



Research article

UDC 34:004:17:004.8:342.7

EDN: <https://elibrary.ru/egkppn>

DOI: <https://doi.org/10.21202/jdtl.2025.7>

Artificial Intelligence in Healthcare: Balancing Innovation, Ethics, and Human Rights Protection

Pedro Miguel Alves Ribeiro Correia ✉

University of Coimbra, Coimbra, Portugal

Ricardo Lopes Dinis Pedro

University of Lisbon, Lisbon, Portugal

Susana Videira

University of Lisbon, Lisboa, Portugal

Keywords

artificial intelligence,
data protection,
ethical regulation,
ethics,
fundamental rights,
healthcare,
human rights,
law,
legal regulation,
predictive analytics

Abstract

Objective: to identify key ethical, legal and social challenges related to the use of artificial intelligence in healthcare; to develop recommendations for creating adaptive legal mechanisms that can ensure a balance between innovation, ethical regulation and the protection of fundamental human rights.

Methods: a multidimensional methodological approach was implemented, integrating classical legal analysis methods with modern tools of comparative jurisprudence. The study covers both the fundamental legal regulation of digital technologies in the medical field and the in-depth analysis of the ethical, legal and social implications of using artificial intelligence in healthcare. Such an integrated approach provides a comprehensive understanding of the issues and well-grounded conclusions about the development prospects in this area.

Results: has revealed a number of serious problems related to the use of artificial intelligence in healthcare. These include data bias, non-transparent complex algorithms, and privacy violation risks. These problems can undermine public confidence in artificial intelligence technologies and exacerbate inequalities in access to health services. The authors conclude that the integration of artificial intelligence into healthcare should take into account fundamental rights, such as data protection and non-discrimination, and comply with ethical standards.

✉ Corresponding author

© Correia P. M. A. R., Pedro R. L. D., Videira S., 2025

This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Scientific novelty: the work proposes effective mechanisms to reduce risks and maximize the potential of artificial intelligence under crises. Special attention is paid to regulatory measures, such as the impact assessment provided for by the Artificial Intelligence Act. These measures play a key role in identifying and minimizing the risks associated with high-risk artificial intelligence systems, ensuring compliance with ethical standards and protection of fundamental rights.

Practical significance: adaptive legal mechanisms were developed, that support democratic norms and respond promptly to emerging challenges in public healthcare. The proposed mechanisms allow achieving a balance between using artificial intelligence for crisis management and human rights. This helps to build confidence in artificial intelligence systems and their sustained positive impact on public healthcare.

For citation

Correia, P. M. A. R., Pedro, R. L. D., & Videira, S. (2025). Artificial Intelligence in Healthcare: Balancing Innovation, Ethics, and Human Rights Protection. *Journal of Digital Technologies and Law*, 3(1), 143–178. <https://doi.org/10.21202/jdtl.2025.7>

Contents

Introduction

1. The use of Artificial Intelligence in the Combat Against Pandemics and Why it Fails
2. Examples of Artificial Intelligence Failures: Lessons Learned and the Path Forward
3. The Problem with Models: Garbage In Garbage Out
4. The Sustainability Problem (almost) no One Talks About
5. Governance as Key: How State Measures and Data Availability Reinforce some Organizational Values and Contribute to the Sustainability of the National Health System
6. Governance as Key: How Governance Reinforces some Organizational Values and Contributes to the Sustainability of Crisis Management
7. Critical Reflections
8. AI and Fundamental Rights
9. Pandemics and Fundamental Rights
10. Possible Relationships between AI and Fundamental Rights in Combating Pandemics

Conclusions

References

Introduction

As practitioners, institutions and governments consider the looming shadow of disease and pandemics, can artificial intelligence (AI) become our panacea and our light? What limits can fundamental rights present when fighting pandemics with the use of AI?

Throughout history, humanity has faced the devastating wrath of epidemics and pandemics. From the bubonic plague to the Spanish flu, these outbreaks have reshaped societies and challenged our very existence. Today, as we navigate an increasingly connected world, the threat of new and rapidly spreading diseases looms large. Globalization is increasingly becoming a double-edged sword, fostering collaboration but also facilitating the swift travel of pathogens across borders ([Jones et al., 2008](#); [Morse et al., 2012](#)).

Artificial intelligence is now viewed as a new, emerging weapon in this age-old fight. This revolutionary, powerful technology (or amalgamation of technologies), once relegated to the realm of science fiction, holds immense potential to revolutionize how humans combat epidemics and pandemics. Poised to become a powerful weapon in our arsenal for combatting disease. By means of artificial intelligence techniques one might analyze vast amounts of data to predict outbreaks and outcomes, accelerate drug discovery, and even personalize treatment strategies ([Syrowatka et al., 2021](#)). Or so it is believed.

For instance, artificial intelligence can be used for forecasting the unforeseen, as it can sift through voluminous information from social media, news reports, and even satellite imagery, potentially identifying early warning signs of an outbreak before it explodes into a full-blown pandemic. It is not farfetched to imagine an artificial intelligence system detecting a surge in searches for flu-like symptoms in a specific region, triggering an immediate investigation that could potentially nip an outbreak in the bud ([Wong et al., 2023](#)).

Another instance is the use of artificial intelligence to accelerate the quest for cures (if that, and not only chronic disease mitigating solutions, perpetual cash cows, are still pursued by pharmaceutical companies). Medical chemical compounds discovery (including vaccines) has traditionally been a slow and arduous process, often taking years to yield results, if at all. Artificial intelligence can be used to revolutionize this process by analyzing molecular structures to identify potential drug candidates or repurpose existing drugs for new significantly reducing the time it takes to get life-saving treatments into the hands of patients ([Matsuzaka & Yashiro, 2022](#)).

Yet another instance is the tailoring of treatments, diverging from the one-size-fits-all approach, as artificial intelligence could be developed to be capable of analyzing patients' individual genetic makeups and medical histories to predict how they might respond to different treatment options, paving the way for personalized medicine and allowing doctors to tailor treatment plans for maximum effectiveness ([Topol, 2019](#)).

An additional example for the use of artificial intelligence, in this context, would be predicting the trajectory of an outbreak or a pandemic, due to the capability of these models to analyze historical data and disease characteristics, allowing public health officials to strategically allocate resources and implement targeted containment measures (Ferguson et al., 2006).

One final example (of numerous others not listed here) is the use of artificial intelligence to enhance contact tracing, as it can analyze contact tracing data and travel information to pinpoint individuals at high risk of infection, therefore helping healthcare workers prioritize testing and quarantine measures, potentially containing the spread of the pathogen (Fetzer & Graeber, 2021).

The path forward, however, is not without its challenges. Biases in data, the «black box» or “gray box” nature of complex algorithms, and the ever-present risk of over-reliance on artificial intelligence all pose potential and significant pitfalls (DeCamp & Tilburt, 2019).

Despite this perfunctory listing of potentialities, this text will delve not so much into the stimulating possibilities that arose with the advent of artificial intelligence in the fight against disease, epidemics, or pandemics, but, much more, into the critical considerations and cautions users should take into account as the use of these approaches is evermore pondered. We will examine the various challenges artificial intelligence faces before it can become, if ever, a beacon of hope in a world ever shadowed by the threat of widespread, generalized, and acute disease outbreaks.

In addition, the impact of AI on fundamental rights and, in particular, the use of AI to combat a pandemic and its relationship with the preservation of fundamental rights are also analysed.

1. The use of Artificial Intelligence in the Combat Against Pandemics and Why it Fails

However, this potential for good comes intertwined with challenges that must be addressed. George Box famously stated that “all models are wrong, but some are useful” (Box, 1979). It should come as no surprise that, despite its strengths, artificial intelligence models’ effectiveness is hampered by several limitations that can turn those solutions into a double-edged sword.

One of the major Achilles heels of artificial intelligence models is that those same models are only as good as the data they’re trained on. Inaccurate, incomplete, or biased data can lead to unreliable and potentially harmful outputs (Gianfrancesco et al., 2018). Limited access to real-time health data in some regions or privacy concerns can further restrict artificial intelligence’s capabilities.

Another one is commonly known as the “black box enigma” or, in a diluted formulation the “grey box” enigma. Understanding the inner workings of intricate artificial intelligence algorithms poses an extremely difficult task, impeding humans’ ability to grasp their

decision-making processes and this opacity can, in turn, undermine confidence and complicate efforts to detect and rectify underlying biases and various other types of problems ([Mittelstadt et al., 2016](#)).

Yet another one is the temptation of over-reliance on artificial intelligence models' significant capabilities without recognizing their fundamental limitations and, thus, relying too heavily on artificial intelligence without paying due attention to fundamental public health strategies such as contact tracing, vaccination efforts, and community awareness campaigns, all of them established approaches that must continue to play a vital role in effectively managing disease outbreaks ([Silva et al., 2022](#)).

2. Examples of Artificial Intelligence Failures: Lessons Learned and the Path Forward

The COVID-19 pandemic has been a testing ground for artificial intelligence in pandemic response, with mixed results. Some examples follow below.

Early in the pandemic, some artificial intelligence models drastically overestimated the spread of the virus due to limited initial data and the rapidly evolving situation. This led to unnecessary panic and resource allocation. In other instances, artificial intelligence powered chatbots designed to answer public health questions were overwhelmed by the surge in demand and provided inaccurate information in some cases. This highlights the need for robust training data and clear limitations set for artificial intelligence applications ([Bajwa et al., 2021](#); [Gürsoy & Kaya, 2023](#)).

It can, therefore, be argued that only by acknowledging the artificial intelligence limitations and focusing on responsible development, can humanity harness its power for a healthier future. Actions as prioritizing data quality and responsible data collection practices or addressing data bias and ensuring data privacy are crucial for building trustworthy artificial intelligence models. The development of robust methods to mitigate bias in artificial intelligence algorithms and techniques based on fairness testing and data augmentation can also help identify and address potential biases from the start. Investing in explainable artificial intelligence (also known as XAI) research can help stakeholders understand how artificial intelligence models arrive at their conclusions, fostering trust and enabling early detection of potential problems ([Jobin et al., 2019](#)).

The promotion of balanced approaches, where artificial intelligence is used alongside traditional public health interventions and complements, not replaces, established public health measures could be the path forward. By addressing these challenges and fostering responsible artificial intelligence development, agents can leverage the power of artificial intelligence to create a world better prepared for future pandemics ([Benke & Benke, 2018](#)). Artificial intelligence can be a powerful weapon in humanity's arsenal, but only if wielded wisely, as evidenced underneath.

3. The Problem with Models: Garbage In Garbage Out

The concept of «garbage in, garbage out» is a fundamental principle in artificial intelligence, particularly in machine learning ([Breiman, 2001](#)).

It helps to understand what are the main types of «garbage» data that exist. Firstly, data can be inaccurate. This includes errors in spelling, typos, factual mistakes, or outdated information ([Halevy et al., 2009](#)). One can easily imagine an artificial intelligence trained on news articles with many typos – it might struggle to understand language properly. Secondly, data can be incomplete. Missing values or data points can skew the model's understanding ([Little & Rubin, 2019](#)). For example, an artificial intelligence for predicting customer churn (why patients choose not to return to a hospital, for instance) might miss crucial data points if customer feedback isn't collected. Thirdly, data can be biased. Data that unfairly represents a certain group can lead to discriminatory outcomes ([Berk, 1983](#)). An artificial intelligence used for hiring decisions trained on resumes with mostly male applicants for nursing jobs might favor male candidates in the future. Fourthly and lastly, data can be irrelevant. Information not relevant to the task at hand can confuse the model ([Greiner et al., 1997](#)). An artificial intelligence for sentiment analysis (understanding emotions in text) in a psychiatric institution might be overwhelmed by irrelevant emojis in a dataset.

It also helps to understand what might be the consequences of «garbage in». First, there is the perpetuation of bias, as artificial intelligence can amplify existing societal biases if the training data reflects those biases ([Bazarkina & Pashentsev, 2020](#)). This can lead to unfair outcomes in areas like loan approvals, facial recognition, and criminal justice predictions. Next, there is the reduction of accuracy and reliability, as models trained on inaccurate data will produce unreliable outputs ([Shin & Park, 2019](#)). Imagine an artificial intelligence for pathology prediction trained on faulty temperature readings from patients – its diagnostics would be inaccurate. Furthermore, resources can be wasted, as time and money spent training models on bad data are significant resource drains ([Hulten, 2018](#)).

Most useful is to know the chief techniques for combating «garbage in». On one hand, it pays off to invest on data cleaning and curation. Techniques like data validation, error correction, and filtering are used to ensure data quality. This can be a labor-intensive process, but crucial for reliable artificial intelligence ([Wang & Shi, 2011](#)). On another hand, the use of data augmentation, that is, creating synthetic data to supplement existing datasets can help address issues like incomplete data ([Mumuni & Mumuni, 2022](#)). For example, generating realistic-looking images with diverse faces can help reduce bias in facial recognition. Another instance is the use of algorithmic bias detection methods to identify and mitigate bias in artificial intelligence algorithms themselves. This can involve analyzing the model's decision-making process to uncover hidden biases

(Kordzadeh & Ghasemaghaei, 2022). Explainable artificial intelligence constitutes another technique that focuses on making artificial intelligence models more transparent, allowing humans to understand how the model arrives at its conclusions, helping to identify potential biases or errors (Arrieta et al., 2020).

Addressing the «garbage in, garbage out» problem is critical for building trustworthy and ethical artificial intelligence in the future. As artificial intelligence becomes more integrated into everyone's lives, ensuring data quality and mitigating bias is essential (Jobin et al., 2019). In this regard, some efforts are well underway. Standardization and regulations won't solve it by themselves but can help. Developing guidelines and regulations for responsible artificial intelligence development and deployment can improve data quality and fairness¹. Public education and awareness are also being pursued. Raising awareness about the potential pitfalls of artificial intelligence and the importance of responsible development can foster public trust (Kandlhofer et al., 2023). Additionally, collaboration between artificial intelligence developers and experts, on an interdisciplinary basis, including data scientists, ethicists, and policymakers is crucial for building robust and responsible artificial intelligence systems (Bisconti et al., 2023). This is another trend on the rise that can prevent or moderate succumbing to the pitfalls of «garbage in, garbage out».

4. The Sustainability Problem (almost) no One Talks About

Sustainability of artificial intelligence solutions will be a key factor.

Energy consumption is paramount. Training a large language model like GPT-3 can consume the same amount of energy as several cars in their lifetime. Some studies estimate the energy consumption of training a single large language model to be around 1.5 MWh². Data centers housing artificial intelligence systems are estimated, conservatively, to consume 1% to 3% of global electricity³.

Also, water consumption is a growing concern. Data centers rely heavily on water for cooling, with estimates suggesting they use up to 1.7 billion gallons of water per year in the United States of America alone⁴. The water footprint of artificial intelligence can be significant, even for individual users. A single query on a large language model can require enough water to fill a small bottle⁵.

¹ European Commission. (2021). Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Legislative Instruments (COM(2021) 206 final). <https://clck.ru/3DzaGK>

² Luccioni, S. (2023, April 12). The mounting human and environmental costs of generative AI. Ars Technica. <https://clck.ru/3DzaKM>

³ AI Now Institute. (2023). Algorithmic Accountability: Moving Beyond Audits. <https://clck.ru/3DzaLM>

⁴ Meredith, S. (2023, December 6). A 'thirsty' generative AI boom poses a growing problem for Big Tech. CNBC. <https://clck.ru/3DzaLy>; Microsoft (2022). 2022 Environmental Sustainability Report. <https://clck.ru/3DzaLy>

⁵ Ibid.

However, data storage limitations will probably be the ultimate limiting factor (Susskind, 2020). The amount of data generated globally is growing exponentially, doubling roughly every two years. Current storage technologies like hard disk drives are reaching their physical limitations in terms of miniaturization and storage capacity⁶.

It's important to note that this data is constantly evolving as technology advances. Researchers are actively developing more energy-efficient artificial intelligence models, water-saving cooling systems for data centers, and new data storage technologies with higher capacities (Chen, 2016).

Predicting exactly when the entire Earth will be needed to store data is difficult. Data growth is exponential, but storage technology is also constantly evolving. However, it is safe to say the entire planet's physical space for data storage will not be needed in the near future. The world is witnessing a race between exponential growth and Moore's Law (Theis & Wong, 2017). Data creation is indeed growing exponentially, doubling roughly every two years. Even if the growth slows down to 20% per year (it is actually closer to 70%), it would still be an unsustainable growth rate, in the long term. Existing infrastructure and energy limitations would make it increasingly difficult to maintain such a pace⁷. However, storage capacity is also increasing rapidly, following a trend similar to Moore's Law (doubling of transistor density on integrated circuits roughly every two years). Data storage isn't a one-to-one process. Compression techniques can significantly reduce the physical space needed to store information. While data growth might outpace storage capacity at some point, advancements in storage technology like solid-state drives and advancements in data compression techniques can help bridge the gap (Chen, 2016). Arguments are also made that not all data needs forever storage in these systems. A significant portion of data doesn't require permanent storage. Logs, temporary files, and certain types of entertainment content can be deleted after a set period. Focusing on efficient, sustainable data management and prioritizing what gets stored permanently can significantly reduce storage needs (Arass & Souissi, 2018). Nonetheless, in this case one starts to consider that the supposed superintelligences are being brought down to human level. Less than perfect memory implies imperfect solutions, more mistakes made and less reliability: in other words, human like performance. Entirely new but yet distant new storage technologies might delay this inevitability, as researchers are exploring alternative solutions with much higher capacities than traditional hard drives. These include technologies like DNA Storage to achieve vast amounts of data stored in a very compact space (as all life form do in their genome). While still in its early stages, DNA storage holds immense potential for long-term data archiving (Goldman et al., 2013). They include holographic storage, as well, using laser technology

⁶ Rydning, D., Reinsel, J., & Gantz, J. (2018). The digitization of the world from edge to core. International Data Corporation. <https://clck.ru/3DzaNN>

⁷ Ibid.

to store data in three dimensions, offering much higher density than traditional methods (Lin et al., 2020).

Despite all this wishful thinking, Vopson (2020), estimates, in a convincingly rigorous calculation, for several information annual rate growth scenarios, how many years it would take for the entire mass of the Earth to be dedicated to data storage. His approach, as it is based on the number of atoms available, becomes mostly independent of storage efficiency management techniques and eventual, hypothetical new storage technologies. The values attained: 4500 years at 1% growth, 918 years at 5% growth, 246 years at 20% growth, and circa 110 years at 50% growth. The author refers to this eventuality as the impending information catastrophe. Considering that the crust of the Earth comprises only 0.7% of the planet's total volume and that that, even if wildly farfetched, is the part of the volume humanity has a more realistic chance to utilize in its entirety, humans are less than a century away (more likely 30 to 50 year away, or even less) from Vopson's information catastrophe. Studies by the AI Now Institute⁸ and the Stanford Institute for Human-Centered Artificial Intelligence⁹ might not agree on the exact numbers or not share the same approach but highlight the same trajectory.

5. Governance as Key: How State Measures and Data Availability Reinforce some Organizational Values and Contribute to the Sustainability of the National Health System

It can be argued that the missing link for adequately bridging the gap between traditional health measures and practices, and an artificial intelligence approach to pandemic response is good governance. State measures aligning to the principles of good governance are in a privileged position to become the foundation for artificial intelligence integration with more conventional methods.

Correia et al. (2020a) explore established public health measures that form the bedrock of a robust national health system during pandemics. For one, these authors address lockdowns, contact tracing, and vaccination campaigns as measures that can be implemented to slow the spread of viruses, protect vulnerable populations, and achieve herd immunity, embodying the value of prioritizing public health and demonstrating that government's responsibility should be towards its citizens. A second important point stressed by these authors is that data plays a crucial role in monitoring infection rates, tracking resource allocation, and understanding patient outcomes, informing decision-making, and that, in turn, reinforces the value of evidence-based practice, and ultimately contributes to the efficient use of resources within the health system.

⁸ AI Now Institute. (2023). Algorithmic Accountability: Moving Beyond Audits. <https://clck.ru/3DzaQC>

⁹ Stanford Institute for Human-Centered Artificial Intelligence. (2023). Sustainability and AI. <https://clck.ru/3DzaRd>

A path, then, emerges. The path of sustainability through efficiency, where artificial intelligence can amplify the impact of traditional measures. While conventional measures are still essential and probably will always be so, artificial intelligence offers the potential to significantly enhance their effectiveness and further contribute to the sustainability of national health systems. One immediate application of this idea can be materialized in the use of artificial intelligence models to analyze historical data, identify patterns, and predict the emergence or spread of future disease outbreaks, epidemics, and pandemics, allowing public health officials to take proactive measures like early warning systems, stockpile of vital supplies, and the strategic deployment of resources. In particular, this optimization of resource allocation by means of artificial intelligence algorithms can be readily put to use in the analysis of real-time data on infection rates, hospital capacity, and material resources availability, allowing for the dynamic allocation of medical staff, equipment, and critical supplies to the areas facing the biggest strain (Correia et al., 2021, 2022), and, therefore, ensuring efficient resource management. More advanced and, consequently, delicate applications encompass personalized treatment plans that require the analysis of individual patient data like medical history and genetic makeup, where artificial intelligence can potentially assist medical professionals in tailoring treatment plans for maximum effectiveness, contributing to faster recovery times, improved patient outcomes, and reduced strain on healthcare resources (Jiang et al., 2017).

It becomes obvious, then, that the effectiveness of governance in pandemic response (whether making use of artificial intelligence or not), hinges on the availability of high-quality, comprehensive data, data gathered through traditional measures and methods like contact tracing and patient records (Wu et al., 2022). However, if governance is good, it will address data privacy concerns regarding the collecting and use of patient data, including for artificial intelligence development, not neglecting the assurance of data anonymization and robust data security protocols to maintain public trust (Smidt & Jokonya, 2021).

Effective use of artificial intelligence in public health requires seamless data sharing between different healthcare institutions and interoperability (O'Reilly-Shah et al., 2020). The need for standardized data formats and secure communication channels to facilitate this data exchange is paramount (Sass et al., 2020).

In conclusion, a symbiotic relationship for sustainable health systems can be implemented. Traditional public health measures, data availability, and artificial intelligence are not separate entities but can be interconnected elements in the fight against pandemics. The existing data infrastructure and experience with traditional measures create a fertile ground for artificial intelligence integration (Baclic et al., 2020). By leveraging the power of artificial intelligence in conjunction with established practices, national health systems can achieve greater efficiency, personalize treatment approaches,

and ultimately ensure their long-term sustainability in the face of future occurrences (Gunasekeran et al., 2021). That is equivalent to affirm that robust management practices and sound organization values can pave the way for future artificial intelligence integration in this crucial domain.

6. Governance as Key: How Governance Reinforces some Organizational Values and Contributes to the Sustainability of Crisis Management

The statement above holds immense significance for the context of artificial intelligence applied to the fight against pandemics. That is because effective governance practices provide the framework and guiding principles for utilizing artificial intelligence responsibly and ethically in crisis management, with pandemics being a prime example.

Governance emphasizes open communication and holding decision-makers accountable for their actions. This is vital for building public trust in artificial intelligence powered solutions used during pandemics, like contact tracing apps. Clear explanations about how artificial intelligence is being used and how data privacy is protected are crucial to avoid public apprehension (Galetsi et al., 2022). Good governance also fosters collaboration, including data sharing, and coordination between diverse stakeholders, including resource allocation, vaccine development, and communication strategies. This collaboration and cooperation comprise government agencies, healthcare institutions, research bodies, and private sector entities (Bulled, 2023).

Good governance translates into specific organizational values that have the potential to shape artificial intelligence development and use in pandemics. Well applied, it promotes equity and fairness, adequate distribution of resources and strives to bridge the digital divide. This, in turn, can ensure that artificial intelligence tools do not exacerbate existing social inequalities (Margetts, 2022). For example, artificial intelligence powered contact tracing apps should be accessible to all demographics and should not unfairly target certain populations. Governance can also be determinant in establishing robust data privacy and security protocols. These types of actions can contribute to protects citizens' information while allowing for responsible data collection and use for artificial intelligence development in pandemic response. Striking a balance between data-driven insights and data security is crucial during a pandemic (Zhang et al., 2022). Moreover, good governance fosters evidence-based decision making, a culture of relying on data and scientific evidence to inform decisions. This aligns perfectly with the core principle of artificial intelligence, which uses data analysis to generate insights and recommendations for, amongst others, public health officials (Rubin et al., 2021).

In addition, effective and sustainable pandemic response requires a forward-looking approach and long-term planning. Good governance practices contribute to the sustainability of crisis management in several ways. Governance is tasked with ensuring long-term investment in infrastructure and maintaining the hardware, software, and expertise needed for artificial intelligence development and deployment in public health. This includes investment in research and development, training programs for artificial intelligence specialists within healthcare institutions, and establishing robust data management systems (Balog-Way & McComas, 2022). Governance also promotes the futureproofing of strategies, the development of flexible frameworks that can adapt to evolving threats and pandemics with novel characteristics. This ensures that artificial intelligence remains relevant and useful for future public health challenges. For example, artificial intelligence algorithms for pandemic prediction need to be adaptable to handle new virus strains and variations. Good governance has, as well, the ability to build and strengthen public trust in government institutions and their use of artificial intelligence during a pandemic (Romano et al., 2021). This trust fuels cooperation with artificial intelligence suggested initiatives, like contact tracing and symptom-tracking apps.

Correia et al. (2020b), explore the principles that create a strong foundation for solutions in pandemic response. By fostering collaboration, prioritizing ethical values, and ensuring long-term sustainability, governance practices pave the way to become a powerful weapon in the arsenal for combatting pandemics and building a more resilient future for public health. The authors propose a six-dimensions, eight-hypothesis model (already validated for very specific circumstances) represented on figure 1.

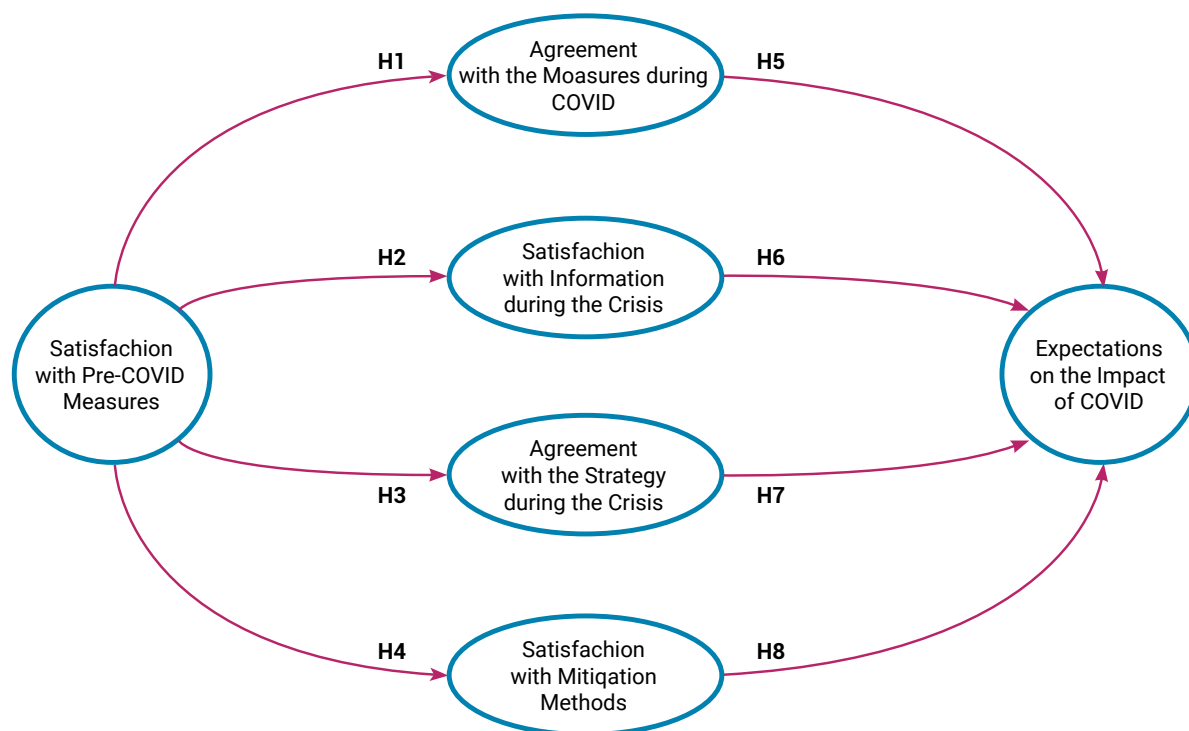


Figure 1. Crisis management: COVID-19 structural model (Correia et al., 2020b)

It can be reasoned that the incorporation on these inputs (models, dimensions and relations between dimensions), the human factors, if you will, is the link that has been missing in the creation of a supportive environment, through effective governance, for artificial intelligence to be leveraged responsibly and ethically, while also increasing their performance in predicting, preparing and warning for outbreaks, in predicting and preparing for individuals' responses to public health measures, in optimizing real time strategic resource allocation based on infection rates, in developing targeted, personalized interventions for high-risk individuals, and in tracking, containing, isolating and limiting spread of pathogens.

The linkage of good governance, crisis management models and artificial intelligent solutions, and the synergies thus generated hold immense potential to improve our preparedness and response to future pandemics, ultimately saving lives and ensuring a more sustainable future for global health.

7. Critical Reflections

Artificial intelligence represents one of the greatest technological innovations of the modern era, with the potential to fundamentally transform society in many aspects (Bostrom, 2014). However, this transformation also brings with it significant challenges and concerns about the fragility of modern human societies.

Digital inequality presents a pressing concern as the integration of artificial intelligence into various sectors may widen socioeconomic gaps, accentuating disparities between individuals equipped with access and proficiency in leveraging technological advancements and those marginalized without such resources (Eubanks, 2018). This phenomenon has the potential to intensify preexisting inequities while giving rise to novel manifestations of digital exclusion.

The advent of artificial intelligence driven automation also poses a significant risk of displacing numerous conventional occupations, particularly those characterized by routine and predictable duties. This transition holds the potential to precipitate widespread job loss, plunging affected individuals into a state of mass unemployment and instigating profound existential crises rooted in the displacement by technological innovations. The health sector appears to be no exception to this peril (Hazarika, 2020).

Moreover, artificial intelligence algorithms, especially in crucial domains such as justice, healthcare, and finance, are prone to manipulation and bias, posing a risk of unfair and harmful consequences, particularly for marginalized and vulnerable communities (Obermeyer, 2019).

Furthermore, concerns regarding data privacy and security are heightened by the extensive collection and analysis of data powering artificial intelligence algorithms, with lack of transparency and control over data usage posing as a predictable consequence and a significant threat to public trust in technology, public policy, and public institutions (Larsson & Heintz, 2020).

Additionally, societies around the world have to deal, increasingly, with polarization and misinformation, manifested by the so-called misuse of digital platforms driven by artificial intelligence, that can undermine social cohesion and erode trust in democratic institutions, leading to an increasingly fragmented and divisive society (Kavanagh & Rich, 2018).

One must not ignore, as well, the potential dangers of becoming overly reliant on technology. The more humanity depends on artificial intelligence for decision-making and task completion, the weaker human's ability to function independently becomes (Bostrom, 2014). This vulnerability to disruptions and systemic failures in artificial intelligence could have devastating consequences.

Also, artificial intelligence is creeping in, or, in a more technical terminology, the erosion of human autonomy and agency is ever more noticeable (Ettliger, 2022). The increasing integration of artificial intelligence into people lives could lead to the erosion of human autonomy and agency, as we increasingly leave important decisions to automated systems. This raises questions about who controls technology and who can one trust to make decisions that affect one's life.

Finally, this first layer, superficial challenges analysis must include not only immediate concerns, but the often-mentioned existential risks, that comprise long-term fears about the development of artificial intelligence, including the common scenarios of artificial superintelligences that surpass human control and threaten the survival of humanity (Bostrom, 2014).

In essence, the convergence of artificial intelligence with the delicate nature of contemporary human societies prompts profound inquiries into ethics, governance, fairness, and human principles. A comprehensive and cooperative approach is vital to tackle these matters, ensuring that artificial intelligence advancement and implementation prioritize human welfare and enduring sustainability (Jobin et al., 2019).

However, it is possible to add several layers of depth in a conscious quest to use artificial intelligence in a responsible, productive, and secure manner.

Artificial intelligence appears to be quite fragile when stressed in just the right way. When stressed in just the right way, artificial intelligence systems can indeed exhibit fragility or vulnerability. This fragility can manifest in various ways, depending on the context and nature of the artificial intelligence systems. For example, adversarial attacks can happen, involving intentionally manipulating input data to artificial intelligence systems in a way that causes them to make mistakes or produce incorrect outputs. These attacks can exploit vulnerabilities in the artificial intelligence algorithms, such as deep learning neural networks, leading to unexpected and potentially harmful behavior (Ruan et al., 2021). Another example, similar to one previously addressed through a different lens, is the use of unrepresentative data to train artificial intelligence models, leading to the above-mentioned biased decisions or predictions. This bias can be exacerbated under certain conditions or when the artificial intelligence system encounters new, unseen data that

differs significantly from the training data sets. As a result, the artificial intelligence system may fail to generalize effectively, leading to fragility in its performance (Navigli et al., 2023). One more example consists in what can be called catastrophic forgetting, as some artificial intelligence systems, particularly those based on artificial neural networks, can exhibit catastrophic forgetting when exposed to new data. This phenomenon occurs when the artificial intelligence system forgets previously learned information as it learns new information, leading to a loss of performance or accuracy over time. This fragility can limit the system's ability to adapt to changing environments or tasks (Kirkpatrick et al., 2017). Yet another example is model fragility, given that artificial intelligence models can be fragile to small changes in input data or model parameters. For example, slight perturbations to input images can cause image recognition models to misclassify objects, leading to potentially dangerous consequences in applications such as autonomous vehicles or medical diagnosis (Chen et al., 2020). One final example of artificial intelligence fragility is the inherent system complexity. As artificial intelligence systems become more complex and interconnected, they can become increasingly fragile to disruptions or failures in individual components. A failure in one part of the system can cascade into other parts, leading to system-wide failures or breakdowns (Chen et al., 2020). Complexity is a double-edged sword as many systems derive their power from their complexity and ability to process vast amounts of data. However, this complexity can also be a vulnerability, as it increases the surface area for potential attacks and makes the system more difficult to understand and secure. This fragility underscores the importance of robustness and resilience in artificial intelligence system design and deployment. Addressing the fragility of artificial intelligence systems requires careful attention to design, testing, and validation processes, as well as ongoing monitoring and maintenance. It also highlights the need for transparency, accountability, and ethical considerations in the development and deployment of artificial intelligence technologies. By addressing these challenges, individuals, institutions, and societies can work towards creating artificial intelligence systems that are more robust, reliable, and trustworthy in a wide range of applications. Just imagine the pandemonium that would result of such types of fragilities occurring separately or in chain, in the health sector, during an epidemic or pandemic crisis.

An additional convolution surfaces when one comprehends that artificial intelligence pathways, even if they appear robust, can often be undermined by the presence of choke points, which are critical junctures where failure or disruption can have cascading effects on the entire system. This duality, where robustness and fragility coexist, is inherent to many complex artificial intelligence systems (Zhou et al., 2024). For instance, as seen before, artificial intelligence models often involve complex networks of interconnected components, such as layers in deep neural networks or nodes in graph-based models. While this interconnectedness can enhance robustness by allowing for redundancy and fault tolerance, it also introduces choke points where failure or disruption in a critical

component can propagate throughout the network (Villegas-Ch et al., 2024). Another instance is the emergence of critical dependencies. Certain components of artificial intelligence applications may serve as critical dependencies, upon which the functionality of the entire system relies. These choke points can include specific layers or nodes in neural networks that play pivotal roles in processing or decision-making. If these critical dependencies fail or malfunction, it can lead to a breakdown in the system's performance (Macrae, 2022). Still another instance is the tendency of artificial intelligence usages to be sensitive to input data, especially in uses such as image recognition or natural language processing. Small perturbations or adversarial inputs at choke points within the pathway can lead to significant changes in the system's output. This sensitivity underscores the fragility of the pathway to specific types of input manipulation (Dhingra & Gupta, 2017). One final instance that illustrates the dangers of choke points are trade-offs in Design. These trade-offs are between robustness and efficiency. Strategies aimed at enhancing robustness, such as adding redundancy or error correction mechanisms, may introduce additional choke points or computational overhead. Conversely, optimizations for efficiency may inadvertently increase the system's fragility by reducing redundancy or resilience. Addressing these specific problems requires a multi-faceted approach that involves identifying and mitigating choke points and enhancing robustness through redundancy and diversity (Goodfellow et al., 2016). By understanding these delicate balances, researchers, engineers and practitioners can work towards creating more resilient and trustworthy artificial intelligence technologies.

Artificial intelligence intricacy expands when the concept of losing control acquires a broader and more nuanced meaning. That is to say, when the concept reflects the complexities and challenges associated with the development and deployment of artificial intelligence technologies. One such manifestation relates to autonomy and decision-making (Wallach & Allen, 2008). As artificial intelligence systems become increasingly autonomous and capable of making decisions without direct human intervention, there is a concern about losing control over the outcomes of these decisions. This can be particularly relevant in high-stakes applications such as autonomous vehicles, or medical appliances where the actions of artificial intelligent systems can have real-world, life-threatening consequences. Another manifestation is the inclination of artificial intelligent systems, especially those based on deep learning and neural networks, to be highly opaque, making it difficult for humans to understand or interpret their internal workings. This lack of transparency can lead to a loss of control over how artificial intelligent systems arrive at their decisions, raising concerns about accountability and trust (Chiao, 2019). One other manifestation is patent in the fact that artificial intelligence systems can exhibit emergent behavior, where complex patterns or behaviors arise from the interactions of simple components. This emergent behavior can be difficult to predict or control, leading to uncertainty about the behavior of the systems in novel or unanticipated situations. That, in itself, can lead to unintended consequences, as the actions or decisions of artificial

intelligence functions produce outcomes that were not anticipated or intended by their creators. This can occur due to unexpected interactions with the environment, or a range of other factors (Bostrom, 2014). Last but not least, the ethical and societal implications must be considered. Losing control over artificial intelligent technologies can also bear broader ethical and societal implications, such as the impact of artificial intelligence on employment, privacy, security, and inequality, as addressed in the course of this text (Thomsen, 2019). These concerns highlight the need for responsible artificial intelligence development and governance to ensure that artificial intelligent technologies are deployed in ways that benefit society as a whole and not only oligarchies, big tech companies and the ruling elites. Addressing these issues requires a holistic approach that encompasses technical, ethical, and regulatory considerations. This includes promoting transparency and accountability in artificial intelligent systems and engaging in ongoing dialogue and collaboration between stakeholders to mitigate risks and maximize the benefits.

Another complication rises as one ponders on the premise of reliability, particularly during times of stress or uncertainty. This premise may warrant urgent reevaluation. One illustration of ill placed premises refers to interconnectedness, as artificial intelligence systems often comprise intricate networks of interconnected components, each contributing to overall functionality. When stress or unexpected conditions arise, such as adversarial attacks, data anomalies, or environmental changes, the complexity of these systems can amplify the likelihood of failure. This underscores the need for a more nuanced understanding of reliability beyond traditional measures (Macrae, 2022). Another illustration is unpredictability. The emergent behavior exhibited by artificial intelligent systems can lead to unpredictability in their responses to stressors. Even minor perturbations or variations in input data can trigger unexpected outcomes, highlighting the challenges of ensuring reliability under diverse conditions. This unpredictability underscores the importance of robustness testing and scenario planning to identify and mitigate potential failure points (Bostrom, 2014). These exact concerns were expressed, in different combinations, above. Yet another illustration of this class of phenomena deals with adaptive and evolving environments, given that artificial intelligence systems operate within dynamic and everchanging settings, where conditions may change rapidly and unpredictably. In such environments, the notion of reliability as a static attribute becomes inadequate. Instead, reliability must be viewed as a dynamic property that adapts to changing circumstances, requiring continuous monitoring, adaptation, and feedback mechanisms (Sundar, 2020). One final illustration focusses on the closely intertwined relation between these types of systems and the human-machine interaction. Human operators play a critical role in monitoring system performance, interpreting outputs, and intervening when necessary. However, under stress or high-pressure situations, human operators may also be prone to errors or cognitive biases, further complicating the reliability of the systems (Hoff & Bashir, 2015). In light of this set of challenges, rethinking the premise of reliability in artificial intelligence necessitates a shift towards more adaptive, resilient,

and context-aware approaches. This may involve incorporating principles of uncertainty quantification, robustness engineering, and human-centered design into the development and deployment of artificial intelligence models and applications. By embracing a broader understanding of reliability and proactively addressing the factors that contribute to failure, humanity can strive towards more trustworthy and dependable artificial intelligence technologies. Preparing and having contingency, artificial intelligence free, plans that allow society to continue functioning in case of technological crisis or collapse should be of the utmost priority. Would most, if not all, developed countries still have a justice system tomorrow, if the internet failed now that justice has been “dematerialized”? Do we really want to take such risks, ones that may bring societies to a halt or to altogether collapse?

A deeper yet layer, one more hurdle to be surpassed is the necessity to shift from a premise of reliability to one of risk regarding artificial intelligence. That is equivalent to a move from fixity to nimbleness in responding to changing circumstances. One good example in the understanding of risk. Reliability is often associated with the notion of deterministic outcomes and predictable behavior. However, in complex and dynamic environments, such as those encountered by artificial intelligence systems, complete reliability may be unattainable. Recognizing this, the focus shifts towards understanding and managing risk – the likelihood and impact of adverse events or uncertainties (Bigham et al., 2019). This embracing of risk implies acknowledging the inherent uncertainty in artificial intelligence systems and embracing adaptive strategies to cope with it. Rather than striving for absolute reliability, artificial intelligence systems should be designed to be resilient and responsive in the face of changing circumstances. This may involve incorporating mechanisms for real-time monitoring, dynamic adjustment, and learning from experience (Syed et al., 2023). This nimbleness in responding to changing circumstances requires agility and flexibility in these applications, incorporating the ability to quickly assess risks, identify opportunities, and adapt behavior or decision-making strategies accordingly. Agile artificial intelligence systems will be increasingly capable of dynamically allocate resources, prioritize tasks, and adjust to new information or objectives as they arise. Shifting from a mindset of reliability to a mindset of risk necessitates the development of robust risk management frameworks. These frameworks must provide a systematic approach to identifying, assessing, mitigating, and monitoring risks throughout the lifecycle of artificial intelligence usages. By proactively managing risks, organizations can enhance resilience and reduce the likelihood of adverse outcomes (Jobin et al., 2019). Naturally, a risk-aware artificial intelligence use needs to be centered in a continuous learning and improvement capacity, leveraging feedback loops, experimentation, and data-driven insights to iteratively enhance performance and adapt to evolving challenges. This iterative approach will increasingly allow artificial intelligence systems to refine their strategies over time and become more effective in managing risks. The final piece of this particular puzzle will be put in place by recognizing the limitations of artificial intelligent

systems in navigating complex and uncertain environments, and, as a consequence, increasing emphasis on human-in-the-loop approaches (Russell & Norvig, 2021). By integrating human judgment, expertise, and oversight, artificial intelligent systems can more adequately complement human decision-making, mitigate risks, and enhance overall system performance (Parasuraman & Riley, 1997). In general, the shift from a premise of reliability to one of risk reflects a broader recognition of the inherent uncertainties and complexities of real-world applications. By embracing risk and fostering nimbleness in responding to changing circumstances, artificial intelligent models can better navigate uncertain terrain, adapt to evolving challenges, and ultimately deliver greater value and impact in diverse domains, to which healthcare is no stranger.

The next in-depth layer is one in which artificial intelligence can appear to disguise weaknesses as strengths, especially when it comes to certain types of machine learning models or algorithms. The most blatant example is the imbalance between generalization and overfitting, occurring when a model learns to perform well on the training data but fails to generalize to new, unseen data (Igual & Seguí, 2024). This can give the illusion of strength because the model appears to perform exceptionally well on the data it was trained on. However, when exposed to new data, the weaknesses of the model become apparent as it fails to make accurate predictions. A visual recognition model can easily learn to identify ties, on pictures, and associate that to males, if trained with a biased Wall Street executive data set. This can happen because artificial intelligent models aim to generalize patterns from training data to make predictions on unseen data. While strong generalization is desirable, over-reliance on specific patterns in the training data can lead to overfitting, where the model fails to generalize effectively. Recognizing the balance between generalization and overfitting is crucial for ensuring robustness. Thoroughly testing artificial intelligence systems on diverse datasets, scrutinizing their decision-making processes, and mitigating biases and vulnerabilities, allows users to uncover, and address weaknesses disguised as strengths, leading to more reliable and trustworthy solutions.

As the analysis keeps going deeper, it arrives at the realization that artificial intelligence cultivates instability. A collective that has come to expect that things have and will operate without interruption and that blockages are either serendipitous or negligible in impact is, undoubtedly, much less prepared when these principles are threatened. The dependency on artificial intelligence is becoming increasingly integrated into various aspects of society and there is a growing dependency on their functionality. Individual, organizations, governments, and supranational institutions (like the United Nations and the World Health Organization) rely evermore on artificial intelligence for decision-making, automation, and optimization of processes (Bostrom, 2014). However, this dependency can create instability if these functions experience failures or disruptions. This widespread adoption of artificial intelligent solutions may foster a collective expectation of continuity and seamless operation. When these solutions function as expected,

they reinforce the perception that contrarities are minimal. However, this expectation can lead to complacency and vulnerability if systems encounter unexpected challenges or malfunctions (Parasuraman et al., 2000). It follows that, when artificial intelligence applications fail or encounter blockages, the impact can be significant, especially if they are relied upon for critical tasks or services. Disruptions in technological-driven processes can disrupt supply chains, financial markets, communication networks, and other essential functions and critical systems (like nuclear power stations), leading to economic losses, social unrest, or even safety risks. Cultivating resilience and adaptability in the face of this potential instability requires proactive measures to anticipate and mitigate risks, as mentioned prior. This may involve diversifying technological dependencies, building redundancy into critical systems, and developing human-centric approaches to decision-making and problem-solving (Bigam et al., 2019; Jobin et al., 2019). In summary, artificial intelligence technologies also pose challenges related to stability and resilience. By recognizing the potential for instability inherent in these systems and taking proactive measures to address risks, stakeholders can better navigate the complexities of an artificial intelligence driven world and build more robust and sustainable systems.

At last arrived at the deepest place, the center of and coldest place in, the universe, according to Aristotle, the center of Earth and Hell, according to Dante, one must contemplate artificial intelligence in the light of what can be called Luciferian semiotics, a voyage into the symbolic or metaphorical implications of artificial intelligence. In various mythologies and belief systems, Lucifer (Luciferian symbolism) is often associated with themes of rebellion, enlightenment, and the pursuit of knowledge. Lucifer is often depicted as the carrier of the light (Hanegraaff, 2013). The term Luciferian may thus connote the pursuit of knowledge or power that challenges established norms or authority structures. Semiotics refers to the study of signs and symbols and their interpretation. In the context of artificial intelligence, semiotics encompasses the symbolic meanings associated with artificial intelligence, including notions of intelligence, autonomy, and control (Binder, 2024). Artificial intelligence is often perceived as a symbol of strength and capability, given its ability to process vast amounts of data, make complex decisions, and automate tasks with efficiency. This perception of strength may be reinforced by the impressive feats accomplished by these solutions in various domains. However, as shown throughout the text, objects or entities that appear strong may, in fact, possess vulnerabilities or weaknesses that are not immediately apparent. This reversal of expectations can be seen as a manifestation of the Luciferian symbolism, where the pursuit of knowledge or power leads to a reevaluation of established truths or assumptions. Artificial intelligence models and systems, despite their perceived strength, can exhibit vulnerabilities or limitations in certain contexts. These weaknesses may become more visible over time as artificial intelligence technologies are subjected to scrutiny, experimentation, and real-world deployment. The exploration of the

Luciferian semiotics raises broader ethical and philosophical questions (Bostrom, 2014) about the nature of power, knowledge, and control in the age of artificial intelligence. It prompts reflections on the unintended consequences of technological advancement and the need for responsible stewardship of all technologies. In summary, the Luciferian semiotics applied to artificial intelligence invites us to consider the symbolic meanings and implications of artificial intelligence, including the ways in which perceptions of strength and power may be subverted or challenged by deeper exploration and understanding. It underscores the importance of critical inquiry and ethical reflection in navigating the complexities of artificial intelligence and its impact on societies, public policies, and political systems.

Given all the above, the use of artificial intelligence in the combat against pandemics, *quid juris*? Maybe, probably, humanities best course of action is to continue to rely on the human factor. Humans make more errors, there is no question about that. But most of those errors are small and inconsequential. They may originate individual tragedies but not global ones. Artificial intelligent models to deal with outbreaks, epidemics, and pandemics, as well as other artificial intelligence reliant medical applications might be almost error free. But that one error may doom all.

8. AI and Fundamental Rights

The impact of AI on fundamental rights is so relevant that it has not gone unnoticed by the AI Act, which in article 27 requires that high-risk systems must be subject to an impact assessment on fundamental rights. The main purpose of the impact assessment of AI systems is to identify and mitigate the potential risks that these systems may pose to people's fundamental rights. This is especially important when it comes to high-risk AI systems, which have the potential to significantly affect people's lives and well-being. To summarise, in view of the negative impacts of AI, we have not opted to abandon AI, but rather, in order to take advantage of its positive impacts, to classify AI systems according to risk and through their prior, concomitant and a posteriori control.

There are several relationships that can be established between AI and fundamental rights, in particular the impact that AI can have on the fulfilment of fundamental rights and, in a different sense, the impact that AI can have on the violation of fundamental rights. In any case, it is reasonable to believe that, as the FRA¹⁰ points out, even in a restricted context, the lack of a large body of empirical data on the wide range of rights involved in AI makes it fragile to provide the necessary safeguards to ensure that the use of AI is effectively in line with fundamental rights.

¹⁰ FRA – European Union Agency for Fundamental Rights, *Getting the future right – Artificial intelligence and fundamental rights* – Report, Publications Office of the European Union, 2020.

The main argument in favour of using AI is efficiency (Pedro, 2023). As far as the challenges are concerned, the main concerns are the violation of fundamental rights. Thus, among the main candidates for fundamental rights potentially harmed by AI are the right to the protection of personal data (Gómez Abeja, 2022) and the right to non-discrimination (Gómez Abeja, 2022), the right to effective judicial protection (Shaelou & Razmetaeva, 2023), the right to freedom of information, the right to suffrage and the right of access to public information (Gómez Abeja, 2022).

Returning to the work of the FRA¹¹, the use of AI can have an impact on fundamental rights, imposing the need to guarantee the non-discriminatory use of AI (right to non-discrimination); the requirement to process data lawfully (right to personal data protection); and the possibility of lodging complaints about AI-based decisions and lodging appeals (right to an effective remedy and to an impartial tribunal).

Finally, it should also be emphasised that the relationship between AI and fundamental rights can be richer, at least in the following dimensions: confronting implicit fundamental rights (Gómez Colomer, 2023), such as the principle of the rule of law and the principle of the natural judge, and the emergence of «new» or «renewed» fundamental rights (Shaelou & Razmetaeva, 2023), such as the right to be forgotten (Gómez Abeja, 2022), the «right not to be subject to automatic decisions and automatic treatment» in the broad sense (Shaelou & Razmetaeva, 2023); the «right to influence your digital footprint» (Shaelou & Razmetaeva, 2023), and new rights, such as the 'right not to be manipulated', the 'right to be informed neutrally online' and the 'right to meaningful human contact', the 'right not to be measured, analysed or trained' (Shaelou & Razmetaeva, 2023).

9. Pandemics and Fundamental Rights

The pandemic situation, as happened with COVID-19, called for the admission of public legal regimes of exceptionality (alongside normality regimes), which is nothing new – before all time (Gomes & Pedro, 2020) – just by looking at the Latin brocardo «Necessitas non habet legem, sed ipsa sibi facit legem». It was this brocardo that justified extraordinary powers in Roman law, exercisable in cases where it was necessary to deal with an unforeseeable situation that required an immediate decision, with no possibility of postponement.

The need for a legal system of exceptionality, and therefore its mobilisation, has become more evident in recent times. The configuration of the current risk society (Beck, 1986) and the fact that we live in a globalised world (economically and socially) in which, despite physical distance, everything seems to be close by, as the (still) current health crisis caused by the outbreak of COVID-19 - which in the space of a few months has spread from its source (China – Wuhan city) to the whole world (Pedro, 2022) – has greatly contributed to this.

¹¹ Ibid.

In the face of disasters of this kind, public law could not and cannot remain indifferent, in other words, given the damaging effects that public disasters have on the «salus populi», it is easily understandable that public bodies must use all the means at their disposal to restore normality (Alvarez Garcia, 1996). Therefore, in order to guarantee the rule of law, it is essential to provide for regimes that are flexible enough to respond to public interests that are threatened – regimes that make it possible to respond to states of public necessity, or, in other words, public law regimes of exceptionality.

Within a framework of real normality, public law is governed by the principle of the legality of public action – which therefore corresponds to a framework of legal normality. The problem arises whenever reality temporarily changes in a radical way, creating situations of imminent or real danger for the community and in which the public law of normality does not offer an adequate response, and the idea of maintaining the democratic rule of law imposes the need for exceptional legal regimes to come into play – «jus extremae necessitatis» – so that normality is restored in the short term and the legal regimes of normality return to force. What is at stake is an alternative legality, an exceptional legality of exceptionality (Correia, 1987) – a substitute and temporary legality.

Thus, as a rule, in exceptional situations, a state of siege or a state of emergency, respecting the principle of proportionality, some fundamental rights can be suspended. Despite this permission, it should be noted that not all fundamental rights can be suspended, as is the case with the rights to life, personal integrity, personal identity, civil capacity and citizenship, the non-retroactivity of criminal law, the right of defence of defendants and freedom of conscience and religion.

10. Possible Relationships between AI and Fundamental Rights in Combating Pandemics

In democratic states governed by the rule of law, the consideration of the use of AI to combat pandemics generally involves respect for fundamental rights. This requires, on the one hand, consideration of the impact that the use of AI has on certain fundamental rights, taking into account the risks that each specific AI system entails and, on the other hand, that the context of a pandemic, as happened with COVID-19, calls for a legality of exception, which may allow for the restriction of certain fundamental rights, with a view to safeguarding values such as public health, in order to restore a situation of normality.

Conclusions

Struggling diseases and pandemics today requires a comprehensive approach, uniting the efforts of doctors, medical organizations and states. Artificial intelligence is a promising tool, capable of radically changing the methods of counteracting epidemics and pandemics. Its potential lies in analyzing big data, forecasting disease outbursts, accelerating development of medication, personalization of treatment and optimization

of resources distribution. Examples of AI use, such as early detection of outbreak through analysis of social networks data, facilitation of search for medication and improved contact tracing, demonstrate its significance in struggling global threats to health.

However, introduction of AI into healthcare is accompanied by a number of challenges. These include data bias problems, algorithm complexity, risks of excessive dependence on technologies, and ethical dilemmas related to fundamental rights. Using AI to struggle pandemics requires observing a balance between innovations, ethics and human rights protection, including the right to privacy, freedom and equal access to medical assistance.

Hence, AI, despite its revolutionary capabilities, is not a panacea. Its use should be accompanied by a critical analysis of potential risks and requires developing legal and ethical mechanisms to ensure safe and fair use of technologies. Only after addressing these aspects, AI may become an effective tool to struggle diseases without threatening fundamental rights and freedoms.

References

- Arass, M., & Souissi, N. (2018). Data lifecycle: from big data to SmartData. In *2018 IEEE 5th International Congress on Information Science and Technology* (pp. 80–87). IEEE. <https://doi.org/10.1109/CIST.2018.8596547>
- Alvarez Garcia, V. (1996). *El concepto de necesidad en derecho público* (1st ed.). Madrid: Civitas. (In Spanish).
- Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Baclic, O., Tunis, M., Young, K., Doan, C., Swerdfeger, H., & Schonfeld, J. (2020). Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, 46(6), 161–168. <https://doi.org/10.14745/ccdr.v46i06a02>
- Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthcare Journal*, 8(2), e188–e194. <https://doi.org/10.7861/fhj.2021-0095>
- Beck, U. (1986). *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp Verlag.
- Balog-Way, D., & McComas, K. (2022). COVID-19: Reflections on trust, tradeoffs, and preparedness. In *COVID-19* (pp. 6–16). Routledge.
- Bazarkina, D. Y., & Pashentsev, E. N. (2020). Malicious use of artificial intelligence. *Russia in Global Affairs*, 18(4), 154–177. <https://doi.org/10.31278/1810-6374-2020-18-4-154-177>
- Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *International Journal of Environmental Research and Public Health*, 15(12), 2796. <https://doi.org/10.3390/ijerph15122796>
- Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48(3), 386–398. <https://doi.org/10.2307/2095230>
- Bigham, G., Adamtey, S., Onsarigo, L., & Jha, N. (2019). Artificial Intelligence for Construction Safety: Mitigation of the Risk of Fall. In K. Arai, S. Kapoor, R. Bhatia (Eds.). *Intelligent Systems and Applications*. Springer. https://doi.org/10.1007/978-3-030-01057-7_76
- Binder, W. (2024). Technology as (dis-)enchantment. AlphaGo and the meaning-making of artificial intelligence. *Cultural Sociology*, 18(1), 24–47. <https://doi.org/10.1177/17499755221138720>
- Bisconti, P., Orsitto, D., Fedorczyk, F., Brau, F., Capasso, M., De Marinis, L., ... & Schettini, C. (2023). Maximizing team synergy in AI-related interdisciplinary groups: an interdisciplinary-by-design iterative methodology. *AI & Society*, 38(4), 1443–1452. <https://doi.org/10.1007/s00146-022-01518-8>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Box, G. (1979). Robustness in the strategy of scientific model building. In R. Launer & G. Wilkinson (Eds.), *Robustness in Statistics* (pp. 201–236). Academic Press. <https://doi.org/10.1016/B978-0-12-438150-6.50018-2>
- Breiman, L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199–231. <https://doi.org/10.1214/ss/1009213726>

- Bulled, N. (2023). "Solidarity:" A failed call to action during the COVID-19 pandemic. *Public Health in Practice*, 5, 100379. <https://doi.org/10.1016/j.puhip.2023.100379>
- Chen, A. (2016). A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electronics*, 125, 25–38. <https://doi.org/10.1016/j.sse.2016.07.006>
- Chen, J., Zhang, R., Han, W., Jiang, W., Hu, J., Lu, X., Liu, X., & Zhao, P. (2020). Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment. *Journal of Advanced Transportation*, 2020, 6649867. <https://doi.org/10.1155/2020/6649867>
- Chiao, V. (2019). Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, 15(2), 126–139. <https://doi.org/10.1017/S1744552319000077>
- Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020a). The combat against COVID-19 in Portugal: How state measures and data availability reinforce some organizational values and contribute to the sustainability of the National Health System. *Sustainability*, 12(18), 7513. <https://doi.org/10.3390/su12187513>
- Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020b). The combat against COVID-19 in Portugal, Part II: how governance reinforces some organizational values and contributes to the sustainability of crisis management. *Sustainability*, 12(20), 8715. <https://doi.org/10.3390/su12208715>
- Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2022). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. *European Journal of Applied Business Management*, 8(1), 1–12.
- Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2021). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. In *European Consortium for Political Research General Conference* (pp. 1–18). United Kingdom.
- Correia, J. M. C. (1987). *Legalidade e autonomia contratual nos contratos administrativos* (pp. 283, 768). Lisboa: Almedina.
- DeCamp, M., & Tilburt, J. (2019). Why we cannot trust artificial intelligence in medicine. *The Lancet Digital health*, 1(8), e390. [https://doi.org/10.1016/S2589-7500\(19\)30197-9](https://doi.org/10.1016/S2589-7500(19)30197-9)
- Dhingra, M., & Gupta, N. (2017). Comparative analysis of fault tolerance models and their challenges in cloud computing. *International Journal of Engineering & Technology*, 6(2), 36–40. <https://doi.org/10.14419/ijet.v6i2.7565>
- Ettlinger, N. (2022). *Algorithms and the Assault on Critical Thought: Digitalized Dilemmas of Automated Governance and Communitarian Practice* (1st ed.). Routledge. <https://doi.org/10.4324/9781003109792>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: Picador, St Martin's Press.
- Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., & Burke, D. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 442(7101), 448–452. <https://doi.org/10.1038/nature04795>
- Fetzer, T., & Graeber, T. (2021). Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 118(33), e2100814118. <https://doi.org/10.1073/pnas.2100814118>
- Galets, P., Katsaliaki, K., & Kumar, S. (2022). The medical and societal impact of big data analytics and artificial intelligence applications in combating pandemics: A review focused on Covid-19. *Social Science & Medicine*, 301, 114973. <https://doi.org/10.1016/j.socscimed.2022.114973>
- Gianfrancesco, M., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine*, 178(11), 1544–1547. <https://doi.org/10.1001/jamainternmed.2018.3763>
- Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B., & Birney, E. (2013). Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, 494(7435), 77–80. <https://doi.org/10.1038/nature11875>
- Gomes, C. A., & Pedro, R. (Coords.). (2020). *Direito administrativo de necessidade e de exceção*. Lisboa: AAFDL.
- Gómez Abeja, L. (2022). Inteligencia artificial y derechos fundamentales. In F. H. Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (1.ª ed., pp. 91–114, 93). Murcia: Ediciones Laborum. (In Spanish).
- Gómez Colomer, J.-L. (2023). *El juez-robot: La independencia judicial en peligro*. Valencia: Tirant lo Blanch. (In Spanish).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Greiner, R., Grove, A., & Kogan, A. (1997). Knowing what doesn't matter: exploiting the omission of irrelevant data. *Artificial Intelligence*, 97(1–2), 345–380. [https://doi.org/10.1016/S0004-3702\(97\)00048-9](https://doi.org/10.1016/S0004-3702(97)00048-9)

- Gunasekeran, D., Tseng, R., Tham, Y., & Wong, T. (2021). Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. *NPJ Digital Medicine*, 4(1), 40. <https://doi.org/10.1038/s41746-021-00412-9>
- Gürsoy, E., & Kaya, Y. (2023). An overview of deep learning techniques for COVID-19 detection: methods, challenges, and future works. *Multimedia Systems*, 29(3), 1603–1627. <https://doi.org/10.1007/s00530-023-01083-0>
- Hanegraaff, W. (2013). *Western Esotericism: A Guide for the Perplexed*. Bloomsbury Publishing.
- Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2), 8–12. <https://doi.org/10.1109/MIS.2009.36>
- Hazarika, I. (2020). Artificial intelligence: opportunities and implications for the health workforce. *International Health*, 12(4), 241–245. <https://doi.org/10.1093/inthealth/ihaa007>
- Hoff, K., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434. <https://doi.org/10.1177/0018720814547570>
- Hulten, G. (2018). *Building Intelligent Systems: A Guide to Machine Learning Engineering*. Apress.
- Igual, L., & Seguí, S. (2024). *Supervised learning*. In *Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications* (pp. 67–97). Springer International Publishing.
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. <https://doi.org/10.1136/svn-2017-000101>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jones, K., Patel, N., Levy, M., Storeygard, A., Balk, D., Gittleman, J., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, 451(7181), 990–993. <https://doi.org/10.1038/nature06536>
- Kandlhofer, M., Weixelbraun, P., Menzinger, M., Steinbauer-Wagner, G., & Kemenesi, Á. (2023). Education and Awareness for Artificial Intelligence. In *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives* (pp. 3–12). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44900-0_1
- Kavanagh, J., & Rich, M. (2018). *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. RAND Corporation. <https://doi.org/10.7249/RR2314>
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521–3526. <https://doi.org/10.1073/pnas.1611835114>
- Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems*, 31(3), 388–409. <https://doi.org/10.1080/0960085X.2021.1927212>
- Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet Policy Review*, 9(2). <https://doi.org/10.14763/2020.2.1469>
- Lin, X., Liu, J., Hao, J., Wang, K., Zhang, Y., Li, H., ... & Tan, X. (2020). Collinear holographic data storage technologies. *Opto-Electronic Advances*, 3(3), 190004. <https://doi.org/10.29026/oea.2020.190004>
- Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data*. John Wiley & Sons.
- Macrae, C. (2022). Learning from the failure of autonomous and intelligent systems: Accidents, safety, and sociotechnical sources of risk. *Risk Analysis*, 42(9), 1999–2025. <https://doi.org/10.1111/risa.13850>
- Margetts, H. (2022). Rethinking AI for good governance. *Daedalus*, 151(2), 360–371. https://doi.org/10.1162/daed_a_01922
- Matsuzaka, Y., & Yashiro, R. (2022). Applications of Deep Learning for Drug Discovery Systems with BigData. *BioMedInformatics*, 2(4), 603–624. <https://doi.org/10.3390/biomedinformatics2040039>
- Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>
- Morse, S., Mazet, J., Woolhouse, M., Parrish, C., Carroll, D., Karesh, W., Zambrana-Torrel, C., Lipkin, W., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *Lancet*, 380(9857), 1956–1965. [https://doi.org/10.1016/S0140-6736\(12\)61684-5](https://doi.org/10.1016/S0140-6736(12)61684-5)
- Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258. <https://doi.org/10.1016/j.array.2022.100258>
- Navigli, R., Conia, S., & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *Journal of Data and Information Quality*, 15(2), 10. <https://doi.org/10.1145/3597307>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>

- O'Reilly-Shah, V., Gentry, K., van Cleve, W., Kendale, S., Jabaley, C., & Long, D. (2020). The COVID-19 pandemic highlights shortcomings in US health care informatics infrastructure: a call to action. *Anesthesia & Analgesia*, 131(2), 340–344. <https://doi.org/10.1213/ANE.0000000000004945>
- Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>
- Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *Systems and Humans*, 30(3), 286–297. <https://doi.org/10.1109/3468.844354>
- Pedro, R. (2022). Traços gerais da indemnização civil extracontratual pública em contextos de excecionalidade. In *Impactos da pandemia da Covid-19 nas estruturas do direito público* (pp. 379–413). Coimbra: Almedina. (In Portuguese).
- Pedro, R. (2023). Inteligência artificial e arbitragem de direito público: Primeiras reflexões. In R. Pedro, & P. Caliendo (Coords.), *Inteligência artificial no contexto do direito público: Portugal e Brasil* (1.ª ed., pp. 105–127). Coimbra: Almedina. (In Portuguese).
- Romano, A., Spadaro, G., Balliet, D., Joireman, J., van Lissa, C., Jin, S., ... & Leander, N. P. (2021). Cooperation and trust across societies during the COVID-19 pandemic. *Journal of Cross-Cultural Psychology*, 52(7), 622–642. <https://doi.org/10.1177/00220221209889>
- Ruan, W., Yi, X., & Huang, X. (2021). Adversarial robustness of deep learning: Theory, algorithms, and applications. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 4866–4869). <https://doi.org/10.48550/arXiv.2108.10451>
- Rubin, O., Errett, N., Upshur, R., & Baekkeskov, E. (2021). The challenges facing evidence-based decision making in the initial response to COVID-19. *Scandinavian Journal of Public Health*, 49(7), 790–796. <https://doi.org/10.1177/140349482199722>
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- Sass, J., Bartschke, A., Lehne, M., Essenwanger, A., Rinaldi, E., Rudolph, S., ... & Thun, S. (2020). The German Corona Consensus Dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. *BMC Medical Informatics and Decision Making*, 20, 341. <https://doi.org/10.1186/s12911-020-01374-w>
- Shin, D., & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 277–284. <https://doi.org/10.1016/j.chb.2019.04.019>
- Shaelou, S. L., & Razmetaeva, Y. (2023). Challenges to fundamental human rights in the age of artificial intelligence systems: Shaping the digital legal order while upholding rule of law principles and European values. *ERA Forum*, 24(3), 567–587. <https://doi.org/10.1007/s12027-023-00777-2>
- Silva, M., Flood, C., Goldenberg, A., & Singh, D. (2022). Regulating the Safety of Health-Related Artificial Intelligence. *Healthcare Policy*, 17(4), 63–77. <https://doi.org/10.12927/hcpol.2022.26824>
- Smidt, H., & Jokonya, O. (2021). The challenge of privacy and security when using technology to track people in times of COVID-19 pandemic. *Procedia Computer Science*, 181, 1018–1026. <https://doi.org/10.1016/j.procs.2021.01.281>
- Sundar, S. (2020). Rise of machine agency: A framework for studying the psychology of human – AI interaction (HAI). *Journal of Computer-Mediated Communication*, 25(1), 74–88. <https://doi.org/10.1093/jcmc/zmz026>
- Susskind, D. (2021). A world without work: Technology, automation and how we should respond. *New Technology, Work and Employment*, 36(1), 114–117. <https://doi.org/10.1111/ntwe.12186>
- Syed, R., Ulbricht, M., Piotrowski, K., & Krstic, M. (2023). A Survey on Fault-Tolerant Methodologies for Deep Neural Networks. *Pomiar Automatyka Robotyka*, 27(2), 89–98. https://doi.org/10.14313/PAR_248/89
- Syrowatka, A., Kuznetsova, M., Alsubai, A., Beckman, A., Bain, P., Craig, K., ... & Bates, D. (2021). Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *npj Digital Medicine*, 4(1), 96. <https://doi.org/10.1038/s41746-021-00459-8>
- Theis, T., & Wong, H. (2017). The end of Moore's law: A new beginning for information technology. *Computing in Science & Engineering*, 19(2), 41–50. <https://doi.org/10.1109/MCSE.2017.29>
- Thomsen, K. (2019). Ethics for artificial intelligence, ethics for all. *Paladyn, Journal of Behavioral Robotics*, 10(1), 359–363. <https://doi.org/10.1515/pjbr-2019-0029>
- Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- Villegas-Ch, W., Jaramillo-Alcázar, A., & Luján-Mora, S. (2024). Evaluating the Robustness of Deep Learning Models against Adversarial Attacks: An Analysis with FGSM, PGD and CW. *Big Data and Cognitive Computing*, 8(1), 8. <https://doi.org/10.3390/bdcc8010008>
- Vopson, M. (2020). The information catastrophe. *AIP Advances*, 10(8), 085014. <https://doi.org/10.1063/5.0019941>

- Wallach W., & Allen, C. (2008). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press.
- Wang, S., & Shi, W. (2011). Data Mining and Knowledge Discovery. In W. Kresse, D. Danko (Eds.), *Springer Handbook of Geographic Information*. Springer Handbooks. https://doi.org/10.1007/978-3-540-72680-7_5
- Wong, F., de la Fuente-Nunez, C., & Collins, J. (2023). Leveraging artificial intelligence in the fight against infectious diseases. *Science*, 381(6654), 164–170. <https://doi.org/10.1126/science.adh1114>
- Wu, D., Xu, H., Yongyi, W., & Zhu, H. (2022). Quality of government health data in COVID-19: definition and testing of an open government health data quality evaluation framework. *Library Hi Tech*, 40(2), 516–534. <https://doi.org/10.1108/LHT-04-2021-0126>
- Zhang, Q., Gao, J., Wu, J., Cao, Z., & Dajun, D. (2022). Data science approaches to confronting the COVID-19 pandemic: a narrative review. *Philosophical Transactions of the Royal Society A*, 380(2214), 20210127. <https://doi.org/10.1098/rsta.2021.0127>
- Zhou, J., Zheng, W., Wang, D., & Coit, D. W. (2024). A resilient network recovery framework against cascading failures with deep graph learning. *Journal of Risk and Reliability*, 238(1), 193–203. <https://doi.org/10.1177/1748006X22112886>

Authors information



Pedro Miguel Alves Ribeiro Correia – PhD in Social Sciences (Specialty in Public Administration), Invited Associate Professor, Faculty of Law, University of Coimbra; Visiting Full Professor, ICET/CUA/UFMT, Barra do Garças
Address: Pátio da Universidade, 3004-528 Coimbra, Portugal;
 Avenida ValdonVarjão, n. 6390, Barra do Garças – MT, CEP: 78605-091, Brazil
E-mail: pcorreia@fd.uc.pt
ORCID ID: <https://orcid.org/0000-0002-3111-9843>
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58223408400>
WoS Researcher ID: <https://www.webofscience.com/wos/author/record/B-2753-2015>
Google Scholar ID: <https://scholar.google.pt/citations?user=KABKPuUAAAAJ>



Ricardo Lopes Dinis Pedro – PhD (Law), Researcher, Lisbon Public Law Research Centre, Faculty of Law, University of Lisbon
Address: Alameda da Universidade, 1649-014 Lisbon, Portugal
E-mail: ricardopedro@fd.ulisboa.pt
ORCID ID: <https://orcid.org/0000-0001-6339-5140>
Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57879177700>
WoS Researcher ID: <https://www.webofscience.com/wos/author/record/AEN-4511-2022>
Google Scholar ID: <https://scholar.google.com/citations?hl=en&user=oJ1ImgUAAAAJ>



Susana Videira – PhD (Law), Associate Professor, Faculty of Law, University of Lisbon; Scientific and Pedagogical Coordinator, Law Degree and the Master's Degree in Judicial Law, European University
Address: Faculdade de Direito da Universidade de Lisboa, Alameda da Universidade, 1649-014 Lisbon, Portugal; Universidade Europeia, Estrada da Correia, n.º 53, 1500-210, Lisbon, Portugal
E-mail: susanavideira@fd.ulisboa.pt
ORCID ID: <https://orcid.org/0000-0002-9246-2557>

Authors' contributions

The authors have contributed equally into the concept and methodology elaboration, validation, formal analysis, research, selection of sources, text writing and editing, project guidance and management.

Conflict of interests

The authors declare no conflict of interests.

Financial disclosure

Regarding the participation of the Author Ricardo Pedro, it should be noted that, to the exact extent of his participation, the work is financed (or partially financed) by national funds through FCT–Foundation for Science and Technology, I.P., under the project UIDP/04310/2020. This work was also supported by Portuguese national funds through FCT–Foundation for Science and Technology, I.P., under project UIDB/04643/2020.

Thematic rubrics

OECD: 5.05 / Law

PASJC: 3308 / Law

WoS: OM / Law

Article history

Date of receipt – June 15, 2025

Date of approval – June 27, 2025

Date of acceptance – March 25, 2025

Date of online placement – March 30, 2025



Научная статья

УДК 34:004:17:004.8:342.7

EDN: <https://elibrary.ru/egkppn>

DOI: <https://doi.org/10.21202/jdtl.2025.7>

Искусственный интеллект в здравоохранении: баланс инноваций, этики и защиты прав человека

Педро Мигель Алвес Рибейро Коррейя ✉

Коимбрский университет, Коимбра, Португалия

Рикардо Лопес Динис Педро

Лиссабонский университет, Лиссабон, Португалия

Сусана Видейра

Лиссабонский университет, Лиссабон, Португалия

Ключевые слова

защита данных,
здравоохранение,
искусственный интеллект,
права человека,
право,
правовое регулирование,
предиктивная аналитика,
фундаментальные права,
этика,
этическое регулирование

Аннотация

Цель: определить ключевые этические, правовые и социальные вызовы, связанные с использованием искусственного интеллекта в здравоохранении, а также разработать рекомендации для создания адаптивных правовых механизмов, способных обеспечить баланс между инновациями, этическим регулированием и защитой фундаментальных прав человека.

Методы: в ходе исследования был реализован многоаспектный методологический подход, интегрирующий классические правовые методы анализа с современными инструментами сравнительного правоведения. Данное исследование охватывает как фундаментальные основы правового регулирования цифровых технологий в медицинской сфере, так и глубокий анализ этических, правовых и социальных импликаций внедрения искусственного интеллекта в систему здравоохранения. Такой комплексный подход позволил обеспечить всестороннее понимание проблематики и сформировать обоснованные выводы относительно перспектив развития данной области.

Результаты: выявлен ряд серьезных проблем, связанных с использованием искусственного интеллекта в здравоохранении. К ним относятся необъективность данных, непрозрачность сложных алгоритмов и риски нарушения неприкосновенности частной жизни. Эти проблемы могут подорвать доверие общества к технологиям искусственного интеллекта и усугубить неравенство в доступе к медицинским услугам. Авторы приходят к выводу, что интеграция искусственного интеллекта в систему здравоохранения должна осуществляться с учетом фундаментальных прав, таких как защита данных и запрет дискриминации, а также соответствовать этическим нормам.

✉ Контактное лицо

© Коррейя П. М. А. Р., Педро Р. Л. Д., Видейра С., 2025

Статья находится в открытом доступе и распространяется в соответствии с лицензией Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/deed.ru>), позволяющей неограниченно использовать, распространять и воспроизводить материал при условии, что оригинальная работа упомянута с соблюдением правил цитирования.

Научная новизна: состоит в предложении эффективных механизмов управления для снижения рисков и максимизации потенциала искусственного интеллекта в кризисных ситуациях. Особое внимание уделяется регулятивным мерам, таким как оценка влияния, предусмотренная Законом об искусственном интеллекте. Эти меры играют ключевую роль в выявлении и минимизации рисков, связанных с высокорисковыми системами искусственного интеллекта, обеспечивая соблюдение этических норм и защиту основных прав.

Практическая значимость: заключается в разработке адаптивных правовых механизмов, которые поддерживают демократические нормы и оперативно реагируют на возникающие вызовы в области общественного здравоохранения. Предложенные механизмы позволяют достичь баланса между использованием искусственного интеллекта для управления кризисными ситуациями и сохранением прав человека. Это способствует укреплению доверия к системам искусственного интеллекта и их устойчивому положительному влиянию на общественное здравоохранение.

Для цитирования

Коррейя, П. М. А. Р., Педро, Р. Л. Д., & Видейра, С. (2025). Искусственный интеллект в здравоохранении: баланс инноваций, этики и защиты прав человека. *Journal of Digital Technologies and Law*, 3(1), 143–180. <https://doi.org/10.21202/jdtl.2025.7>

Список литературы

- Arass, M., & Souissi, N. (2018). Data lifecycle: from big data to SmartData. In *2018 IEEE 5th International Congress on Information Science and Technology* (pp. 80–87). IEEE. <https://doi.org/10.1109/CIST.2018.8596547>
- Alvarez Garcia, V. (1996). *El concepto de necesidad en derecho público* (1st ed.). Madrid: Civitas. (In Spanish).
- Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Baclic, O., Tunis, M., Young, K., Doan, C., Swerdfeger, H., & Schonfeld, J. (2020). Challenges and opportunities for public health made possible by advances in natural language processing. *Canada Communicable Disease Report*, 46(6), 161–168. <https://doi.org/10.14745/ccdr.v46i06a02>
- Bajwa, J., Munir, U., Nori, A., & Williams, B. (2021). Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthcare Journal*, 8(2), e188–e194. <https://doi.org/10.7861/fhj.2021-0095>
- Beck, U. (1986). *Risikogesellschaft: Auf dem Weg in eine andere Moderne*. Frankfurt am Main: Suhrkamp Verlag.
- Balog-Way, D., & McComas, K. (2022). COVID-19: Reflections on trust, tradeoffs, and preparedness. In *COVID-19* (pp. 6–16). Routledge.
- Bazarkina, D. Y., & Pashentsev, E. N. (2020). Malicious use of artificial intelligence. *Russia in Global Affairs*, 18(4), 154–177. <https://doi.org/10.31278/1810-6374-2020-18-4-154-177>
- Benke, K., & Benke, G. (2018). Artificial Intelligence and Big Data in Public Health. *International Journal of Environmental Research and Public Health*, 15(12), 2796. <https://doi.org/10.3390/ijerph15122796>
- Berk, R. A. (1983). An introduction to sample selection bias in sociological data. *American Sociological Review*, 48(3), 386–398. <https://doi.org/10.2307/2095230>
- Bigam, G., Adamtey, S., Onsarigo, L., & Jha, N. (2019). Artificial Intelligence for Construction Safety: Mitigation of the Risk of Fall. In K. Arai, S. Kapoor, R. Bhatia (Eds.). *Intelligent Systems and Applications*. Springer. https://doi.org/10.1007/978-3-030-01057-7_76
- Binder, W. (2024). Technology as (dis-)enchantment. AlphaGo and the meaning-making of artificial intelligence. *Cultural Sociology*, 18(1), 24–47. <https://doi.org/10.1177/17499755221138720>

- Bisconti, P., Orsitto, D., Fedorczyk, F., Brau, F., Capasso, M., De Marinis, L., ... & Schettini, C. (2023). Maximizing team synergy in AI-related interdisciplinary groups: an interdisciplinary-by-design iterative methodology. *AI & Society*, 38(4), 1443–1452. <https://doi.org/10.1007/s00146-022-01518-8>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Box, G. (1979). Robustness in the strategy of scientific model building. In R. Launer & G. Wilkinson (Eds.), *Robustness in Statistics* (pp. 201–236). Academic Press. <https://doi.org/10.1016/B978-0-12-438150-6.50018-2>
- Breiman, L. (2001). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199–231. <https://doi.org/10.1214/ss/1009213726>
- Bulled, N. (2023). "Solidarity:" A failed call to action during the COVID-19 pandemic. *Public Health in Practice*, 5, 100379. <https://doi.org/10.1016/j.puhip.2023.100379>
- Chen, A. (2016). A review of emerging non-volatile memory (NVM) technologies and applications. *Solid-State Electronics*, 125, 25–38. <https://doi.org/10.1016/j.sse.2016.07.006>
- Chen, J., Zhang, R., Han, W., Jiang, W., Hu, J., Lu, X., Liu, X., & Zhao, P. (2020). Path Planning for Autonomous Vehicle Based on a Two-Layered Planning Model in Complex Environment. *Journal of Advanced Transportation*, 2020, 6649867. <https://doi.org/10.1155/2020/6649867>
- Chiao, V. (2019). Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice. *International Journal of Law in Context*, 15(2), 126–139. <https://doi.org/10.1017/S1744552319000077>
- Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020a). The combat against COVID-19 in Portugal: How state measures and data availability reinforce some organizational values and contribute to the sustainability of the National Health System. *Sustainability*, 12(18), 7513. <https://doi.org/10.3390/su12187513>
- Correia, P., Mendes, I., Pereira, S., & Subtil, I. (2020b). The combat against COVID-19 in Portugal, Part II: how governance reinforces some organizational values and contributes to the sustainability of crisis management. *Sustainability*, 12(20), 8715. <https://doi.org/10.3390/su12208715>
- Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2022). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. *European Journal of Applied Business Management*, 8(1), 1–12.
- Correia, P., Pereira, S., Mendes, I., & Subtil, I. (2021). COVID-19 Crisis management and the Portuguese regional governance: Citizens perceptions as evidence. In *European Consortium for Political Research General Conference* (pp. 1–18). United Kingdom.
- Correia, J. M. C. (1987). *Legalidade e autonomia contratual nos contratos administrativos* (pp. 283, 768). Lisboa: Almedina.
- DeCamp, M., & Tilburt, J. (2019). Why we cannot trust artificial intelligence in medicine. *The Lancet Digital health*, 1(8), e390. [https://doi.org/10.1016/S2589-7500\(19\)30197-9](https://doi.org/10.1016/S2589-7500(19)30197-9)
- Dhingra, M., & Gupta, N. (2017). Comparative analysis of fault tolerance models and their challenges in cloud computing. *International Journal of Engineering & Technology*, 6(2), 36–40. <https://doi.org/10.14419/ijet.v6i2.7565>
- Ettlinger, N. (2022). *Algorithms and the Assault on Critical Thought: Digitalized Dilemmas of Automated Governance and Communitarian Practice* (1st ed.). Routledge. <https://doi.org/10.4324/9781003109792>
- Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: Picador, St Martin's Press.
- Ferguson, N., Cummings, D., Fraser, C., Cajka, J., Cooley, P., & Burke, D. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 442(7101), 448–452. <https://doi.org/10.1038/nature04795>
- Fetzer, T., & Graeber, T. (2021). Measuring the scientific effectiveness of contact tracing: Evidence from a natural experiment. *Proceedings of the National Academy of Sciences of the United States of America*, 118(33), e2100814118. <https://doi.org/10.1073/pnas.2100814118>
- Galets, P., Katsaliaki, K., & Kumar, S. (2022). The medical and societal impact of big data analytics and artificial intelligence applications in combating pandemics: A review focused on Covid-19. *Social Science & Medicine*, 301, 114973. <https://doi.org/10.1016/j.socscimed.2022.114973>
- Gianfrancesco, M., Tamang, S., Yazdany, J., & Schmajuk, G. (2018). Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data. *JAMA Internal Medicine*, 178(11), 1544–1547. <https://doi.org/10.1001/jamainternmed.2018.3763>
- Goldman, N., Bertone, P., Chen, S., Dessimoz, C., LeProust, E. M., Sipos, B., & Birney, E. (2013). Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature*, 494(7435), 77–80. <https://doi.org/10.1038/nature11875>
- Gomes, C. A., & Pedro, R. (Coords.). (2020). *Direito administrativo de necessidade e de excepção*. Lisboa: AAFDL.

- Gómez Abeja, L. (2022). Inteligencia artificial y derechos fundamentales. In F. H. Llano Alonso (Dir.), J. Garrido Martín & R. Valdivia Jiménez (Coords.), *Inteligencia artificial y filosofía del derecho* (1.^a ed., pp. 91–114, 93). Murcia: Ediciones Laborum. (In Spanish).
- Gómez Colomer, J.-L. (2023). *El juez-robot: La independencia judicial en peligro*. Valencia: Tirant lo Blanch. (In Spanish).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Greiner, R., Grove, A., & Kogan, A. (1997). Knowing what doesn't matter: exploiting the omission of irrelevant data. *Artificial Intelligence*, 97(1–2), 345–380. [https://doi.org/10.1016/S0004-3702\(97\)00048-9](https://doi.org/10.1016/S0004-3702(97)00048-9)
- Gunasekaran, D., Tseng, R., Tham, Y., & Wong, T. (2021). Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. *NPJ Digital Medicine*, 4(1), 40. <https://doi.org/10.1038/s41746-021-00412-9>
- Gürsoy, E., & Kaya, Y. (2023). An overview of deep learning techniques for COVID-19 detection: methods, challenges, and future works. *Multimedia Systems*, 29(3), 1603–1627. <https://doi.org/10.1007/s00530-023-01083-0>
- Hanegraaff, W. (2013). *Western Esotericism: A Guide for the Perplexed*. Bloomsbury Publishing.
- Halevy, A., Norvig, P., & Pereira, F. (2009). The unreasonable effectiveness of data. *IEEE Intelligent Systems*, 24(2), 8–12. <https://doi.org/10.1109/MIS.2009.36>
- Hazarika, I. (2020). Artificial intelligence: opportunities and implications for the health workforce. *International Health*, 12(4), 241–245. <https://doi.org/10.1093/inthealth/ihaa007>
- Hoff, K., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434. <https://doi.org/10.1177/0018720814547570>
- Hulten, G. (2018). *Building Intelligent Systems: A Guide to Machine Learning Engineering*. Apress.
- Igual, L., & Seguí, S. (2024). *Supervised learning*. In *Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications* (pp. 67–97). Springer International Publishing.
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. <https://doi.org/10.1136/svn-2017-000101>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jones, K., Patel, N., Levy, M., Storeygard, A., Balk, D., Gittleman, J., & Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, 451(7181), 990–993. <https://doi.org/10.1038/nature06536>
- Kandlhofer, M., Weixelbraun, P., Menzinger, M., Steinbauer-Wagner, G., & Kemenesi, Á. (2023). Education and Awareness for Artificial Intelligence. In *International Conference on Informatics in Schools: Situation, Evolution, and Perspectives* (pp. 3–12). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44900-0_1
- Kavanagh, J., & Rich, M. (2018). *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. RAND Corporation. <https://doi.org/10.7249/RR2314>
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., ... & Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521–3526. <https://doi.org/10.1073/pnas.1611835114>
- Kordzadeh, N., & Ghasemaghahi, M. (2022). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems*, 31(3), 388–409. <https://doi.org/10.1080/0960085X.2021.1927212>
- Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet Policy Review*, 9(2). <https://doi.org/10.14763/2020.2.1469>
- Lin, X., Liu, J., Hao, J., Wang, K., Zhang, Y., Li, H., ... & Tan, X. (2020). Collinear holographic data storage technologies. *Opto-Electronic Advances*, 3(3), 190004. <https://doi.org/10.29026/oea.2020.190004>
- Little, R. J., & Rubin, D. B. (2019). *Statistical analysis with missing data*. John Wiley & Sons.
- Macrae, C. (2022). Learning from the failure of autonomous and intelligent systems: Accidents, safety, and sociotechnical sources of risk. *Risk Analysis*, 42(9), 1999–2025. <https://doi.org/10.1111/risa.13850>
- Margetts, H. (2022). Rethinking AI for good governance. *Daedalus*, 151(2), 360–371. https://doi.org/10.1162/daed_a_01922
- Matsuzaka, Y., & Yashiro, R. (2022). Applications of Deep Learning for Drug Discovery Systems with BigData. *BioMedInformatics*, 2(4), 603–624. <https://doi.org/10.3390/biomedinformatics2040039>
- Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716679679>

- Morse, S., Mazet, J., Woolhouse, M., Parrish, C., Carroll, D., Karesh, W., Zambrana-Torrel, C., Lipkin, W., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *Lancet*, 380(9857), 1956–1965. [https://doi.org/10.1016/S0140-6736\(12\)61684-5](https://doi.org/10.1016/S0140-6736(12)61684-5)
- Mumuni, A., & Mumuni, F. (2022). Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258. <https://doi.org/10.1016/j.array.2022.100258>
- Navigli, R., Conia, S., & Ross, B. (2023). Biases in Large Language Models: Origins, Inventory, and Discussion. *Journal of Data and Information Quality*, 15(2), 10. <https://doi.org/10.1145/3597307>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. <https://doi.org/10.1126/science.aax2342>
- O'Reilly-Shah, V., Gentry, K., van Cleve, W., Kendale, S., Jabaley, C., & Long, D. (2020). The COVID-19 pandemic highlights shortcomings in US health care informatics infrastructure: a call to action. *Anesthesia & Analgesia*, 131(2), 340–344. <https://doi.org/10.1213/ANE.0000000000004945>
- Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>
- Parasuraman, R., Sheridan, T., & Wickens, C. (2000). A model for types and levels of human interaction with automation. *Systems and Humans*, 30(3), 286–297. <https://doi.org/10.1109/3468.844354>
- Pedro, R. (2022). Traços gerais da indemnização civil extracontratual pública em contextos de excecionalidade. In *Impactos da pandemia da Covid-19 nas estruturas do direito público* (pp. 379–413). Coimbra: Almedina. (In Portuguese).
- Pedro, R. (2023). Inteligência artificial e arbitragem de direito público: Primeiras reflexões. In R. Pedro, & P. Caliendo (Coords.), *Inteligência artificial no contexto do direito público: Portugal e Brasil* (1.ª ed., pp. 105–127). Coimbra: Almedina. (In Portuguese).
- Romano, A., Spadaro, G., Balliet, D., Joireman, J., van Lissa, C., Jin, S., ... & Leander, N. P. (2021). Cooperation and trust across societies during the COVID-19 pandemic. *Journal of Cross-Cultural Psychology*, 52(7), 622–642. <https://doi.org/10.1177/00220221209889>
- Ruan, W., Yi, X., & Huang, X. (2021). Adversarial robustness of deep learning: Theory, algorithms, and applications. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management* (pp. 4866–4869). <https://doi.org/10.48550/arXiv.2108.10451>
- Rubin, O., Errett, N., Upshur, R., & Baekkeskov, E. (2021). The challenges facing evidence-based decision making in the initial response to COVID-19. *Scandinavian Journal of Public Health*, 49(7), 790–796. <https://doi.org/10.1177/140349482199722>
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- Sass, J., Bartschke, A., Lehne, M., Essenwanger, A., Rinaldi, E., Rudolph, S., ... & Thun, S. (2020). The German Corona Consensus Dataset (GECCO): a standardized dataset for COVID-19 research in university medicine and beyond. *BMC Medical Informatics and Decision Making*, 20, 341. <https://doi.org/10.1186/s12911-020-01374-w>
- Shin, D., & Park, Y. (2019). Role of fairness, accountability, and transparency in algorithmic affordance. *Computers in Human Behavior*, 98, 277–284. <https://doi.org/10.1016/j.chb.2019.04.019>
- Shaelou, S. L., & Razmetaeva, Y. (2023). Challenges to fundamental human rights in the age of artificial intelligence systems: Shaping the digital legal order while upholding rule of law principles and European values. *ERA Forum*, 24(3), 567–587. <https://doi.org/10.1007/s12027-023-00777-2>
- Silva, M., Flood, C., Goldenberg, A., & Singh, D. (2022). Regulating the Safety of Health-Related Artificial Intelligence. *Healthcare Policy*, 17(4), 63–77. <https://doi.org/10.12927/hcpol.2022.26824>
- Smidt, H., & Jokonya, O. (2021). The challenge of privacy and security when using technology to track people in times of COVID-19 pandemic. *Procedia Computer Science*, 181, 1018–1026. <https://doi.org/10.1016/j.procs.2021.01.281>
- Sundar, S. (2020). Rise of machine agency: A framework for studying the psychology of human – AI interaction (HAI). *Journal of Computer-Mediated Communication*, 25(1), 74–88. <https://doi.org/10.1093/jcmc/zmz026>
- Susskind, D. (2021). A world without work: Technology, automation and how we should respond. *New Technology, Work and Employment*, 36(1), 114–117. <https://doi.org/10.1111/ntwe.12186>
- Syed, R., Ulbricht, M., Piotrowski, K., & Krstic, M. (2023). A Survey on Fault-Tolerant Methodologies for Deep Neural Networks. *Pomiary Automatyka Robotyka*, 27(2), 89–98. https://doi.org/10.14313/PAR_248/89
- Syrowatka, A., Kuznetsova, M., Alsubai, A., Beckman, A., Bain, P., Craig, K., ... & Bates, D. (2021). Leveraging artificial intelligence for pandemic preparedness and response: a scoping review to identify key use cases. *npj Digital Medicine*, 4(1), 96. <https://doi.org/10.1038/s41746-021-00459-8>
- Theis, T., & Wong, H. (2017). The end of Moore's law: A new beginning for information technology. *Computing in Science & Engineering*, 19(2), 41–50. <https://doi.org/10.1109/MCSE.2017.29>

- Thomsen, K. (2019). Ethics for artificial intelligence, ethics for all. *Paladyn, Journal of Behavioral Robotics*, 10(1), 359–363. <https://doi.org/10.1515/pjbr-2019-0029>
- Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
- Villegas-Ch, W., Jaramillo-Alcázar, A., & Luján-Mora, S. (2024). Evaluating the Robustness of Deep Learning Models against Adversarial Attacks: An Analysis with FGSM, PGD and CW. *Big Data and Cognitive Computing*, 8(1), 8. <https://doi.org/10.3390/bdcc8010008>
- Vopson, M. (2020). The information catastrophe. *AIP Advances*, 10(8), 085014. <https://doi.org/10.1063/5.0019941>
- Wallach W., & Allen, C. (2008). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press.
- Wang, S., & Shi, W. (2011). Data Mining and Knowledge Discovery. In W. Kresse, D. Danko (Eds.), *Springer Handbook of Geographic Information*. Springer Handbooks. https://doi.org/10.1007/978-3-540-72680-7_5
- Wong, F., de la Fuente-Nunez, C., & Collins, J. (2023). Leveraging artificial intelligence in the fight against infectious diseases. *Science*, 381(6654), 164–170. <https://doi.org/10.1126/science.adh1114>
- Wu, D., Xu, H., Yongyi, W., & Zhu, H. (2022). Quality of government health data in COVID-19: definition and testing of an open government health data quality evaluation framework. *Library Hi Tech*, 40(2), 516–534. <https://doi.org/10.1108/LHT-04-2021-0126>
- Zhang, Q., Gao, J., Wu, J., Cao, Z., & Dajun, D. (2022). Data science approaches to confronting the COVID-19 pandemic: a narrative review. *Philosophical Transactions of the Royal Society A*, 380(2214), 20210127. <https://doi.org/10.1098/rsta.2021.0127>
- Zhou, J., Zheng, W., Wang, D., & Coit, D. W. (2024). A resilient network recovery framework against cascading failures with deep graph learning. *Journal of Risk and Reliability*, 238(1), 193–203. <https://doi.org/10.1177/1748006X22112886>

Сведения об авторах



Коррейя Педро Мигель Алвес Рибейро – PhD в области общественных наук (государственное управление), приглашенный доцент, юридический факультет, Коимбровский университет; приглашенный профессор, ICET/CUA/UFMT, Барра до Гарсас

Адрес: Португалия, 3004-528, г. Коимбра, Патио да Универсидаде; Бразилия, 78605-091, Авенида Валдон Варжан, 6390, Барра до Гарсас – MT, CEP

E-mail: pcorreia@fd.uc.pt

ORCID ID: <https://orcid.org/0000-0002-3111-9843>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=58223408400>

WoS Researcher ID: <https://www.webofscience.com/wos/author/record/B-2753-2015>

Google Scholar ID: <https://scholar.google.pt/citations?user=KABKPuUAAAAJ>



Рикардо Лопес Динис Педро – PhD в области права, научный сотрудник, Лиссабонский исследовательский центр в области публичного права, юридический факультет, Лиссабонский университет

Адрес: Португалия, 1649-014, г. Лиссабон, Аламеда де Универсидаде

E-mail: ricardopedro@fd.ulisboa.pt

ORCID ID: <https://orcid.org/0000-0001-6339-5140>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=57879177700>

WoS Researcher ID: <https://www.webofscience.com/wos/author/record/AEN-4511-2022>

Google Scholar ID: <https://scholar.google.com/citations?hl=en&user=oJ1ImgUAAAAJ>



Сусана Видейра – PhD в области права, доцент, юридический факультет, Лиссабонский университет; координатор по науке и образованию, Европейский университет в Лиссабоне

Адрес: Португалия, 1649-014, г. Лиссабон, Аламеда де Универсидаде; Португалия, 1500-210, г. Лиссабон, Эстрада да Коррейя, 53

E-mail: susanavideira@fd.ulisboa.pt

ORCID ID: <https://orcid.org/0000-0002-9246-2557>

Вклад авторов

Авторы внесли равный вклад в разработку концепции, методологии, валидацию, формальный анализ, проведение исследования, подбор источников, написание и редактирование текста, руководство и управление проектом.

Конфликт интересов

Авторы сообщают об отсутствии конфликта интересов.

Финансирование

Участие автора Рикардо Лопес Динис Педро частично финансировалось Фондом науки и технологий Португалии (Foundation for Science and Technology, FCT) в рамках проекта UIDP/04310/2020. Исследование также было поддержано тем же Фондом в рамках проекта UIDB/04643/2020.

Тематические рубрики

Рубрика OECD: 5.05 / Law

Рубрика ASJC: 3308 / Law

Рубрика WoS: OM / Law

Рубрика ГРНТИ: 10.15.59 / Права и свободы человека и гражданина

Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 15 июня 2024 г.

Дата одобрения после рецензирования – 27 июня 2024 г.

Дата принятия к опубликованию – 25 марта 2025 г.

Дата онлайн-размещения – 30 марта 2025 г.