BALANCING ARTIFICIAL INTELLIGENCE INNOVATION WITH FUNDAMENTAL RIGHTS

Johannesburg, 3.09.2025

I. INITIAL REMARKS

1. Greetings and thanks.

2. I think it's very important that we're here debating the major issues of justice, at a time when the independence and empowerment of the Judiciary has become very important for preserving the rule of law in different countries around the world.

3. I have divided my presentation in three parts: I. The digital revolution and the rise of AI; II. Benefits, risks and the use of AI by Courts; and III. AI and fundamental rights.

Parte I

THE DIGITAL REVOLUTION AND THE RISE OF ARTIFICIAL INTELLIGENCE

I. THE THIRD INDUSTRIAL REVOLUTION

1. We are still under the aegis of the Third Industrial Revolution, that began in the last decades of the past century. It is defined by the shift from analog to digital technology.

2. Known as the Digital Revolution, this transformation paved the way for the widespread adoption of personal computers, smartphones, and the internet, connecting billions of people worldwide.

3. All of us, especially those from earlier generations, have had to develop a new vocabulary. The language of our times includes a set of terms that describe tools and utilities that did not exist until very recently, yet have become indispensable to our daily lives.

⇒ To name a few: Google, Facebook, Instagram, TikTok, YouTube, WhatsApp, Telegram, Twitter, Waze, Spotify, Netflix, Uber, Dropbox, Skype, and FaceTime. For singles, there’s also Tinder.

II. A BRAVE NEW WORLD AND A NEW ECONOMY

1. Innovations and technological advancements are shaping this brave new world of information technology, biotechnology, nanotechnology, robotics[1], 3D printing[2], quantum computing, autonomous vehicles, and the Internet of Things[3]. The concept of the algorithm is becoming one of the most important of our time.

⇒ Thus, a new economy emerges – the economy of knowledge, information, and intellectual property.

2. The world's most valuable companies are no longer those that extract oil – like Shell and Exxon – or manufacture cars, such as General Motors and Ford, or large equipment, like General Electric. Today, the most valuable companies include Amazon, Apple, Facebook, Microsoft, Google, and Nvidia - businesses driven by technology and data.

3. In this environment, major technology companies (the big techs - Apple, Alphabet, Meta, Microsoft) have flourished, alongside ventures focused on innovation and the development of new products and services, such as startups and fintechs (Financial Technology) like Nubank and Nomad.

4. New business models have emerged, encompassing e-commerce (Amazon, Alibaba, Magalu), video platforms (YouTube), streaming services for movies and series (Netflix, HBO, Amazon Prime), music streaming platforms (Spotify, Apple Music), cloud storage services (Dropbox, Google Drive), payment solutions (e.g., PayPal, PagSeguro), shared goods and services platforms (Uber, Airbnb), social media networks (Facebook, Instagram, TikTok), news websites (UOL, Globo.com), among many others.

5. Evidently, this New Economy brings numerous legal challenges and concerns across various areas: constitutional, tax, labor, contractual, intellectual property,

[1] Many sophisticated surgeries today are already performed by robots, although they are still controlled by human hands.
[2] With a photo of your foot, a custom-made shoe can be created.
[3] You'll walk into your home, and the air conditioner will ask if the temperature feels right.

criminal, civil liability, and more. Legislators, judges, and lawyers will certainly not run out of work.

III. ARTIFICIAL INTELLIGENCE AND THE FOURTH INDUSTRIAL REVOLUTION

1. We are now at the dawn of the Fourth Industrial Revolution, with Artificial Intelligence at its core. This revolution also encompasses Biotechnology[4], Genetic Engineering[5], the Internet of Things[6], and the Internet of Senses[7], among other technological advancements.

⇒ We will be focusing particularly on Artificial Intelligence.

2. Artificial Intelligence is comparable to some innovations and inventions that have changed the course of human history, transforming the way we live, extending life expectancy, or making life significantly better and easier.

⇒ A brief list of some of these groundbreaking inventions or discoveries include: fire, wheel, writing, electricity, internal combustion engine and medical advancements (antibiotics, vaccines and anesthesia).

Parte II

ARTIFICIAL INTELLIGENCE: BENEFITS AND RISKS

[4] Biotechnology (Biotech), as the name suggests, is a combination of biology and technology, utilizing living organisms, cells, and plants to develop products such as synthetic insulin, synthetic hormones, biofuels, and many others. Beyond healthcare, it has significant applications in agriculture and various industries.

[5] Genetic Engineering is a branch of Biotechnology that involves manipulating DNA and RNA, the structures that determine the hereditary traits of living organisms, for various purposes. These range from therapeutic cloning to human cloning, which is prohibited in most countries. Additionally, concerns are raised about the risks of its use for eugenic purposes.

[6] The Internet of Things (IoT) is a network connecting objects to the internet for data transmission and reception. It enables the creation of smart homes, where lights turn on when the resident arrives, air conditioning adjusts the temperature, and pantries notify about missing products. It also powers smart cars, which suggest the best routes and warn of road hazards, and smart cities, providing traffic flow updates, security camera monitoring, and efficient waste collection systems.

[7] The Internet of Senses enables online sensory experiences by integrating the five human senses: sight, sound, touch, taste, and smell. Some practical applications include: in medicine, allowing doctors to examine patients by virtually "feeling" them; in e-commerce, enabling consumers to "touch" fabrics online to sense their texture; in virtual tourism, experiencing the scent of a location; and even tasting food virtually.

I. WHAT IS ARTIFICIAL INTELLIGENCE

1. Basic concepts

1. Simply defined, *Artificial Intelligence* consists of softwares that transfer human capabilities to computers. These capabilities involve cognitive tasks and decision-making, usually based on the data, instructions, and goals with which they are provided.

2. It's still common wisdom that Artificial Intelligence lacks self-awareness, discernment of right or wrong, and emotions, feelings, morality, or even common sense. In other words, it is entirely dependent on human intelligence to guide it, including ethical values.

3. This idea starts to be challenged by some authors, but we are not going down this path right now.

2. Types of Artificial Intelligence

This general label Artificial Intelligence encompasses diverse fields of expertise and applications, including:

1. *Machine learning*
2. *Natural language processing*
3. *Computer vision*
4. *Robotics*
5. *Affective computing*
6. *Generative Artificial Intelligence:* The ability to create original content, such as images, texts, music, and art in general. It demonstrates creativity across various domains, from science to literature.

⇒ Its most well-known commercial version is ChatGPT, but it also includes competitors like Claude, DeepSeek, Perplexity, Copilot, Llama, Notebook and others.

7. This is the latest frontier in Artificial Intelligence, stunning the world, and is also referred to as a Large Language Model.

II. Benefits of Artificial Intelligence

1. The current applications and potential of Artificial Intelligence are so vast that organizing them is no easy task. Here are a few significant examples:

⇒ 1. Improved decision-making capacity in many areas (Google); 2. Automation; 3. Language; 4. Research and innovation; 5. Applications in medicine; 6. Applications in the justice system; 7. Education and culture; 8. Other useful applications of AI; a) Practical everyday uses (Uber, Wase, Google Maps); b) Protection of the environment; c) Personalization of business and other relationships.

III. Risks of Artificial Intelligence

1. However, new technologies have a darker side that must be acknowledged and addressed. Digital platforms, driven by an unrestrained quest for engagement, have facilitated extreme polarization, disinformation, hate speech, and the spread of conspiracy theories.

3. The following is an illustrative list of problems and risks associated with AI and its potential exponential growth:

⇒ 1. Impact on the labor market; 2. Use of AI for war purposes; 3. Mass disinformation; 4. Violation of privacy; 5. Algorithmic discrimination; 6. Intellectual property and copyright issues; 7. Risks to reasoning and critical thinking in individuals; 8. Singularity (loss of control by humans).

Parte III

THE NEED FOR REGULATION OF AI, ITS APPLICATIONS IN THE JUSTICE SYSTEM AND CONCERNS WITH FUNDAMENTAL RIGHTS

I. THE NEED FOR REGULATION OF AI

1. Considering everything discussed so far, it is evident that regulating Artificial Intelligence has become essential. However, this task is not straightforward and involves significant challenges and complexities. Below, we aim to outline some of these challenges.

1. *Regulation must be implemented while the train is in motion.* Appeals from scientists for a pause for the establishment of "a set of shared safety protocols" were unsuccessful. There is a major competition among nations, researchers, and entrepreneurs over advancements in this field and nobody wants to pause.

2. *The pace of these transformations is astonishing.* This makes it extremely challenging to foresee what lies ahead and to incorporate new realities into legal frameworks, which risk becoming obsolete in a short time. Illustrating this point is not difficult.

The traditional landline took 75 years to reach 100 million users. The mobile phone took 16 years. The Internet took 7 years. Meanwhile, ChatGPT reached 100 million users in just two months[8]. It is clear that legislation and regulation struggle to keep pace with the rapid rate of innovation.

3. *Risks of excessive regulation.* While regulation has become essential, as mentioned earlier, it also brings its own risks. One key concern is that restrictions and civil liability should not be so severe as to stifle innovation.

4. *Information and power asymmetry between companies and regulators.* AI technology is largely controlled by the companies developing it, which possess expertise that exceeds that of potential regulators.

[8] The Feed. ChatGPT witnesses massive rise, chatbot gains 100 million users in two months. *Economic Times*, 5 mar. 2023. Available at: https://economictimes.indiatimes.com/news/new-updates/chatgpt-witnesses-massive-rise-chatbot-gains-100-million-users-in-two-months/articleshow/98428443.cms?from=mdr. Accessed on March 31, 2024.

5. *The need for global harmonization of regulation*. AI is a predominantly private technology that transcends national borders. This involves the challenges of enforcing fundamental rights against private companies as well as of extraterritoriality.

I. APPLICATIONS OF AI IN THE JUSTICE SYSTEM

1. Different uses of Artificial Intelligence by courts

1. AI brings the prospect of profound transformations in legal practice and the delivery of judicial services, including:

a) *Precedent research*: In a context where precedents are becoming increasingly significant, even in civil law countries, AI proves immensely valuable for efficient jurisprudence research.

b) *Draft preparation*: The ability for AI to draft legal documents for lawyers, opinions for the Public Prosecution Office, and decisions for judges, based on researched and generated templates, will simplify workflows and shorten processing times. Evidently, all such work will remain under strict human supervision, as the responsibility ultimately lies with these professionals.

c) *Court tools:* In courts, AI programs that group cases by subject and summarize voluminous case files help optimize the time and effort of judges.

d) *Process and procedure automation:* In Brazil, nearly all legal processes are now electronic. The automation of procedures – such as summons and asset searches - as well as the automatic enforcement of decisions, greatly simplifies operations (e.g., decisions granting social security benefits are automatically entered into the INSS [National Institute of Social Security] system).

e) *Online dispute resolution:* An important frontier in the Judiciary is the online resolution of disputes, with the potential to make justice more agile and efficient.

⇒ In Brazil, across various courts, there are over a hundred projects utilizing AI to enhance judicial services.

2. Risks in judicial application.

1. All those uses raise a controversial and particularly intriguing issue: the use of AI in drafting judicial decisions. Concerns abound, and not without justification, regarding risks such as bias, discrimination, lack of transparency, and explainability. Equally troubling is the potential absence of social sensitivity, empathy, and compassion in such processes.

2. It is essential to recognize that human judges are equally prone to risks such as bias, discrimination, and undue influence. This highlights another perspective: the possibility that AI could be better equipped to deliver impartial decisions, less influenced by personal interests, political pressures, or intimidation.

3. This potential advantage is particularly significant in less developed countries, where judicial independence may be limited, and corruption more prevalent[9].

⇒ Nonetheless, at the current stage of societal and technological development, human oversight by a judge remains indispensable. To ensure accountability, judges might face a heightened duty to justify their decisions in cases where they deviate from outcomes proposed by AI.

3. Concerns regarding fundamental rights, democracy and governance

1. In conclusion, I mention concerns regarding the use of Artificial Intelligence, on the one hand, and the protection of fundamental rights, defense of democracy and good governance, on the other.

(i) Fundamental rights under risk

a) *Privacy and data protection*. The use of AI must respect the personal data of individuals and legal entities, and it should not use such data without consent. Invasive surveillance, such as facial recognition, biometrics, and location tracking, should be employed in a restricted and controlled manner. Given the vast amounts of data used to train AI, appropriate security mechanisms must be in place to prevent data breaches.

[9] On the issue, refer to ARIEL GUSTAVO. Are artificial intelligence courts a discrimination risk? *European AI alliance*, Aug 31, 2021. Available at: https://futurium.ec.europa.eu/en/european-ai-alliance/open-discussion/are-artificial-intelligence-courts-discrimination-risk. Accessed on April 3, 2024.

b) *Right to equality and non-discrimination.* The equality of all individuals, in its formal, material, and recognition dimensions, is one of the core pillars of contemporary civilization. As previously discussed, the risks of algorithmic discrimination have been raised. AI regulation must prevent individuals from being unjustly treated based on suspect categories that amplify vulnerabilities, such as gender, race, sexual orientation, religion, age, and other characteristics. There are concerning precedents in this area.[10]. **(Poliafetivas)**

c) *Freedoms (individual autonomy, cognitive freedom and freedom of expression.* In terms of *individual autonomy*, the use of neuroscience and targeted advertising (microtargeting) has the potential to manipulate behavior and decision-making by exploiting emotions such as fear, prejudice, euphoria, and other cognitive biases. This can lead individuals to make choices - such as purchasing goods, hiring services, or adopting behaviors - that go against their best interests, thereby violating their cognitive freedom or mental self-determination. Additionally, *the right to information and freedom of expression* can be compromised by recommendation or moderation algorithms that filter, direct, and exclude content, effectively acting as private censorship.

(ii) Protecting democracy

a) *Combating disinformation.* Democracy is a system of collective self-governance that depends on the informed and thoughtful participation of citizens. The spread of disinformation and conspiracy theories misleads people or fosters unfounded fears, undermining their ability to make sound decisions. As previously highlighted, this issue is worsened by deepfakes, which create fake videos and speech that appear

[10] Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, REUTERS, Oct. 10, 2018, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G/; Jeff Larson et al., *How We Analyzed the COMPAS Recidivism Algorithm*, PROPUBLICA, May 23, 2016, https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm; Ziad Obermeyer et al., *Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations*, SCIENCE, Oct. 25, 2019, https://www.science.org/doi/full/10.1126/science.aax2342; Will Douglas Heaven, *Predictive Policing is Still Racist—Whatever Data it Uses*, MIT TECHNOLOGY REVIEW, Feb. 5, 2021, https://www.technologyreview.com/2021/02/05/1017560/predictive-policing-racist-algorithmic-bias-data-crime-predpol.

convincingly real. We are all conditioned to trust what we see and hear. Such manipulations are inherently destructive to democracy[11].

b) *Fighting hate speech.* Since the establishment of universal suffrage, democracy has been founded on the equal participation of all individuals. Hate speech includes attacks on vulnerable groups, such as racist, discriminatory, or ableist rhetoric targeting Black people, LGBTQ+ individuals, people with disabilities, Indigenous people, and others. By aiming to discredit, weaken, or silence certain social groups, hate speech undermines the core values of democracy.

c) *Combating attacks on democratic institutions.* Social media, supported by AI, has been instrumental in orchestrating attacks on democratic institutions, with the goal of destabilizing them. Insurrectionary acts, such as those on January 6, 2021, in the United States, and January 8, 2023, in Brazil, involving coup attempts to undermine election results, pose a serious threat to democracy and must not be tolerated.

(iii) Promoting good governance

1. In light of the international, regional, and domestic recommendations and normative acts mentioned earlier, as well as the ongoing public debate in academia, civil society, and the media, several overlapping consensuses regarding AI governance can be identified, outlined in the five guidelines presented below:

1. *Focus on the common good.* AI should be developed with a focus on the well-being of individuals, countries, and the planet. Its benefits must be distributed fairly among all, and its negative impacts should be mitigated through legislation and regulation;

2. *Plural governance.* AI governance should incorporate, at various stages and with appropriate proportionality, the participation of a diverse range of stakeholders, including public authorities, scientists and researchers, civil society, academia, businesses, and human rights organizations. The diversity of perspectives and

[11] Harari advocates for a ban.

the balancing of values and interests are crucial for the legitimacy of decisions and the development of appropriate regulations;

3. *Transparency and explainability.* Transparency refers to providing users with a basic understanding of how the system operates and informing them when they are interacting with an AI system. Explainability involves making the reasons behind decisions intelligible, enabling users to question and, if necessary, challenge the outcomes.

⇒ In this context, there is an increasing call for the recognition of a new human right: *the right to an explanation*[12].

4. *Security*. AI systems must be internally secure to prevent errors that lead to undesirable outcomes, while also being protected against external attacks. Security in AI usage involves impact assessments, ensuring data quality, implementing robust cybersecurity measures, and mapping the processes and decisions that make up the AI lifecycle (traceability).

5. *Control and accountability.* Human oversight and control are essential to ensure that AI operates within the boundaries of legality, ethics, and justice. Despite the relative autonomy of its decision-making processes, responsibility always lies with a natural or legal person. In cases of misuse or malicious activity, one or both will be subject to civil, administrative, and criminal liability.

Conclusion

1. The Technological or Digital Revolution has drastically transformed our lives, reshaping the course of history. Among its many advancements are the Internet, digital platforms, and Artificial Intelligence.

2. Artificial Intelligence holds immense potential to enhance quality of life but also presents significant risks that must be proactively managed. While regulation is essential, it must strike a balance to avoid stifling research and innovation. This

[12] Yuval Noah Harari, *Nexus*, 2024, p. 331.

transformative technology has the power to profoundly redefine the role and meaning of the human condition.

3. Artificial Intelligence can be extremely useful for the Judiciary. As a general rule, however, human oversight and the judge's ultimate responsibility for decisions cannot be dispensed with.

⇒ Amid all the advancements and transformations, the core existential values that should guide a meaningful life remain unchanged and must be preserved: goodness, justice, the pursuit of truth, and human dignity.