that will be impacted by this technology, so should the people involved in its construction be, with different experiences, paths and training. Our future with AI must be the result of a connection between technology and human sciences, technology and specialties, and technology and society.

REFERENCES

Awad, Edmon. (2017). *Moral machine: Perception of moral judgment made by machines* (Master's thesis). Retrieved from: https://www.media.mit.edu/publications/moral-machine-perception-of-moral-judgment-made-by-machines/

Buolamwini, J. & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research, 81*, 1-15. Retrieved from: http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18 a.pdf

Caliskan, A., Bryson, J. J. & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science, 356*(6334), 183-186. Retrieved from: http://science.sciencemag.org/content/356/6334/183

Li, Fei-Fei. (2018, March 7). How to make A.I. that's good for people. *The New York Times*. Retrieved from: https://www.nytimes.com/2018/03/07/opinion/artificial-intelligence-human.html

Wattenberg, M., Viégas, F. & Hardt, M. (2016). Attacking discrimination with smarter machine learning. Retrieved from: https://research.google.com/bigpicture/attacking-discrimination-in-ml/

World Economic Forum. (2017). Shaping the future implications of digital media for society valuing personal data and rebuilding trust. Retrieved from: http://www3.weforum.org/docs/WEF_End_User_Perspective_on_Digital_Media_Survey_Summary_2017.pdf

Governance Challenges in Artificial Intelligence

Sara Rendtorff-Smith is the Applied Research Lead for Data-Driven Governance and AI Policy at the Massachusetts Institute of Technology (MIT). In this interview, she comments on artificial intelligence opportunities as well as some of the problems associated with their use and how to deal with them.

Sara Rendtorff-Smith
is a researcher at the Massachusetts Institute of Technology (MIT).

I.S.O._ How can we deal with mistakes, bias and unintended results related to artificial intelligence?

S.R.S._ First of all, it is important to consider two things: One, how mistakes, bias and unintended results occur in the first place; and two, how we might audit for these and investigate how errors were made after the fact. It's important to note that these issues are located squarely at the intersection of human-machine interaction. In this way, mistakes and bias are often introduced by humans as they select data, make design choices and trade off across costs borne by different stakeholders.

A useful example for exploring these different concepts is autonomous vehicles. Engineers will program the vehicle to trade off risk between the driver and anyone the vehicle might risk colliding with. The visual recognition technology may also inadvertently be trained in such a way that it performs better on certain demographics than others. This has been an unintended consequence where the tech sector has used demographically skewed data, e.g., where women and people of color were underrepresented. This, in turn, has led to poor accuracy of image recognition for these populations. These issues are often raised in the context of placing legal liability, but also matter for designing frameworks for preventing and mitigating mistakes, bias and unintended results.

In the case of true errors and unintended consequences, on the other hand, it becomes important that AI systems be auditable in a manner that allows investigators to identify the source of an error or adverse consequence. This is part of a much broader conversation about black box AI systems and their place in society.

I.S.O._ What is the role of the different stakeholders for building ethical AI solutions? How should end users be engaged in this discussion?

S.R.S._ AI sits at the heart of sociotechnical relations and it therefore requires the active and informed involvement of multiple stakeholders to build ethical

AI solutions. Academics and developers should be trained to be aware of their own responsibility related to the development of AI systems with direct impacts on society. Governments and citizens, for their part, have a critical responsibility for determining what values should govern the deployment of AI systems in the public sphere and how issues of liability should be regulated. End users or individuals and communities directly impacted by a particular AI technology can be involved at three critical stages of AI development by participating in the design, testing and auditing of new AI systems.

I.S.O._How can AI be used for evidence-based policymaking? Please provide examples. What are the challenges for this?

S.R.S._ One of the most prominent policy areas where AI has been leveraged to promote data-driven, or so-called "evidence-based," policymaking and decision-making is the criminal justice sector. In the U.S., more than 60 police departments are currently implementing some form of AI-powered predictive policing system. Meanwhile, a growing number of jurisdictions have adopted risk scores generated by AI systems to assist with a number of decision points across the criminal justice chain, including pretrial release, post-conviction sentencing, probation and parole.

Notable examples of the associated challenges include bias in the assessment of the predicted future behavior of defendants and incarcerated persons based on race, and over-policing of minority neighborhoods as a result of historically biased policing practices. The ironic thing is that the former was actually introduced in the first place as a strategy for mitigating human bias in judicial decision-making. These and many other challenges tend to emerge partly as a result of the black box nature of the particular type of AI technology deployed, which has predominantly been machine learning and deep learning systems.

These systems, particularly when we talk about unsupervised learning systems, are by nature obscure, making it difficult to uphold key principles of transparency, explainability and accountability. Another challenge has been that governments often procure and deploy proprietary AI technology developed by private industry actors.

My lab at MIT is currently conducting research into how another form of AI, based on probabilistic computing, can help policymakers identify patterns in data, including dependencies among variables, similarities across localities and communities, and outliers. I am hoping this can help public officials and policymakers learn from each other and build alliances, as well as highlight localities and communities that are experiencing unusually low levels of public service delivery or unusually bad socioeconomic outcomes, in order to promote equity.

I.S.O._There is a lot of discussion about AI accountability. How should this happen? Which stakeholders should be engaged in this process?

S.R.S._ This is indeed a very active and timely discussion that is happening at the moment. If we are talking about AI being deployed in a manner that affects the

general public, or perhaps even about an instance where government is leveraging AI technology as part of its system of governance, the threshold of accountability is naturally higher than where a private enterprise deploys AI for its internal operations.

Other approaches to AI accountability include upstream approaches such as teaching developers and academics about their ethical responsibilities and training them to address these already at the design stage, or midstream approaches such as a code of conduct for ethical innovation or public oversight of predeployment product testing.

Interview II

Ethics, Law and Artificial Intelligence

Eduardo Magrani, a lawyer and coordinator of the Institute for Technology and Society (ITS Rio), comments on the challenges brought about by the development of artificial intelligence, in terms of ethics, legislation and the future of work.

Eduardo Magrani is a coordinator of the Institute for Technology and Society (ITS Rio).

I.S.O._ What are the main ethical and legal challenges that should be considered in the use of AI?

E.M._ There is a cultural gap between how society looks at ethical themes within and outside of the topic of AI. In Brazil, ethics are still not much discussed in schools and universities: Civil and contractual discussions are more common. For example, whether it is a good idea for robots to have a personality is not discussed. Given the newness of the topic, there are few people studying AI academically from a critical perspective, which further hinders production related to the topic. Since Brazil is in the very early stages of discussing ethics and AI, there is not yet much production in this regard. We live in a context where technology is fetishized. We incorporate new technologies into our daily lives without having a critical outlook on them, and only consider them from the point of view of utility, without addressing issues of privacy, security and the ethical impacts of these new technologies.

I.S.O._ Outside of Brazil, what are the most troubling issues regarding the intersection between ethics and AI, in terms of legislation?

E.M._ There is a vigorous discussion on how programmers and software engineers give ethical and moral input to machines. Technology is not neutral, algorithms are not neutral, and there is often a morality intrinsic to that technical artifact, inserted by a programmer. So what values will software engineers be inserting into these technical artifacts, which are also increasingly autonomous and unpredictable? Another discussion involves reparation for harm: If an autonomous robot, programmed by a programmer, causes harm that was not anticipated by the company, who should be held accountable?

I.S.O._ Is it possible to draft laws that protect people's rights without limiting the development of these technologies?

E.M._ When you say it must be a "law guided by human rights," it is not as obvious as it seems, and that is what I am defending. Yet, there are also those who advocate technological advance for the sake of efficiency and not necessarily for the protection of human rights. You still see this a lot in Brazil: A large part of the technological progress that we are currently experiencing is not guided by protection of privacy, security and human dignity, as expressed in our Constitution. This advance is guided merely by efficiency and utility. However, I am advocating that the advance of AI must be driven by a human-centric vision, since we need to move away from the idea of enlightened law, which views human beings as endowed with powers of action. Law must move forward and also consider non-human agents that interact with us and generate democratic, constitutional and civil effects; by recognizing these different human and non-human agents, we will be able to draft better regulations for this hyperconnectivity or AI scenario.

I.S.O._ What are the current challenges in terms of legislation and AI, considering our national context and the international scenario?

E.M._ The most obvious problem is jurisdiction. The GDPR in the European Union is a very strong law. There can only be an international flow of data and information between countries with the same adjustment of privacy rules, which is not the case for Brazil. So how can you control the flow of data between countries with different regulations for privacy, for example? This is a major challenge today, which has to do with the context of the Internet of Things, where increasingly smart things are generating data all the time, which is transferred between companies and even internationally.

Another problem is the lack of strong digital literacy and training on digital issues. This isn't found in our schools nor in universities. Therefore, citizens are completely unaware of many risks and vulnerabilities and, consequently, do not apply political pressure for better laws or for governments to observe what is happening in other countries and try to incorporate them in Brazil. The topic of digital law is still discussed very little here and is only in the fledgling stages. So this lack of adequate regulations gives rise to a very abusive state of affairs.

I.S.O._ Are there best practices or guidelines to follow in terms of ethics and AI?

E.M._ There are currently various initiatives seeking to map best practices, such as the report published by the Institute of Electrical and Electronics Engineers (IEEE)[13] on the principles that should guide AI practices[14]. There is also the 2017 Montreal Declaration for a Responsible Development of Artificial Intelligence[15] and, in Europe, a set of guidelines is being published on this topic, but still nothing yet in Brazil.

---

[13] The IEEE, an association dedicated to the advance of innovation and technological excellence for the benefit of humanity, is the largest technical professional association in the world. It was designed to serve professionals in all aspects of the electric, electronic and computing fields and related areas of science and technology that underlie modern civilization.

[14] The executive summary, in Portuguese, can be viewed at <https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_executive_summary_portuguese_v1.pdf>

[15] Available at <https://www.montrealdeclaration-responsibleai.com/the-declaration>

It is possible to apply international guidelines in different national contexts if guided by human rights. And since these declarations operate on a very macro level, they can be incorporated into different national legal systems.

I.S.O._ *Regarding the future of employment, there is a concern about "machines stealing jobs" through automation and the advance of technology. What is your opinion on this? How can society prepare to adapt to these transformations, leaving no one behind?*

E.M._ We are living in the era of robotics. In very repetitive mechanical jobs, there is no doubt that humans will be replaced. What should human beings do? Adapt and seek to develop new skills that robots are unable to do, in order to keep themselves in the job market. The problem is that in Brazil resources are often lacking to give people this training. We may have a large generational gap among people who did not have the time or resources to adapt to this new technological context, whereas other countries will do so more quickly to avoid this.

I have no doubt that AI, this new era of automation, will also generate new demands for jobs. You can already see the number of areas that work with filtering algorithms and big data analyses to boost their business; this generates new demand, even from customers. The problem is those who lack the time or resources to adapt. Training in digital law and in new technologies is the main path for reducing this gap.

Domain Report

The dynamics of the 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 16 largest country code Top-Level Domains (ccTLDs) in the world. Combined, they exceed 97.8 million registrations.

In October 2018, the domains registered under .tk (Tokelau) reached 21.19 million, followed by Germany (.de), China (.cn) and the United Kingdom (.uk), with 16.21 million, 11.11 million and 9.83 million records, respectively[16]. Brazil reached the mark of 4 million registrations under .br, occupying the seventh place on the list. With 1.93 million registrations, Spain (.es) ranked 16th, as can be seen in Table 1.

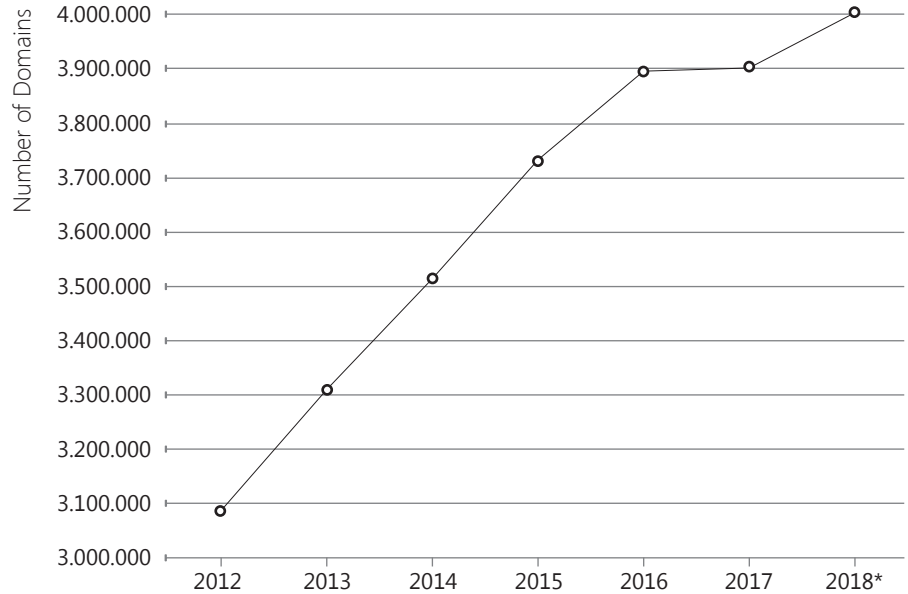**Table 1 – REGISTRATION OF DOMAIN NAMES IN THE WORLD – OCTOBER 2018**

Position	ccTLD	Domains	Ref.	Source
1	Tokelau (.tk)	21.196.458	Oct-18	http://research.domaintools.com/statistics/tld-counts/
2	Germany (.de)	16.211.149	Oct-18	www.denic.de/
3	China (.cn)	11.113.061	Oct-18	http://research.domaintools.com/statistics/tld-counts/
4	United Kingdom (.uk)	9.832.194	Aug-18	https://www.nominet.uk/uk-register-statistics-2018/
5	Netherlands (.nl)	5.815.231	Oct-18	www.sidn.nl
6	Russia (.ru)	5.059.321	Oct-18	www.cctld.ru
7	**Brazil (.br)**	**4.002.574**	**Oct-18**	**registro.br/estatisticas.html**
8	European Union (.eu)	3.661.862	Oct-18	http://research.domaintools.com/statistics/tld-counts/
9	France (.fr)	3.252.404	Oct-18	https://www.afnic.fr/en/resources/statistics/detailed-data-on-domain--names/
10	Australia (.au)	3.162.708	Oct-18	www.auda.org.au
11	Italy (.it)	3.145.839	Oct-18	www.nic.it/
12	Canada (.ca)	2.779.778	Oct-18	www.cira.ca/
13	Poland (.pl)	2.585.723	Oct-18	www.dns.pl/english/zonestats.html
14	Switzerland (.ch)	2.164.079	Jun-18	https://www.nic.ch/reg/cm/wcm-page/statistics/index.html?lid=em*
15	United States (.us)	1.950.307	Oct-18	research.domaintools.com/statistics/tld-counts/
16	Spain (.es)	1.931.707	Oct-18	www.dominios.es

[16] It is important to note that variations exist among ccTLD reference periods, although it is always the most updated one for each country that is used.

Graph 1 shows the performance of .br since 2012.

Graph 1 – TOTAL NUMBER OF DOMAIN REGISTRATIONS PER YEAR FOR .BR – 2012 TO 2018*

Data in reference to October 2018.
Source: Registro.br

In October 2018, the five generic Top-Level Domains (gTLD) totaled more than 169 million registrations. With 137.48 million registrations, the .com ranked first, as shown in Table 2.

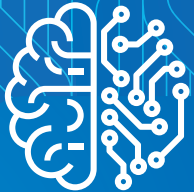**Table 2 – MAIN GTLDS – OCTOBER 2018**

Position	gTLD	Domains	Source
1	.com	137.486.743	research.domaintools.com/statistics/tld-counts/
2	.net	13.985.155	research.domaintools.com/statistics/tld-counts/
3	.org	10.385.742	research.domaintools.com/statistics/tld-counts/
4	.info	5.440.794	research.domaintools.com/statistics/tld-counts/
5	.biz	2.252.993	research.domaintools.com/statistics/tld-counts/

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

A lot of new things?

We have prepared a list with the main terms and concepts associated with artificial intelligence and ethics. Check it out!

Artificial Intelligence

A field of study focused on building systems capable of exhibiting behaviors normally associated with people, such as learning and solving problems. Some lines of study in artificial intelligence are more focused on reproducing how people think and reason, whereas others concentrate on understanding and simulating behavior.

Machine learning

A set of techniques within artificial intelligence, machine learning is the science and art of using computers so that they can learn to perform tasks through experiences (data). It is a field that gives computers the ability to learn without the need for explicit programming.

Accountability

Artificial intelligence systems are subject to unexpected results or flaws that can cause harm. In such cases, forensic ability is essential for ensuring that accidents and similar failures do not occur again and to determine responsibility and accountability. Accountability challenges include: understanding who is responsible for the behavior of an automated system; being able to supervise or audit automated decision-making; and even determining judicial and legal liability and implications.

Transparency

This term is related to how interpretable or explainable an AI system is. Some techniques used are so complex that they work like a black box and leave the automatic decision-making process opaque, even to the person who built the system. Design techniques and standards aimed at transparency help determine how reliable the AI system is, as well as to understand if it will function well and even explain its operational logic. The goal is to have model decisions that are not only good, but also interpretative.

Bias in automated decisions

Algorithms are a reflection of human behavior; even when not explicitly expressed, certain biases found in data and in design decisions can influence the behavior of the system. Therefore, in some cases, algorithm-based decisions can reproduce or reinforce a negative bias, with patterns of discrimination and maintenance of stereotypes, as an inheritance of people's prejudiced decisions or simply because the data reflects cultural, historical and sociodemographic aspects in society.

General Data Protection Regulation (GDPR)

European law regulating the protection of personal data. Its objective is to offer users greater control and transparency in relation to personal information stored in company databases. Individuals must have the right to know what information they are providing to the companies that they use. In addition, an organization must explain why it is requesting certain data from clients and for what purposes it will be used.

Brazilian Data Protection Law No. 13709

This law regulates the use, protection and transfer of personal data in Brazil. It ensures citizens greater control over their personal information. Consent is required for collecting and using data – by the government and private sector – and it stipulates giving users options to view, correct and exclude this data. It also prohibits the treatment of personal data for illegal or abusive discriminatory practices (crossing an individual's or group's information to provide input for commercial decisions, public policies or the activities of government organizations).

Fairness

A concept that encompasses techniques and practices to ensure that AI systems can make fair decisions with the least amount of undesirable bias possible. Automatic decisions related to credit, mortgages, insurance, education, the labor market and others can contain biases and reproduce unjust or discriminatory behavior of people based on individual merit-related attributes. The goal of fairness is to maintain high accuracy in learning algorithms and, at the same time, reduce the degree to which they discriminate against individuals belonging to certain groups.

/Credits

The Internet Sectoral Overview is also available in Portuguese at cetic.br/publicacoes/indice/panoramas/

UNESCO
United Nations
Educational, Scientific and
Cultural Organization

cetic.br
Regional Centre of Studies for the
Development of the Information
Society under the auspices of UNESCO

nic.br
Brazilian Network
Information Center

cgi.br
Brazilian Internet
Steering Committee

STRIVING FOR A BETTER INTERNET IN BRAZIL

CGI.BR, MODEL OF MULTISTAKEHOLDER GOVERNANCE
www.cgi.br

nic.br cgi.br