

PARTIAL ALGORITHMS AS A NEW FORM OF DISCRIMINATION:
A PRINCIPLE-BASED REGULATION PROPOSAL SUPPORTED BY THREE AXES*

Wilson Engelmann**

How does artificial intelligence absorb discriminatory patterns and what can science do to avoid these distortions? What are the legal and ethical impacts? These two problems will guide my presentation in this class. I am not an expert in Data Sciences or Data Computing, but a researcher in the field of Law concerned with the advances and impacts that the increasing use of Artificial Intelligence, from now on, just AI, could generate in our lives.

I will make use of the contribution of Michel Serres, “Petite Poucette”, for its relevance, especially for the guideline given by the subtitle: “a new way of living in harmony, of thinking about institutions, of being and knowing”. Driven by media interference, the “Petite Poucette” and all of us, we now inhabit the virtual. So much so that Luciano Floridi will speak about a “hyperconnected era” (Floridi, 2009). Now we can write with your thumbs, faster than I could ever write with all of my fingers.

We are gradually moving from the “Industrial Age”, where the “Newtonian” or macroscopic vision predominates, to the “Age of information and knowledge”, where the “relativistic and quantum vision” is observed (Kim, 2010). If before the quantum perspective, two bodies could not occupy the same space at the same time, with advances in quantum vision, two bodies will be able to occupy the same space simultaneously (Kop, 2020)

We talk about the nanoscale, that is, the construction of things and life itself from the billionth part of a meter. The “chip”, for example, that supports the development of

* This study is a partial result of the following research projects: a) CNPq* call n. 09/2023 - Research Productivity Grants - PQ, project entitled: “Experiments in Law: challenges and possibilities for regulation based on the principles of artificial intelligence and its testing in a Regulatory Living Lab”; b) Call for CNPq* Universal 2023, project entitled: “Human Rights and Artificial Intelligence: from the violation of personality rights to the need to regulate new technologies” (Brazil). This work is also linked to the research carried out by Engelmann at the following international research centers: CEDIS - Centro de Investigação & Desenvolvimento sobre Direito e Sociedade at the Universidade Nova de Lisboa Law School, Portugal; Instituto Jurídico Portucalense at the Universidade Portucalense University in Porto, Portugal; CEAD (Center for Advanced Studies in Law - Centro Francisco Suárez), at the Lusófona de Lisboa University, Portugal, and CESEG - Centro de Estudios de Seguridad at the Santiago de Compostela University, Spain.

** MPH, PhD, Public Law, Postgraduate Program in Law at the Vale do Rio dos Sinos University - UNISINOS, Brazil; Post-Doctoral Internship in Public Law-Human Rights (Centro de Estudios de Seguridad - CESEG) at the Santiago de Compostela University, Spain; Faculty member and researcher of the Graduate Program in Law - Master’s and Doctorate degrees, and the Graduate Program - Professional Master’s Degree in Company and Business Law, both from UNISINOS (Brazil); CNPq* Research Productivity Scholarship; Leader of the JUSNANO Research Group; E-mail: wengelmann@unisinos.br; ORCID: <https://orcid.org/0000-0002-0012-3559>.

* CNPq - The National Council for Scientific and Technological Development is a foundation linked to the Ministry of Science and Technology (MCT), to support Brazilian research.

algorithms, with which AI is developed, is on this nanometric size/measurement scale. We now live in the so-called “Fourth Industrial Revolution”, according to Klaus Schwab (2016).

And then we must ask: “What is Artificial Intelligence?”. This is the title of a thought-provoking article by John McCarthy (1956); from Computer Science Department, Stanford University:

“Q. What is artificial intelligence?

A. It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

Q. But what is intelligence?

A. Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.

Q. Isn't there a solid definition of intelligence that doesn't depend on relating it to human intelligence?

A. Not yet. The problem is that we cannot yet characterize in general what kinds of computational procedures we want to call intelligent. We understand some of the mechanisms of intelligence and not others.

Q. Is intelligence a single thing so that one can ask a yes or no question ‘Is this machine intelligent or not?’?

A. No. Intelligence involves mechanisms, and AI research has discovered how to make computers carry out some of them and not others. If doing a task requires only mechanisms that are well understood today, computer programs can give very impressive performances on these tasks. Such programs should be considered ‘somewhat intelligent’.

Q. Isn't AI about simulating human intelligence?

A. Sometimes but not always or even usually. On the one hand, we can learn something about how to make machines solve problems by observing other people or just by observing our own methods. On the other hand, most work in AI involves studying the problems the world presents to intelligence rather than studying people or animals. AI researchers are free to use methods that are not observed in people or that involve much more computing than people can do”.

What does algorithmic discrimination mean? Algorithms have become a key element that supports crucial services and infrastructures in information societies, which, according to

Luciano Floridi, is transformed into the *OnLife society*, where life is on the web constantly, regardless of whether it is On-Line or Off-Line (Floridi, 2009). Individuals interact daily with recommendation systems - algorithmic systems that make suggestions about what a user might like - whether to choose a song, a movie, a product or even a friend. In this technological scenario, several ethical, social, political and legal impacts are observed.

There is evidence that new forms of discrimination are being formed: I would highlight two of them:

- a) the lack of access for everyone in society to a quality Internet network and also adequate equipment to be on the World Wide Web;
- b) the AI system itself shows discriminatory biases.

How is “bias” defined? In this context, it will be important to define what the “algorithm” is, which is at the basis of AI: there is a debate about whether such a machine can “think”. At this point, the very concept of Intelligence that comes into play is controversial, as an algorithm that makes decisions is a structure based on mathematical equations submitted to a previously designed processing system - precisely the algorithm, as a kind of recipe - that is capable of to present solutions in millionths of seconds. If there is an error, whether due to poor collection, poor processing, or the unfeasibility of the algorithmic solution to the problem, the risk of an unbiased decision by that algorithm arises. It can be said that algorithmic bias indicates “inconsequential” failure and could cause harm. In any case, the “antecedent” issue must also be considered, that is, the causal path of the algorithmic process must be analyzed so that it is possible to find out whether the decision riddled with bias was taken due to a failure that occurred in a previous stage, which has tainted subsequent processing stages (Sedlakova e Trachsel, 2023; Floridi, 2009).

However, expanding the research, I observed that this definition is not so simple, it is more complex, because if the algorithm were a kind of “cake recipe”, for example, it would be enough to always follow the same route, and you would have similar results (Dellunde, Pujol e Vitrià, 2023). The idea of discrimination occurs at a level of abstraction that requires the concept of agency and from this point of view we believe that the use of the term “algorithm” is not particularly appropriate, because of its reductionist meaning. The classic concept of “algorithm”, understood as a sequence of instructions (as if it were a cake recipe) that serve to solve a problem, evokes a deductive approach that does not represent current artificial intelligence (AI) systems. AI systems are placed in a special category of systems that, following Daniel Dennett’s theoretical framework, according to his book: *The*

intentional stance (1989), must be understood from an intentional approach and that have a different nature from classical algorithms (Dennett, 1989).

According to Daniel Dennett's (1989) proposal we have three alternatives when we want to understand a complex system: a) the first alternative, the physical approach, uses the laws of physics at a certain level of abstraction to model the system based on its constituents and the interactions that we can observe; b) the second alternative, the design approach, allows us to understand a system based on the assumption that it has been designed for a purpose and that therefore its behavior can be expected to conform to this purpose; c) the third alternative, or intentional approach, is applied to those systems that are better understood as rational agents, which can be assumed to have beliefs, a purpose and even a certain representation of the world that allows them to achieve their purpose.

The main element of any theory of AI (in)equity is that data-driven AI models build a predictive system for decision-making inductively, unlike classical methods, which are based on a process deductive. It is this difference that determines, in most cases, the use of design or intentional approaches for its understanding. The inductive process is complex and involves a great variety of elements of a different nature, a fact that reinforces the concept of a system instead of that of an algorithm. On the other hand, this process can be the source of biases that result in discrimination problems (Dennett, 1989; Mittelstadt, Wachter e Russell, 2023).

In the literature (Kaplan e Haenlein, 2019; Ester-Sánchez, 2023), various stages of AI development also highlight new levels of risk, the main one being the possibility of carrying out human actions in all areas. Stefano Rodotà (2014), mentions that we live in the "dictatorship of the algorithm", with the possibility of its deepening and expansion, generating a depersonalization of society, where the person of the decision maker, in fact replaced by the automated procedure; he disappears and the person considered in himself, transformed into an object of uncontrollable powers. As a result, it has a character that borders on being considered "scary", given the possible impacts (Martin e Waldman, 2022).

How is the situation of "algorithmic discrimination" recognized, considering advances in artificial intelligence? In Germany, the following situation emerged: the risk of predictive policing programs based on inadequate information or inappropriate predictions, which may turn out to be biased, always covering the same areas and the same people (or specific group of people). A specific case, with this factual basis, was taken to the German Constitutional Court (German..., 2023): "can surveillance software, oriented in its use for predominantly predictive purposes, be used by police forces to prevent and combat the occurrence of

phenomena of a criminal nature?” The trial of this case began at the end of December 2022 and is concluded in 2023. Predictive use is based, to this extent, on mathematical models and other techniques of an analytical nature in the law enforcement process, with the ultimate scope of identifying potential criminal activities and, as well, its potential and predictable authors. In this real case, life is “imitating fiction”, which was portrayed in the film “Minority Report”, the German Constitutional Court was facing, in the real world, a complaint filed by the German Society for Civil Rights, about the legality of using of the powerful “Gotham” software from the American company Palantir and intended to promote police surveillance in the states of Hesse and Hamburg. Privacy concerns, given the sensitive and personal nature of the data collected, risks associated with the security of this data and its illegitimate sharing with third parties for other purposes, risk of lack of accuracy, as the reliability of predictive policing is closely linked to quality data and the integrity of its users. It is enough, for example, to think that criminal statistics collected based on racist or xenophobic policies will tend to encourage racist predictions, leading to excessive and disproportionate policing of certain population segments and feeding a true vicious circle of permanent stigmatization of the same groups and individuals. Due to these characteristics, the German Constitutional Court decided that there would be harm to fundamental rights and prohibited the use of the surveillance system (Novelli, Casolari e Rotolo, 2023).

Other situations can also be highlighted: AI can be used to analyze the entire financial health, including expenses, savings and other debts, to reach a broader panorama of the consumer (Muir, 2023): if designed correctly, these systems can provide broader access to accessible credit. But one of the biggest dangers is unintentional bias, in which algorithms end up denying loans or accounts to certain groups, including women, migrants or people of color. Part of the problem is that most AI models can only learn from historical data that they have been fed, meaning they will learn what type of customers were previously lent to and which customers were marked as untrustworthy, in addition to having historical data of movements largely of white people only. As such, there is a danger that they will be biased in terms of what a ‘good’ borrower looks like: gender and ethnicity often play a significant role in the AI’s decision-making processes based on the data it has been taught on. Additionally, some models are designed to be blind to so-called protected characteristics, meaning they should not consider the influence of gender, race, ethnicity, or disability.¹ But these AI models can

¹ Regarding the challenges and risks of using facial recognition cameras in Brazil, see an interesting study carried out by Nunes, 2024.

still discriminate as a result of analyzing other data, such as zip codes, which can correlate with historically disadvantaged groups who have never applied for, secured, or repaid loans or mortgages. And in most cases, when an algorithm makes a decision, it is difficult for anyone to understand how it reached that conclusion, resulting in what is commonly called “black box” syndrome (Martín-Ríos, 2022). This means that banks, for example, may struggle to explain what an applicant could have done differently to qualify for a loan or credit card, or whether changing an applicant’s sex from male to female could result in a different result (Creel e Hellman, 2022).

There are other examples mapped out in specific literature: the use of the algorithm in the selection of candidates for job vacancies, where the “preference” for male candidates is observed, due to the fact that the historical data collected by the algorithm shows the historical predominance of men in the world of work (Ogunniye *et al.*, 2021). In addition to these issues, attention is drawn to the possibility that AI will eventually destroy almost every job that now exists. In the world's work environment there will be increasing unemployment due to automation and the expansion of the so-called "Industry 4.0" (Rahn, 2024).

In these cases, the absence of typical human conceptions, such as empathy, is evident and leads to the datafication of the world and people, a true digitalization of life and people.

The recipe for preconception or discrimination:

When using machine learning systems, society runs the risk of inadvertently perpetuating biases, thanks to the common sense that views mathematics as neutral. An American data engineer, Fred Benenson, coined a term to define this risk: “mathwashing”. It was based on “greenwashing”, the use of marketing strategies by companies to simulate environmental concern. Likewise, the idea that algorithms are neutral also benefits and exempts the companies that use them from responsibility (Makortoff, 2022; Tunes, 2020).

It turns out that artificial intelligence systems are fed by data, and those who select this data are human beings - who can be driven by prejudices unconsciously or intentionally. An example of this was explained by a study published in October in the journal Science, led by a researcher from the School of Public Health at the University of California at Berkeley, in the United States. In a large hospital in that country, the North American group found that the algorithm responsible for classifying patients most in need of follow-up - as they were at greater risk² - privileged white people over black people. This happened because the system

² In France, I found in my research an interesting publication by “Laboratoire d’innovation numérique da la CNIL” (LINC), at Avril 2022, a *Dossier: Sécurité des Systèmes d’IA*, and I can underline: “AI systems generate specific security risks in comparison to traditional information systems, as new machine learning capabilities

was based on payments to health plans, which are higher for people who have more access to medical care, and not on the probability of each person having serious or chronic illnesses (Obermeyer *et al.*, 2019).

According to a publication by Luciano Floridi and colleagues (Floridi, 2021): “Bias has also been reported in algorithmic advertisement, with opportunities for higher-paying jobs and jobs within the field of science and technology advertised to men more often than to women. While solutions to these issues are being discussed and designed, the number of algorithmic systems exhibiting ethical problems continues to grow”.

Who is responsible for discrimination generated by the algorithm? And how can these situations be regulated? Create laws, but here we have the difficulty of harmonizing the “time of Law” and the “time of technology”. Principles-based regulation? Standards of good conduct or even technical standards, perhaps, even data-based regulation.

All technological innovations, such as AI, must be guided, according to Erich Fromm’s concept, by a “humanistic management” which recognizes: consideration must always be given to the consequences (even if potential) that could harm each participant affected by the technology under development, not admitting move forward without contemplation and contingency of all negative possibilities, at least the known ones (Fromm, 1968).

The importance of explainability, as a fundamental principle for the AI era; transparency (Floridi, 2023) to know all the movements of the algorithm when making a certain decision; duties of care; preservation of the centrality of the human person; the principle of good faith; the principle of trust; perhaps a “Trust Code”; the development of new assumptions for civil liability, which should be inspired by the “Law of Damages”, that is, civil liability in anticipation of damage, based on the precautionary principle.

Additionally, the principles defined by the European Union’s AI Regulation must be incorporated, which establishes, right at the beginning of the document: “[...] promote the adoption of human-centered and trustworthy artificial intelligence (AI), ensuring simultaneously a high level of protection of health, safety, fundamental rights enshrined in the Charter of Fundamental Rights of the European Union ('Charter'), in particular democracy, the rule of law and environmental protection, protection against harmful effects of AI systems in the Union, and to support innovation. (Regulamento...2024).” This combination of

increase the “attack surface” of these systems, by introducing numerous (and new!) vulnerabilities. The LINC offers you a triptych of articles in order to see things more clearly.” Available at: https://linc.cnil.fr/sites/linc/files/atoms/files/linc_cn timer dossier-securite-systemes-ia.pdf.

principles seems to be a good way to work with the creation of codes of conduct by companies that research and develop AI systems.

For the area of Law, the revaluation of the normative value of principles will be a creative alternative. Through the principles, it will be possible to connect the different and varied ways of regulating AI internally, opening up networks for regulating AI advances, in a position to connect researchers and companies linked to this powerful technology, which presents broad positive and negative possibilities.

This perspective can be constructed from three structuring axes. The first axe: “Praecepta iuris”, by *Justinian*, extracted from the Digest: That nucleus equivalent to the three precepts of law, the *praecepta iuris* of the Romans: the “*honeste vivere*”, the “*alterum non laedere*” and the “*sum cuique tribuere*”: “Justice is a steady and enduring will to render unto everyone his right. 1. The basic principles of right are: to live honorably, not to harm any other person, to render to each his own” (Watson, 1985). The first, which has to do with the concepts of general or social justice, goes beyond the literal translation of a mere honest living, given that this precept, belonging to the world of law, and not of morality, means, in a substantial way, that No one should abuse their powers, their rights, leading to the prohibition of both the abuse of rights and the abuse of freedom. The second, not harming anyone, relates to the categories of commutative or synalagmatic justice, implying both an idea of exchange and the general principle that whoever harms, whoever damages the other has to erase the damage he caused to that other, has to indemnify or indemnify. The third, attributing to each person what belongs to them, has to do with the categories of distributive justice, implying giving to each person according to their needs. In fact, the third precept compensates for the requirement of the first, imposing that each person also gives to the whole according to their possibilities (Maltez, 2024).

The second axe can be found in John Finnis’ book “Natural Law and Natural Rights”, when he refers to the existence of “basic human goods” and “requirements of practical reasonableness”; the preponderance of the “common good” stands out from the second axe: the search for the realization of the “common good” signals material actions for individual and collective well-being, that is, “[...] a set of conditions that enable members of a community to achieve reasonable goals for themselves, or to reasonably realize for themselves the value (or values) for which they have reason to mutually collaborate (positively and/or negatively) in a community” (Finnis, 2011; Colombo *et al.*, 2023). This concept highlights the overlap between public and private so that the “community” can

achieve reasonable objectives. That is to say, objectives constructed collaboratively, where private organizations now have a renewed role.

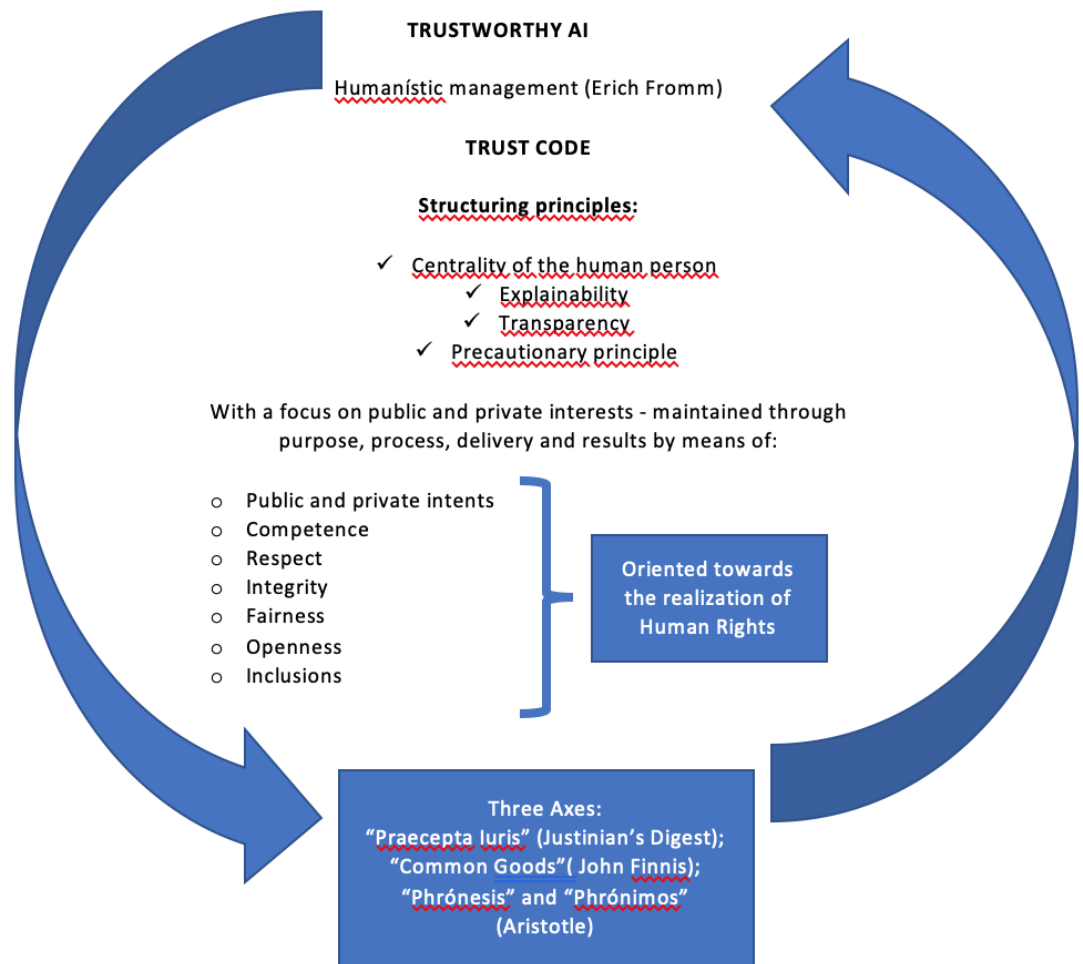
This notion of “common good”, and then we have the third axis, should be combined with Aristotle’s perspective, who, before saying what “phronesis” means, presents the characters of “phronimos”: It is thought that it is characteristic of a person who has “phronesis” to “be able to deliberate well about what is good and convenient for themselves, not in relation to a particular aspect – for example, when one wants to know which species are of things that contribute to health and physical vigor – but about the types of things that lead us to live well in general” (Aristóteles, 2001). It is evident that “phronesis” appears as a type of practical knowledge, which points to the prudence of person in his actions, whose objective is the realization of his good. However, this is not an individualistic conception of good, but its contextualization designed for the entirety of person. It appears, therefore, that “phronesis” aims to harmonize partial goods, in such a way as to enable harmony as a whole. This search requires learning, aiming to develop “the ability to identify and order the goods of the good life, whose achievement implies the ordering of all these other sets of goods, requires character training in and for these excellences, a type of training whose meaning emerges only throughout the process” (MacIntyre, 2001).

Any new technology will only be acceptable if it manages to respect these three axes of the legal tradition. The good of the community must always be sought, avoiding risks that could pose a threat to the full development of human beings. This set of structuring elements contain several principles that concern algorithmic discrimination. The principles have a deontic character in their structural composition, which is very close to the character that the rules also present (Engelmann, 2023). Therefore, in my research I have asked myself: if the legal tradition of principles shows this characteristic of principles, what is the reason for always waiting for the elaboration of the rule through slow and politically bureaucratic legislative activity of the State? The legal role that the principles present, and which serve to regulate the issues raised by AI, should be revalued.

The partial algorithms as a new form of discrimination can be addressed through the principle-based regulation proposal supported by three axes and which is summarized in the following figure (Figure 1)³:

³ Inspired from: Organisation for Economic and Development, 2017; Soeteman-Hernández, *et al.*, 2021.

Figure 1 - Principle-based regulation proposal supported by three axes



Source: Prepared by the author.

Reliable AI must be based on so-called humanistic management (Erich Fromm). Based on this assumption, the Code of Trust is organized as a Voluntary Code of Conduct, structured based on the following principles: Centrality of the human person; Explainability; Transparency; Precautionary principle. With a focus on public and private interests - maintained through purpose, process, delivery and results by means of the following seven trust drivers for emerging tech governance, as AI: Public and private intentions; Competence; Respect; Integrity; Fairness; Openness and Inclusions. These focal elements must be oriented towards the realization of Human Rights, understood from John Gerard Ruggie's (2011, 2007, 2018) framework: protect, respect and remedy. This principle-based regulatory architecture is

organized and can be practiced through three legal-philosophical axes: “Praecepta Iuris” (Justinian’s Digest); “Common Goods” (John Finnis) and “Phrónesis” and “Phrónimos” (Aristotle).

REFERENCES

ARISTÓTELES. 2011. *Ética a Nicômacos*. Traduzido do grego, introdução e notas de Mário da Gama Kury. 4. ed. Brasília, UnB.

COLOMBO, C.; ENGELMANN, W.; WERNER, D. A.; VALONGO, A. 2023. Intelligenza artificiale nei sistemi giuridici brasiliano e italiano: osservazioni sull’oggi, riflessioni sul domain. *Le Corti Salernitane*, 20(3):591-646.

CREEL, K.; HELLMAN, D. 2022. The algorithmic Leviathan: arbitrariness, fairness, and opportunity in algorithmic decision-making systems. *Canadian Journal of Philosophy*, 52(1):26-43.

DELLUNDE, P; PUJOL, O; VITRIÀ, J. 2023. Cerrando una brecha: una reflexión multidisciplinar sobre la discriminación algorítmica. *Daimon Revista Internacional de Filosofía*, 90.

DENNETT, D. 1989. *The intentional stance*. Massachussets, MIT Press.

ENGELMANN, W. 2023. *Direito natural, ética e hermenêutica*. 2. ed. rev., atual. e ampl. Rio de Janeiro, Lumen Juris.

ESTER-SÁNCHEZ, T. 2023. El desafío de la inteligencia artificial a la vigencia de los derechos fundamentales. *Cuadernos Electrónicos de Filosofía del Derecho*, 48:111-139. Available online: <https://doi.org/10.7203/CEFD.48.25863>.

FINNIS, John. 2011. *Natural Law & Natural Rights*. 2. ed. Oxford: Oxford University Press.

FLORIDI, L. *et al.* 2011. The ethics of algorithms: key problems and solutions. *AI & Society*, 37:215-230. Available online: <https://doi.org/10.1007/s00146-021-01154-8>

FLORIDI, L. 2023. On Good and Evil, the mistaken idea that technology is ever neutral, and the importance of the double-charge thesis. *Philosophy & Technology*, 36(60). Available online: <https://doi.org/10.1007/s13347-023-00661-4>.

FLORIDI, L. 2009. *The onlife manifesto: being human in a hyperconnected era*. Oxford, Springer.

FROMM, E. 1968. *The revolution of hope: toward a humanized technology*. New York: Harper & Row.

GERMAN Constitutional Court blocks police use of surveillance software. 2023. Available at: <https://iapp.org/news/a/german-constitutional-court-blocks-police-use-of-surveillance-software/>

KAPLAN, A.; HAENLEIN, M. 2019. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62:15-25.

KIM, Y-G. 2010. A shift of higher educational paradigm with scientific development from isolation to integrative/holistic global education in the twenty-first century. *Educational Research*, 1(4): 075-087. Available online: <http://www.interesjournals.org/ER>.

KOP, M. 2020. Regulating Transformative Technology in The Quantum Age: Intellectual Property, Standardization & Sustainable Innovation, 2 TTLF Newsletter on Transatlantic Antitrust and IPR Developments Stanford-Vienna Transatlantic *Technology Law Forum*, Stanford University Available at: <https://law.stanford.edu/publications/regulating-transformative-technology-in-the-quantum-age-intellectual-property-standardization-sustainable-innovation/>.

LABORATOIRE d'innovation numérique da la CNIL” (LINC) 2022. *Dossier: Sécurité des Systèmes d'IA*. Available at: https://linc.cnil.fr/sites/linc/files/atoms/files/linc_cnil_dossier-securite-systemes-ia.pdf.

MACINTYRE, A. 2001. *Justiça de quem? Qual racionalidade?* Traduzido por Marcelo Pimenta Marques. 2. ed. São Paulo: Loyola.

MAKORTOFF, K. 2022. Risks posed by AI are real: EU moves to beat the algorithms that ruin lives. *The Guardian*. Available in: <https://www.theguardian.com/technology/2022/aug/07/ai-eu-moves-to-beat-the-algorithms-that-ruin-lives>.

MALTEZ, J.A. 2024. *Praecepta Iuris*. Available at: <https://maltez.info/aaanetnovabiografia/Conceitos/Praecepta%20iuris.htm>.

MARTIN, K.; WALDMAN, A. 2022. Are algorithmic decisions legitimate? The effect of process and outcomes on perceptions of legitimacy of AI decisions. *Journal of Business Ethics*, 1-18.

MARTÍN-RÍOS, P. 2022. Empleo de big data y de inteligencia artificial en el ciberpatrullaje: de la tiranía del algoritmo y otras zonas oscuras. *Revista de Internet, Derecho y Política*, 36: 1-13. Available at: <https://doi.org/10.7238/idp.v0i36.394511>.

MCCARTHY, J. 1956. *What is artificial intelligence?*. Available at: <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>

MITTELSTADT, B.; WACHTER, S.; RUSSELL, C. 2023. *The Unfairness of Fair Machine Learning: Levelling down and strict egalitarianism by default*. Available at SSRN: <https://ssrn.com/abstract=4331652>.

MUIR, M. 2023. AI: the answer to everyday finance? *Financial Times*, Available at: <https://www.ft.com/content/59b1a3cf-ae0-42b0-85fb-1a00a9162696>.

NOVELLI, C.; CASOLARI, F.; ROTOLO, A. 2023. Taking AI risks seriously: a new assessment model for the AI Act. *AI & Society*. Available at: <https://doi.org/10.1007/s00146-023-01723-z>.

NUNES, P. 2024. *Reconhecimento facial e a segurança pública para poucos*: Brasil segue avançando acriticamente nesse campo. Available at: <https://www.jota.info/opiniao-e-analise/colunas/ia-regulacao-democracia/reconhecimento-facial-e-a-seguranca-publica-para-poucos-10072024>.

OBERMEYER, Z. *et al.* 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464): 447-453. Available at: 10.1126/SCIENCE.AAX2342.

ORGANISATION FOR ECONOMIC AND DEVELOPEMNT. 2017. *Public Governance Reviews*: Trust and Public Policy how better governance can help rebuild Public Trust. Available at: <https://www.oecd-ilibrary.org/docserver/9789264268920-en.pdf?expires=1720817062&id=id&accname=guest&checksum=D7876A4491534798B74FFAA523B2603A>.

OGUNNIYE, G. *et al.* 2021. Understanding user perceptions of trustworthiness in E-recruitment systems. *IEEE Computer Society*. Available at: <https://ora.ox.ac.uk/objects/uuid:b9740967-82bc-4861-b440-18a64261f54b>.

RAHN, R. W. 2024. AI will eventually destroy almost every job that now exists. *The Washington Times*. Available in: <https://www.washingtontimes.com/news/2024/jul/8/ai-will-eventually-destroy-almost-every-job-that-n/>

REGULAMENTO (UE) 2024/1689 do Parlamento Europeu e do Conselho: que cria regras harmonizadas em matéria de inteligência artificial e que altera os regulamentos (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e as diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (Regulamento da Inteligência Artificial). 2024. Available at: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=OJ:L_202401689

RODOTÀ, S. 2014. *Il mondo nella rete*: quali i diritti, quali i vincoli. Roma, Laterza.

RUGGIE, J. G. 2007 Business and Human Rights: the evolving international agenda. *The American Journal of International Law*, 101(4):819-840. Available at: <https://www.jstor.org/stable/40006320>.

RUGGIE, J.G. 2011. Human Rights Council Seventeenth Session; Agenda item 3: Promotion and protection of all human rights, civil, political, economic, social and cultural rights, including the right to development. In: REPORT of the Special Representative of the Secretary-General on the issue of human rights and transnational corporations and other business enterprises. Available at: <https://www.right-docs.org/doc/a-hrc-17-31/>.

RUGGIE, J.G. 2018. Multinationals as global institution: power, authority and relative autonomy. *Regulation & Governance*, 12:317-333. Available at: <https://doi-org.ez101.periodicos.capes.gov.br/10.1111/rego.12154>.

SCHWAB, K. 2016. *The Fourth industrial revolution*: what it means, how to respond, 2016. Available at: <https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/>.

SEDLAKOVA, J.; TRACHSEL, M. 2023. Conversational Artificial Intelligence in Psychotherapy: A New Therapeutic Tool or Agent? *The American Journal of Bioethics*, 23(5):4-13. Available at: DOI: 10.1080/15265161.2022.2048739.

SOETEMAN-HERNÁNDEZ, L.G. *et al.* 2021. Modernizing innovation governance to meet policy ambitions through trusted environments. *NanoImpact*, 21. Available et: www.elsevier.com/locate/nanoimpact. <https://doi.org/10.1016/j.impact.2021.100301>.

TUNES, S. 2020. Algoritmos parciais. *Revista Pesquisa FAPESP*, 287:62-67. Available in: <https://revistapesquisa.fapesp.br/algoritmos-parciais/>.

WATSON, A. 1985. *The Digest of Justinian*. Philadelphia: University of Pennsylvania v. 1.