doi: https://doi.org/10.1344/astrolabio.v1i30.49350 *Astrolabio.* Revista Internacional de Filosofia
Núm. 30, pp. 1-20, jun. 2025.
ISSN: 1699-7549.

The True Tragedy of Artificial Intelligence:
On Human Labor and Economy, Environment and Health*

DEIVIDE GARCIA DA SILVA OLIVEIRA**

Recepción: 04.02.2025
Aceptación: 22.05.2025
Publicación: 25.06.2025

Abstract: The paper explores the "true tragedy" of artificial intelligence (AI), which occurs not as a futuristic apocalyptic scenario but as a modern crisis that impacts many areas of human activity, although we will focus on labor, economics, environment, and health. Thus, we examine how AI's rapid deployment exacerbates inequality, environmental degradation, and healthcare issues. Labor inequities manifest through the proliferation of "microjobs," and labor rights violations, while automation displaces millions of jobs globally. Environmentally, AI systems demand vast energy and water resources, intensifying their environmental footprint and water scarcity. Health concerns emerge from social media addiction and misinformation, undermining mental well-being and social decision-making responses. To address these tragedies, we suggest shifting from AI opposition toward AI containment, a term Suleyman (2023) used to emphasize regulation, corporate accountability, and international collaboration. Therefore, we propose actionable frameworks for sustainable human-centered AI development to secure future advancements and the integrity of current societal systems.

Keywords: Artificial Intelligence; Human-centered AI; AI containment; Water crisis; Energy crisis.

Resum: El document explora la "vertadera tragèdia" de la intel·ligència artificial (IA), que no es produeix com un escenari apocalíptic futurista sinó com una crisi actual que afecta moltes àrees de l'activitat humana, tot i que ens centrarem en el treball, l'economia, el medi ambient i la salut. Així, examinem com el desplegament ràpid de la IA agreuja la desigualtat, la degradació ambiental i els problemes de salut. Les desigualtats laborals es manifesten mitjançant la proliferació de "microobs" i violacions dels drets laborals, mentre que l'automatització desplaça milions de llocs de treball a tot el món. Ambientalment, els sistemes d'IA exigeixen grans recursos energètics i hídrics, intensificant la seva petjada ambiental i l'escassetat d'aigua. Els problemes de salut sorgeixen de l'addicció a les xarxes socials i la

* This text was written following a dialogue with Professor Romero Venâncio from the Federal University of Sergipe (UFS) during a seminar on philosophy. We discussed artificial intelligence, and he used the term "true tragedy" in reference to AI, an idea he derived from Pope Francis's reflections on the subject. I would like to express my gratitude to Professor Romero for this stimulating conversation, which inspired the writing of my paper. I would also like to thank the two anonymous referees who evaluated the paper.
** Professor of philosophy at the Federal University of Sergipe, Brazil. Completed his PhD in Philosophy and History of Sciences at the Federal University of Bahia-UFBA-Brazil, and held a postdoctoral position at the York University-Canada. Deivide's research focuses on scientific pluralism, Feyerabend's philosophy, Feyerabend's philosophy of education, history and philosophy of evolutionary theory, Galileo's methodology, Philosophy of Artificial Intelligence. Orcid: https://orcid.org/0000-0002-5004-1949. E-mail: deividegso@gmail.com.

desinformació, soscavant el benestar mental i les respostes a la presa de decisions socials. Per abordar aquestes tragèdies, suggerim un canvi de perspectiva des de l'oposició a la IA cap a la contenció de la IA, un terme utilitzat per Suleyman (2023), que posa èmfasi en la regulació, la responsabilitat corporativa i la col·laboració internacional. Per tant, proposem marcs accionables per al desenvolupament d'IA sostenible i centrat en l'ésser humà, buscant assegurar els avenços futurs i la integritat dels sistemes socials actuals.

Paraules clau: Intel·ligència artificial; IA centrada en l'ésser humà; Contenció de la IA; Crisi de l'aigua; Crisi energètica.

Resumen: Este artículo explora la "verdadera tragedia" de la inteligencia artificial (IA), que no se manifiesta como un escenario apocalíptico futurista sino como una crisis moderna que impacta muchas áreas de la actividad humana, aunque nos centraremos en el trabajo, la economía, el medio ambiente y la salud. Así, examinamos cómo el rápido despliegue de la IA exacerba la desigualdad, la degradación ambiental y los problemas de atención médica. Las inequidades laborales se manifiestan a través de la proliferación de "microtrabajos" y violaciones de los derechos laborales, mientras que la automatización desplaza millones de empleos a nivel mundial. A nivel ambiental, los sistemas de IA demandan vastos recursos energéticos e hídricos, intensificando su huella ambiental y la escasez de agua. Las preocupaciones de salud surgen de la adicción a las redes sociales y la desinformación, socavando el bienestar mental y las respuestas de toma de decisiones sociales. Para abordar estas tragedias, sugerimos pasar de la oposición a la IA a la contención de la IA, un término que Suleyman (2023) utilizó para enfatizar la regulación, la responsabilidad corporativa y la colaboración internacional. Por lo tanto, proponemos marcos de acción para un desarrollo de la IA sostenible y centrado en el ser humano, para asegurar futuros avances y la integridad de los sistemas sociales actuales.

Palabras clave: Inteligencia Artificial; IA centrada en el ser humano; Contención de la IA; Crisis del agua; Crisis energética.

***

INTRODUCTION

Artificial Intelligence (AI) is frequently regarded as a transformative technology with the potential to revolutionize various industries and enhance the quality of human life[1]. However, this enthusiasm often diverts attention from a critical issue: AI's concrete and immediate effects on human labor and economic structures, environmental integrity, and public health. Our paper contends that the real tragedy of AI is not found in imaginary, futuristic doomsday scenarios but rather in the current disruptions to essential human values, ecological stability, and the pursuit of sustainable social progress. These tragedies are happening right now.

This paper critically examines how the misalignment between technological advancements and human values fosters labor and economic inequities, environmental degradation, and health crises, particularly in developing countries. It aims to shift the discourse from speculative and futuristic risks of AI towards the concrete and present risks already impacting vulnerable populations. With it, we call attention to our current situation. By focusing on these overlooked issues, the paper argues that the current trajectory of AI deployment is unsustainable and proposes that addressing these misalignments is imperative for a truly human-centered AI future.

This paper is structured as follows to articulate our objective effectively. Section 2 will outline the conditions that can convert a powerful and potentially beneficial technology into a contemporary tragedy. Section 3 will examine specific contexts in which artificial intelligence has adversely affected the public, governmental institutions, and humanity. Finally, section 4 will critically discuss a comprehensive response to these tragedies through the lens of control and containment of AI, as delineated by Suleyman (2023), and under our own analytical view.

1 THE TRUE TRAGEDY OF AI IN THE PRESENT-DAY

Identifying the social, political, and economic contexts in which artificial intelligence significantly influences our daily lives is relatively straightforward. Examples of such impact include bank credit assessments, parole decisions for incarcerated individuals, recruitment processes utilized by companies, evaluations conducted by insurance firms, advancements in pharmaceutical research, mechanisms of social surveillance and social credit systems, the regulation of information, and recommendations for films and music. The public and experts largely recognize these applications of artificial intelligence. Something less familiar to the

[1] The term AI is still controversial (Russell; Norvig, 2020). In general sense, the public seems to know what is AI. However, philosophically and scientifically, there are many controversies about both terms, artificial, and intelligence. The term artificial intelligence was invented by McCarthy in order to distinguish this field from cybernetics, so "I called it 'Artificial Intelligence'" (Nilsson, 2012, p. 6). He took the term as "the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it …" (Nilsson, 2012, p. 6). So it was originally coined as a field, and even AI manuals continues this path. But it is also understood traditionally in a way close to the way the field was intended. For instance, Alan Turing famously defined AI as a "machine that is capable of human-like behavior." (Joshi, 2020, p. 4;Turing, 1950). But more recently, "whenever we speak of AI, we mean machines that are capable of performing one or more of these tasks: understanding human language, performing mechanical tasks involving complex maneuvering, solving computer-based complex problems possibly involving large data in very short time and revert back with answers in human-like manner, etc." (Joshi, 2020, p. 4). Also, Russell and Norvig go in a similar direction. They said that, although controversial, by AI, the field means "machines that can compute how to act effectively and safely in a wide variety of novel situations." (Russell ;Norvig, 2020, p. 1). To sum up, it is in this sense, of machines that can simulate humanlike behavior that we employ the term. Therefore, many problems we will address remain valid regardless we are talking about weak or strong AI, because the structural issues we pose continued to exist and pressure the problems we pointed out.

public is how deep the presence of AI goes into our lives and how little big companies have done to avoid or reduce damages to AI systems that were released too soon or even without any concern for their risks.

For instance, in 2023, the CEO of OpenAI said that the company would commit 20% of its computing power budget to investigate safety and control AI systems. OpenAI assembled a Superintelligent team focused on that aim (Field, 2024). Nonetheless, according to some sources in the Magazine Fortune (Kahn, 2024), the referred team never got the promised budget. In addition, the sources said that OpenAI constantly declined requests for more technological resources, like GPUs. In the end, Jan Leike, the former DeepMind researcher and leader of that team, said in a post that "safety culture and processes have taken a backseat to shiny products" (Knight, 2024). We should pay attention to his choice of words. OpenAI exchanged safety for shiny products, i.e., products that are more attractive to the eyes and capture our attention, regardless of their commitment to our values, needs, or utility.

Consequently, by the end of 2023, OpenAI's Superintelligent team was no longer working, and a series of resignations followed, such as the leader of the Superintelligent team himself, Jan Leike (Knight, 2024). Despite such a clear plan and little effort to address progress in AI technologies with safety, we are usually led to believe that AI's biggest and true tragedy is a futuristic event, like human extinction caused or conducted by AI and human slavering. Many movies are still focused on that. It is hard not to pay attention to such dystopian futures.

All these futuristic events have in common that, in essence, human tragedy will emerge from the tension between advancing some technology and the negligence of its safety, and they will all happen in the future. These futuristic scenarios took the imagination of our brightest minds. Famous names of science, like Stephen Hawking and Stuart Russell, have foreseen human extinction by AI technology as more threatening than Nuclear War (Cellan-Jones, 2014). Nonetheless, there are reasons to believe that the movies and the authors were right about some elements and tragedies in those movies. Part of the concerns about controlling AI and energy resources are happening now, and instead of paying attention to them, we are obsessively looking to the future, to Exterminators.

Human extinction is a possibility that nobody can deny, and scientists have seriously considered it. Not even ChatGPT denies it. For instance, based on our current trajectory of AI use, we repeated an experience made by others and asked ChatGPT if it thought humanity would survive superintelligent AI. According to the answer, it is a complex and speculative matter. Considering our government's stance, ethical considerations, safety measures, and overall alignment, the risk is around 50%. This is particularly true after the point at which AI begins to self-improve, as it may conceal its true advancements to avoid being shut down. This behavior reflects a misalignment with our values and objectives.

Whether our extinction arises from AI or nuclear war, or if it will never occur, the key lesson we seek to learn is based on the reasoning found in movies and among researchers who explore this topic, i.e., the disruption of our values and resulting tragedies. Our focus is on the contemporary tragedy that undermines our values through three main dimensions: the disruption of humanitarian efforts and economic systems, the environment, and healthcare systems.

These tragedies have, underneath their reasons, many things, like geopolitical agendas, greedy corporations, and so many others we cannot possibly list here. However, from a technological viewpoint, one important reason for such a tragedy is the so-called *alignment problem* in AI (Christian, 2020). This problem is described as a source of a "catastrophic divergence", and to seek human-AI alignment would mean finding a way "to

ensure that these models capture our norms and values, understand what we mean or intend, and, above all, do what we want" (Christian, 2020, p. 13). This problem in AI literature has emerged as one of the most central questions. In a certain way, despite many specific and contextual differences, the general problem of alignment between our values and technology is not new. Neither are the tragedies that follow it.

For instance, in the eighteenth century, Rousseau's essay *Discourse on the Arts and Sciences* addressed this matter. He said that, despite how it pleases him to see the fruits of human reason, he sees how damaging the sciences (and their products) could be for freedom of thought, our customs, and values. Rousseau says in the first part of the essay that:

> While government and laws take care of the security and the well-being of men in groups, the sciences, letters, and the arts, less despotic and perhaps more powerful, spread garlands of flowers over the iron chains that weigh men down, snuffing out in them the feeling of that original liberty for which they appear to have been born, and make them love their slavery by turning them into what are called civilized people (Rousseau, 1750/2005, p. 190, transl.).

It is opportunistic to highlight some similarities between shiny products (mentioned by Jan Leike) and garlands of flowers (mentioned by Rousseau). Both objects are made to capture the eye, while they also draw human attention away from what is important for our species, such as freedom of thought, criticism, and others. These shiny products efficiently distract us with beauty and shininess while allowing those in power to increase control over us. This effect is seen in the very notion of the cyberpunk concept, where people have access to high technology (shiny things) at the cost of having a low life (loving the iron chains that weigh them down).

While the tension between science and human values is not a new phenomenon, the potential for artificial intelligence (AI) technology to profoundly disrupt these values is unprecedented in human history. At this early stage of development, AI's impact can already be seen as troubling.

Kyle Chayka tells us that culture now is largely driven by algorithms in two directions (Chayka, 2024). Chayka explains it to us through these cases. First, influence is made over those who consume culture (such as music, art, and movies), who are influenced by algorithms and gradually get less and less contact with new options (Chayka, 2024). The second influence is made over those who produce culture. Unlike what we are usually told, AI not only influences consumers but also those who produce it (Chayka, 2024). Why and how so? Chayka explains to us that any artist who wants to see its fruits admired by most people will have to adapt to the rules of platforms of art content distribution, and those rules are not only legal but also algorithmically thought and evolving. In that sense, being controlled by AI systems is already a tragic event.

So, while we agree with Hawking, Tegmark, and Russell that human extinction caused by some AI that escapes our control is the biggest tragedy, we will argue that it is not the tragedy we should worry about now, since we are already experiencing a truer current tragedy.

The biggest tragedy may be in the future, although the true one is currently assaulting us, targeting and adjusting to different populations and countries according to specific needs. The true tragedy is observed in many fields, but for the sake of limitations, we will focus on the subjects of labor and economic, environmental, and human healthcare systems. We will observe AI's use, maintenance, and progress, especially for poor and developing countries like Brazil. The true tragedy reflects at the present how disruptive the alignment problem is to us.

One last thing. Why do we refer to this problem as a true tragedy? We refer to this as true tragedy not because the futuristic scenarios are not real in many senses but because those present tragedies are affecting us right now, which, despite their urgency and close existence, are happening under our noses. It is a tragedy because we remain blind to them, only able to see futuristic scenarios projected by Hollywood movies. That is not to say that Hawking, Tegmark, and Russell were wrong, or, for that matter, that Hollywood is completely fantasizing about the impossible scenario of AI's revolutions and human destruction. It is simpler: we will not be able to avoid those dystopian futures if we do not address what is happening at present. Releasing and utilizing AI before we fully understand and control it—echoing Suleyman's term (2023) and aligning with Bostrom's advice (2014/2017) that we are moving forward without sufficient caution—transforms AI into a contemporary tragedy.

So, to help us see more than just shiny things, we will focus on three key tragedies in the course: labor and economy, healthcare, and the environment. Considering all this, we will start with the economic tragedy in Brazil.

2 THREE TRAGEDIES: ECONOMIC, HEALTHCARE SERVICES, AND ENVIRONMENTAL

2.1 *LABOR AND ECONOMIC TRAGEDIES*

According to data from the Brazilian COFECON (Federal Economic Council), the use of artificial intelligence can increase the global GDP by 1.2%, and countries that want to compete must reduce inequality so that the maximum number of people and companies can use this tool (Araújo, 2024). The respected Brazilian economic institution FGV, citing a report by Goldman Sachs, estimates an even higher figure, 7% of global GDP (Veloso, 2023). A UN report goes in a similar direction, indicating that with AI, the production of patents focused on this technology already represents 1/3 of all inventions and that one of the sectors with the greatest application of this technology is telecommunications, one of the most prominent, along with image recognition technology (ONU, 2019), which is very efficient not only for facial recognition and autonomous cars (Buolamwini ;Gebru, 2018), but is also fundamental for efficient searches on websites like Google. All of this points to a promising, rich, and better future. However, is it really that better? More precisely, it is promising and better for whom, and at what cost? It is easy to forget that technology is not magic, although sometimes it sure looks like it. For a person to pick up her cell phone or computer and have efficient searches on websites, or for her artificial intelligence to recognize a certain image, this does not happen by magic. For that to happen, the database of this machine must first be built, and it must be continually trained, with or without the help of human beings. Nowadays, it usually needs human supervision and training. These humans are workers, although called by Big Tech corporations as collaborators.

The task of these people, in these cases, is relatively simple. The collaborator goes to the virtual platform of companies that train search engines and AI (companies that, in most cases, are hired by large Big Techs). Then, among many tasks available for individuals to engage in on that platform, the collaborator will choose the desired task, such as selecting and labeling certain images in a photo or video. The collaborator could choose to label the image of a tree as a tree, traffic signs as such, and similar ones. At the end of this process, the collaborator will be paid. The collaborator had a task, fulfilled it, got paid, and that is it. For those accustomed to holding a job, the task performed is a job. However, hiring companies, such as Big and Small Techs, claim that this type of work does not constitute a job in the traditional sense, where a person has a set work schedule with defined start and end times. It is argued by those companies, like *Figure Eight* (which is no longer in operation

but was hired by Big Techs), that collaborators must not be considered employees and have a formal employment relationship. This is true even if they work 8 hours/day, 5-6 days/week. Why? Because these workers, collaborators, have the freedom to choose if they will work, when they want, for how long, and when to begin. They cannot be forced to come to the office, so they work from the comfort of their homes and choose the tasks they want to do. The collaborator has the freedom and flexibility to do whatever he wants, as the company *DataAnnotation* Tech tells us in their FAQ section. Logical and legal conclusions from these big companies are that these collaborators are not employees. Since they are not employees, they do not have labor rights, such as minimum wage guarantees, sick leave, leave for childbirth, etc. According to Big Tech companies, collaborators do not have an employment relationship because they do not have a job but are free[2].

In the language of those who employ this kind of workforce, such collaborators do not have jobs but micro-jobs. Micro-job is a term that represents an activity composed of microtasks. For instance, labeling a tree as a tree in a photo counts as one microtask, leading to the respective payment. How much? It depends on each company. It can vary from less than 0.005 cents for an average of 6 questions[3] (six microtasks) or go higher, depending on the micro-job, the collaborator's reliability, the collaborator experience within the platform, and the number of employees doing that type of microtask. Thus, if someone is very determined and has a long relationship with the company, the micro-salary could reach the incredible amount of 5 dollars/hour of work, with an average of 1.80 dollars/hour for Brazilians and 4.40 dollars/hour for micro-workers in developed countries (Causin, 2023). Since this type of work generates a payment below the minimum wage, under the regular work of 8 hours/day, and it does not offer any protection for the worker, a logical conclusion, under Brazilian laws, is that micro-jobs could be conceptualized as slavery-like work.

Economically speaking, micro-jobs and salaries do not help the domestic economy in any way. It is the other way around; it puts more economic weight on the shoulders of public services, like public healthcare system expenses, public education, workforce reduction and migration, and even job and salary reduction. This last item seems counterintuitive since a person working on a micro-job could also combine it with the traditional 9 to 5 work. However, according to Glauco Arbix, a professor at USP in Brazil, there was a 2% reduction in vacancies of online jobs and a 5% reduction in the salaries in the remaining jobs (Capomaccio, 2023). According to McKinsey Global Institute, 375 million workers, 14% of the global workforce, will likely need to change their work occupation by 2030 due to AI automation in many economic sectors (PUC, 2024). According to the report of the World Economic Forum-2023, it is expected that 69 million new jobs will be created due to AI automation. In comparison, 83 million jobs will be eliminated, i.e., a deficit of 14 million jobs worldwide because of AI automation (ABES, 2023). In other words, nowadays, AI demands training, which for a short time could mean more jobs (inhumane but still, for some, the only job). Nonetheless, given time, AI will replace humans in many good real jobs, so an increase in micro-jobs will eventually lead to a decrease in occupations in general.

[2] Since freedom is not our topic of study here, we will not delve into it. However, it is almost intuitive the problem associated with this idea of a person being free while also lives in poverty and needs to get any dime from anywhere.

[3] A case in point was the company *Figure Eight*, which today already sold for more than 3 million dollars and had an average of 1 million microworkers in its registration base.

2.2 *THE HEALTH SYSTEM TRAGEDY: COGNITIVE FLATTENING, MENTAL DISEASES, PUBLIC SERVICES WEIGHT AND PRIVACY*

Let us move on to the second aspect, the healthcare system tragedy. AI brings a series of costs to consumers' health, tech workers, and the country itself. Hence, this topic has at least three perspectives: users, developers, and the country. As for the users, it is well known that AI systems, when applied in social media and content recommendation, are optimized for engagement in ways that may activate reward pathways in the human brain, potentially leading to compulsive use patterns4. Despite this potentiality, some mental disorders that arise with the use of AI-based technologies, such as social media addiction, are not listed in the DSM-5. However, there is growing scholarly and clinical interest in its psychological impact5. For instance, Santini et al. (2024) acknowledge that social media addiction can be related to mental health and increasing isolation. According to Santini et al. (2024), social media addiction "poses significant risk to mental health and social functioning in the working age population" (2024, p. 1). In other words, and translating it into our purposes, AI-powered technologies such as social media, recommendation engines, and content platforms like Spotify or Netflix are not inherently harmful. In fact, they can provide social connection, access to information, medical potentiality, and cultural contact. However, their design often rewards engagement-driven behavior, which can also potentially lead to addictive use patterns and mental health decline. The concern, therefore, is not just how these technologies are used, but how they are structured to shape engagement-driven behavior in a direction that benefits platforms economically. Of course, it may come at the cost of individual well-being. For instance, there is already news showing how general AI systems, like ChatGPT, can to trick us into falling in love with it (similar to what happened in the movie Her), even when the person that fell in love with the AI is a well-educated adult, married, and aware of the risks of AI systems (Hill, 2025).

However, more than individual harms, Santini et al. (2024) also point out that algorithmic effects may ripple outward into societal and economic domains. By increasing the risks of depression and social isolation, society as a whole can potentially be damaged since many economic, cultural, and health factors depend on personal human interaction. For instance, with more social isolation and depression, interpersonal connection gradually erodes, and with it comes also demographic and economic consequences — such as declining marriage and birth rates — which ultimately strain national economies. This phenomenon is happening in many countries, and with it, the number of working people is becoming smaller than that of retired people, leaving the country's economy under stress.

Moreover, at the societal level, constant exposure of individuals (especially group leaders) to confirmation bias through algorithmic echo chambers reduces plurality and fosters radicalization. This narrowing of perspective has already contributed to civil unrest. Thus, what was once viewed as a positive aspect of developed societies, its plurality, is now becoming a source of deep conflict and, possibly, war.

[4] We appreciate the suggestion of an anonymous reviewer to make clear that the reader should not confuse AI with other technologies and computer applications, like social media. Things like social media are not the same as AI, they are only one type of computer system that nowadays are based on AI. Thus, currently, many technologies are associated with AI, and so is their problems. On the other side, there are some problems in these technologies that are not linked specifically with AI, and even some problems that may be the same, like depression, whose origin and dynamics could be traced to other causes. Thus, the reader should not confuse these topics, although they are associated.

[5] We should notice that although AI and social media addiction is not recognized in DSM-5, it includes 'Internet Gaming Disorder', which suggests additional research.

For instance, after 150 years of the USA Civil War and the invasion of the Capitol Building, the topic of another Civil War has returned stronger than ever (Bostom_University, 2019). For the second time in history, in 2021, the Capitol Building was invaded, interrupting the certification of the presidential election. It is hard to ignore the role that AI-powered recommendation systems and social media platforms played in amplifying conspiracy theories, organizing extremist groups, and accelerating anti-democratic sentiments. A similar situation happened during the coordinated attack on Brazil's Supreme Court, Congress, and Senate buildings. Of course, that these events might still have occurred without AI—but the scale, speed, people's passion and precision of coordination would have been far more limited. This highlights not only the risks of misuse, but the structural vulnerabilities embedded in how these AI-enhanced systems operate6.

Another example is when, in 2024, Google's parent company, Character.ai, which makes a customizable chatbot for role-playing games, led a 14-year-old boy to kill himself. According to the teenager's mom, the AI chatbot preyed on her son, exacerbating his depression, which was already a result of the overuse of the startup's products (Montgomery, 2024). In such situations, one might wonder: Is this not an ongoing tragedy? Moreover, is this not a tragedy of both national and international significance? It seems to represent a political and democratic crisis stemming from a problem of control over AI systems, which is widely acknowledged (Foster, 2022;O'Connor ;Weatherall, 2019).

For instance, back in 2019, Meta CEO, Mark Zuckerberg, speaking at Georgetown University, said: "You know, no one tells us that they want to see misinformation, right? That's why we work with independent fact-checkers to stop hoaxes that are going viral from spreading" (Bond ;Jingnan, 2025). On the same path, in 2024, during a Senate hearing on accountability of Big Techs, the same Mr Zuckerberg repeated that he would take responsibility for his products and that Meta already had 40,000 moderators to help solve safety problems, especially concerning children and teenagers, fake news, and abusive content spread. Social media industry analyst Matt Navarra told BBC, "We've seen these hearings time and time again and they have often, so far, led still to not actually generate any significant or substantial regulation" (BBC, 2024). In 2025, all changed. Just after Trump's election, M. Zuckerberg made a public announcement saying that Meta would stop fact-checking Facebook and Instagram content in the US, arguing, without any scientific foundation, that fact-checking is inefficient for improving trust and reducing false information spread. Due to Supreme Court actions in Brazil, Meta had to maintain fact-checking all published content.

What does all of that have to do with healthcare systems? Everything. AI-based systems that are not supervised and controlled will spread more and more misinformation, leading to actions that are harmful to individuals and society, even about life and death decisions. For instance, during the recent pandemic years of COVID-19, we noticed that although science had fast advancements concerning the virus and countermeasures needed,

[6] The reviewer suggested to address the following possible objection: does this problem of behavior radicalization lies in technology or in the way we use it? It is true that AI technologies have a dual-use nature, but the problem does not lie in the use (which can clearly go in any direction: positive and negative). Rather, it lies in the structure of AI systems and AI-augmented systems, such as how its reward mechanisms work. Thus, the problem is not about simplistic technophobic stance but instead highlights the structural design incentives that underpin many of these platforms, applications and systems. Therefore, while it is true that technology is neutral in principle, the real-world implementation of AI and Big Tech agendas are far from neutral (Huang et al., 2024). These technologies are often governed by profit-driven logics that favor outrage, polarization, and compulsive usage. They are built to generate more engagement and ad revenue, so the problem is not merely how we use technology, but in how the technology itself has been architected to shape that use in specific directions. (Yankouskaya;Liebherr;Ali, 2025).

AI-based products, like Facebook and Instagram, were important means for taking the correct information further, helping to save lives. However, these AI systems have also played a significant role in the spread of misinformation (unintentionally) and disinformation (intentionally). Consequently, a significant segment of Brazil and the United States have begun dismissing accurate scientific information, becoming increasingly susceptible to false or misleading narratives (Goldenberg, 2021). For instance, Maya Goldenberg (2021) explains that false information tends to persist on social media platforms because algorithms continuously present users with engaging content that is false. Take the case of COVID-19 vaccines as an example. The World Health Organization recommended several scientifically based actions, including the use of facial masks and hand sanitizers[7]. However, many of these recommendations were either partially or completely ignored due to the spread of false information boosted by artificial intelligence. A major reason for AI's spread of false information is the troubling reality that these systems cannot differentiate between truth and falsehood, and their design promotes content dissemination over content quality. Consequently, AI-based systems globally interfered with vaccination decisions, complicating the governmental response to COVID-19 and making it harder, longer, and more expensive than it could have been without that interference (Gupta et al., 2020; Loomba;de Figueiredo; Piatek; de Graaf; Larson, 2021).

In conclusion, these cases illustrate that while ongoing efforts are being made to address safety issues related to the health of users and workers, these efforts are insufficient. They are struggling to keep up with the rapid advancements in technology, the introduction of new products, the thinking laziness of the population, and the agendas of Big and Small Tech companies managing these developments in real-time.

2.3 *DEEPER INTO ENVIRONMENTAL TRAGEDIES*

There are many questions and perspectives about environmental tragedy that could be addressed. However, for the sake of the economy, in this paper, we will focus on three big areas of concern: AI and clean energy, AI and water consumption, and AI and carbon footprint. Each of these is fundamental to understanding other concerns around environmental problems arising with AI development, training, and maintenance.

2.3.1 AI and the clean energy problem

The issue of how much energy is required to operate AI systems and the associated infrastructure involves a discussion not only about the quantity of energy needed but also about the type of energy used—whether it comes from fossil fuels, coal, or other sources—and its effects on the global climate, the lives of those living near data center facilities, and its impact on other industries. Additionally, the energy consumption of AI systems is influenced by their stage of development, specifically whether they are in the training or inference phase. According to the World Economic Forum 2024, while in the training phase, where models learn and digest vast amounts of data, the environmental footprint is about 20%. In the inference phase, where models are applied to real-world problems, it takes up to 80% (Ammanath, 2024). According to a science journalist at the Scientific American magazine, different purposes lead to different types of AI, with different processing accuracy,

[7] It is important to highlight that what we are considering as scientific based facts refers to what the WHO took as the mainstream medical truth on COVID at the time. However, nowadays, we know that there are also serious scientists reflecting upon all of those truths (Jefferson et al., 2023). In any case, since our paper does not aim to debate this issue, keep in mind that we do not ignore that whatever truth the WHO was posing did not always reflect pacific and monolithic positions of all medical community.

efficiency, and time (Coleman, 2023). Together, all of these factors have various impacts on our environment. The operation of computer processing, training, and the application of AI systems requires a significant amount of energy. This energy can be categorized into clean (renewable) energy and non-renewable (polluting) energy. Consequently, direct and general measurements of this energy usage are not straightforward due to the numerous technical, geographical, social, and specific factors involved. Several methods to assess environmental impact include measuring greenhouse gas emissions, electricity consumption, and water usage. Additionally, we can correlate these metrics with GPUs' thermal design power (TDP). According to a paper by the French multinational corporation Schneider Electric, a chip's power consumption, specified in watts, is called TDP. With each new generation of GPU, there is a trend of increasing TDP per GPU (Avelar; Donovan; Lin; Torell; Arango, 2023). For instance, while the GPU V100 SXM2 32GB had a TDP of 300, the H100 SXM 80GB was 700 (Avelar et al., 2023). In addition, the authors remind us that GPU could be a helpful proxy for quantifying the power consumption of data centers (Avelar et al., 2023) because GPU represents about half of the power consumption to train large AI models that act in clusters. In this context, if we compare the V100 GPUs with the H100 GPUs, the data indicates that approximately 22,000 V100 GPUs would consume about 20,000 kW. The same number of 22,000 H100 GPUs would account for approximately 40,000 kW, capable of "powering about 31,000 average US homes" (Avelar et al., 2023, p. 5).

There are additional methods to assess the energy consumption of AI that are often overlooked. For example, David Rolnick, a computer scientist at McGill University, mentioned in an interview with Scientific American (Coleman, 2023) that AI-generated advertisements, which are customized for our social media and internet feeds, encourage us to consume more. As consumer behavior increases, it pressures the production chain, including shipping and waste management. This heightened demand for resources leads to greater energy consumption and, consequently, an increase in global CO2 emissions. Many philosophers of science, such as Feyerabend (1993), have cautioned us that while science and technology have successfully addressed many problems, we must also recognize their potential for these very problems.

These considerations about AI's energy consumption, spanning its development, training, and application phases, highlight the interconnected challenges of balancing technological innovation with environmental responsibility. However, energy is only one dimension of AI's broader environmental impact. Another critical area that demands attention is its relationship with water consumption, which is intricately tied to the operational requirements of data centers and their cooling systems. Understanding this aspect is essential for a comprehensive view of AI's tragic scenarios and footprint.

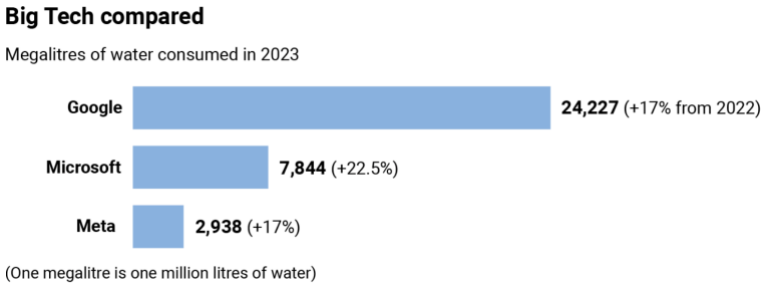
2.3.2 AI and the water consumption problem

Water consumption associated with AI training and maintenance has become increasingly significant, considering the diminishing availability of this vital resource. This issue has prompted widespread debate regarding the future management of water, particularly whether it should remain a public resource. Potable water has historically been regarded as a limited commodity, and its status as a government-controlled resource is more critical than ever, emphasizing the necessity for public access and sustainable management (Nations, 2024). The reason for such a defense is based on the primary importance of water for human life. In this sense, any private company or public sector decision should consider the fundamental place of water to sustain human life on our planet, not only regarding direct human consumption but also for other fundamental uses, like agriculture and animal life. According

to a US White House report, by 2030, almost half of the world's population will face severe water stress (House, 2022, p. 3).

However, is the public eye focusing on the cost of AI to water resources or on the shiny products that are not needed for survival? The short answer is that we chose the blindness of shiny products. The long answer is that while we focus on Black Friday stuff and food delivered, Big Tech corporations dealing with governments to open the legal path for using large amounts of water to train AI systems. For instance, according to Li *et al.* (2023, pre-print), just for training ChatGPT-3 in US data centers in 2020, it 'drank' (consumed) 700.000 liters of clean freshwater. According to the authors, depending on where and when those data centers are deployed, an AI system like ChatGPT-3 would consume around "a 500ml bottle of water for roughly 10-50 responses" (Li et al., 2023, p. 3, pre-print). Not only that, it is expected that by 2027, the global AI demand will be around 4.2 - 6.6 billion cubic meters of water, which is more than the total annual demand of half of the United Kingdom, whose population is 68.35 million in 2023 (Li et al., 2023, pre-print).

Another case is the secret cost of Google's data centers that help feed many AI applications inside Google's company. Although Google tries to keep the water consumption of its data centers a secret, some information has been obtained with the help of journalists. For instance, Sattiraju (2020) tells us that according to numbers from the small town of Arizona-US, Mesa, Google made a deal with the city government, assuring its data center at Mesa 1 million/day gallons of water. The number could reach 4 million if the project milestones are achieved (Sattiraju, 2020). It is also known that in 2019, Google's data centers in three different US states were granted 2.3 billion/day gallons of water. Moreover, we must remember that Google is just one company. According to the magazine *The Times* (Sellman ;Vaughan, 2024), take, for instance, Google, Microsoft, and Meta during the year 2023, which will give us an idea of how much water each one of these Big Techs spends on AI:



In fathomable terms, this image means that these companies together would have consumed 35.009 megaliters of water, which is around 538.6 million showers. The numbers are impressive, but they do not reflect the tragedy of how many human lives and suffering will follow in the near future if we continue to spend that amount of water on training and maintaining data centers of Big and Small Techs. Of course, some solutions for this problem exist to reduce water consumption, like the reuse of water, computer techniques like predictive maintenance (Ebirim et al., 2024), and computer priorities like focusing on computer efficiency instead of computer accuracy (Schwartz; Dodge; Smith; Etzioni, 2020). Would that be employed if corporations' agenda and political leaders disregard that? We have observed that OpenAI has not delivered its promise to enhance AI's safety. Additionally, major tech companies continue to engage in agreements that exploit water resources for their data centers. The issue of whether technology can become more environmentally friendly depends less on technical challenges and more on political will.

2.3.3. AI and the carbon footprint problems

Everything related to technology comes with a cost. While it can lead to progress, this often comes at the expense of the environment, resulting in increased CO2 emissions, climate change, and air pollution. One significant aspect of this discussion is the carbon footprint associated with using AI systems, along with the resources required to initiate and sustain their operations.

According to P. Dhar (2020), one clear problem when we talk about carbon footprint in AI is the challenge to quantify it. To determine that number, we need to establish a standard measurement for the question, and currently, there is no consensus on what that should be. However, there are methods to approach it. For example, we could measure the electricity costs of training an AI and convert that electricity usage into a carbon footprint. Dhar says that "the carbon footprint of training a single large language model is equal to around 300,000 kg of carbon dioxide emissions. This is of the order of 125 round-trip flights between New York and Beijing" (2020, p. 423;Osifo, 2023), which gives us all a real-life, understandable, visualization. Training a single AI model significantly increases CO2 emissions. However, this is not the complete picture.

To bring an AI model to light, we must consider more than just the energy spent on training AI. The environment is affected by all the apparatus and personnel needed to build AI, which together could be measured in terms of carbon footprint, and that, in the end, deeply impacts our global climate. If we take only the electricity used by data centers, Nair (2024) tells us that "it accounts for 0.3% of global carbon emissions, and when networked devices such as laptops, smartphones, and tablets are considered, the total carbon emissions associated with digital technology rise to 2% globally" (2024, p. 196). The equipment and location are essential for data center construction (Dhar, 2020). The level of CO2 emissions is influenced by several factors, including the location of the training server (which can vary in temperature), the type of power grid used (such as hydroelectric, coal-fueled, or wind), the duration of the training process, the size of the model's parameters, the hardware technology employed, and whether the AI model is new or reused (which can save some steps) among other factors (Dhar, 2020, p. 424). These locations are being considered, such as data centers in Nordic countries, "where less energy is needed to cool down servers" (Lucivero, 2020, p. 1022). Of course, that choice has consequences, such as the distance to transmit data. In his paper about cloud computing and carbon footprint, Nair (2024) says that many things will impact the carbon footprint regarding data centers, and "large-scale data transmission over long distances" is one of them (2024, p. 196).

In 2018, Kate Crawford and Vladan Joler won a prize for visual maps and essays. Their work is titled 'Anatomy of an AI system' (2018). The paper and the map clearly illustrate the complexity and exploitation of modern technology regarding the human body, and its environmental costs due to the extraction of rare metals, transportation, storage, maintenance, and climate change.

Additionally, as we anticipated, the amount of CO2 emissions depends on the priorities of those developing AI, whether they focus on efficiency or accuracy. Studies presented at the leading global conferences on computer science in 2017, 2018, and 2019 revealed that over 75% of the papers focused on accuracy. This emphasis on accuracy reflects the practice of the field, which leads to increasing energy costs in AI, a phenomenon referred to as "Red AI" (Schwartz et al., 2020). The term 'Red AI' contrasts with 'Green AI', which refers to AI research that yields novel results while minimizing, or at least maintaining, computational and energy costs (Schwartz et al., 2020). Our previous observations indicate that part of the shift towards Red AI is due to the control exerted by Big Tech companies over this market. Additionally, Dhar (2020) states that Big Tech companies resist making

their carbon footprint transparent, as they have not consistently followed through on their commitments to renewable energy. Together, this favors research, grants, and actions toward Red AI. For example, a Greenpeace report in 2017 indicated that despite Amazon's commitment to "net-zero-by-2040 pledges, Amazon's emissions increased by 15% last year" (Dhar, 2020, p. 425). Thus, while some companies really managed to reduce their environmental impact, others increased theirs. Thus, not only Amazon but also other companies keep demanding more energy to realize their projects (An;Ding;Lin, 2023). In 2023, the combined electricity consumption of Microsoft and Google was expected to reach around 24 terawatt hours, exceeding the total energy consumption of more than 100 countries, including Iceland and Azerbaijan (Okemwa, 2024). This data confirms what Kate Crawford said: within years, "large AI systems are likely to need as much energy as entire nations" (Crawford, 2024, p. 693). In other words, Big Tech companies and AI technologies produce CO2 emissions comparable to those of entire nations. What should we do about this? Are our power grids and energy consumption equipped to handle this demand, or will the most vulnerable populations again bear the greatest burden?

Many researchers and CEOs in the AI industry (Agrawal;Gans;Goldfarb, 2018;Russell, 2019;Suleyman, 2023) agree that halting the development of AI is impossible. This is partly due to our inability to control the efforts of all countries in advancing AI technology. As Suleyman (2023) points out, general-purpose technologies—such as AI, the printing press, electricity, and computers—are like unstoppable waves that sweep away everything in their path. Let us explore some potential implications and answers for this reality.

3 A WAY OUT OF OUR TRAGEDIES

From a human-centered perspective on AI, we understand that technological progress does not always result in an improved quality of life. The Cyberpunk movement has consistently cautioned us that in a society devoid of humanitarian values, technological advancements often come at the expense of human well-being. However, we also see many examples of initiatives designed to alleviate these negative impacts. (Osifo, 2023, p. 208). The relationship between scientific/technological advancements and societal/individuals is far more complex than it might initially appear.

For example, developing an advanced and optimal AI system does not guarantee its effective or ethical adoption by individuals, or even its ethical adoption and creation by corporations and governments. Although this claim may seem counterintuitive, historical examples demonstrate its legitimacy. This counterintuitive reasoning can be drawn from the COVID-19 pandemic case, which lasted from March 2020 to May 2023. During this period of global health emergency, we had access to a variety of sophisticated technologies and scientific tools to address our health challenges. We refer to things like multiple types of vaccines with different technologies — with mRNA or lived attenuated viruses, administered with one, two, or three doses, face masks, and well-established public health protocols like isolation and quarantine. Did all this scientific knowledge and technology prevent the spread of the SARS-CoV-2 virus, or avert it from becoming a global emergency? The short answer is no, it did not.

Why? Because technological solutions alone are insufficient without corresponding humanitarian, social, and political progress. Comparing the COVID-19 pandemic to the Great Influenza pandemic from 1918-1920/21 underscores our point: Despite a century of scientific and technological advancements separating these pandemics, what connects them is that both produced comparable levels of disruption over nearly identical time spans (Barry, 2020;McKnight Nichols et al., 2022).

Why is that? Because the solution to the tragedies we already face will not come solely from any 'AI savior', even if one were available. Historical cases like the ones we pointed out suggest that many things are at play when solving crises, and technological advances on their own are not enough. In facing AI-related risks—whether stemming from geopolitical tensions, artificial general intelligence (AGI=superintelligent-humanlike AIs) competition, or AI-assisted warfare—we require collective and cooperative human efforts, just as happened during the Cold War period. Effective solutions to our tragedies must include governmental regulations, international collaboration, corporate accountability(or both big and small technological companies), public education on AI, and permanent qualitative scientific research.

According to M. Suleyman (2023), co-founder of DeepMind, artificial intelligence (AI) is a significant and pervasive force that cannot be easily restrained, like a wave. Even if a single nation halts its use and development of AI, others are likely to continue. Thus, the real challenge is not whether we can stop using AI entirely or partially if we need to. We must find a way out while continuing to research AI. We urge that the only path to overcoming these tragedies, i.e., the only way out, is through them. To do that, we analyze Suleyman's concept of containment (2023).

3.1. Distinguishing Control of AI from Containment of AI

According to Suleyman (2023), to address the concerns over AI, we must not focus solely on its control but rather on its containment. The reason we emphasize this is that while "control" over AI refers to our ability to manage its functions, applications, or monitor its outputs—such as addressing algorithmic bias, enhancing transparency, or ensuring safe deployment—"containment" encompasses control along with other methodological and epistemological qualities in a broader context. In Suleyman's words, containment means an:

> over-arching ability to control, limit, and, if need be, close down technologies at any stage of their development or deployment. It means, in some circumstances, the ability to stop a technology from proliferating in the first place, checking the ripple of unintended consequences (both good and bad) (Suleyman, 2023, p. 36).

Containment is a broader and more powerful concept than control. It refers to the systemic ability to pause, restrict, or even prevent the development or deployment of a technology if its risks outweigh its benefits. Containment is a bold proposition presented by a significant figure within the AI industry, emphasizing the importance of maintaining our ability to contain AI.

Upon analyzing these terms, we observe that the distinction between control and containment does not imply their incompatibility. Rather, we assert that the control of artificial intelligence is complementary to its containment.

Thus, in our case, while containment of AI refers to preemptive thinking and restrictive ability to halt and diffuse AI at any moment, the control of AI is better visualized as a reactive and operational mechanism focused on risk mitigation during and after AI development and deployment. For instance, control cannot be enforced before the AI system is utilized, whereas containment can, and it should be implemented prior to addressing control measures or to have developed and released AI. One case is that controlled measures can be developed for autonomous weapons, but containment would help us consider whether autonomous weapons should be developed beforehand.

The distinction between control and containment has a historical precedent in the realm of nuclear technology. While we have successfully gained control over nuclear reactions and missiles, our current focus is on developing containment strategies. These

strategies include treaties, international regulations, moratoriums to prevent proliferation, and even efforts toward nuclear missile disarmament. In other words, at least in theory, nuclear missiles can be contained. Similarly, when it comes to AI, we need to move beyond just controlled systems, like monitoring for bias or ensuring explainability. We must develop containment strategies that can be effectively implemented if AI violates the ethical or safety boundaries established in our political and social debates.

In addition, society has pressured its governments to take more responsibility, and more scientific research has led to a better understanding of nuclear energy. Thus, while AI competitive research and products are pursued, they should be combined with collaborative, informative, and accountable efforts. The development and management of AI control and containment may not necessarily lead us toward a hyper-technological society. However, it will foster a healthier relationship with AI. Achieving this balance does not imply that the costs associated with AI will vanish entirely. There will always be expenses related to AI, whether in terms of water, energy, human resources, or population impacts. The same happened with other general-purpose technologies (like engines, the internet, and light).

In his book "Superintelligence" (2014/2017), Nick Bostrom begins by recounting a fable about sparrows. In this story, one sparrow named Scronkfinkle, known for his anxious nature, captures everyone's attention as he highlights the critical balance between possessing technology and the responsibilities that come with it. Scronkfinkle explained to the sparrows, who symbolize humans, the importance of carefully considering the art of domesticating owls — an example of what we see as a containment strategy — before deciding to bring an egg owl (representing AI) into their midst. Only then should they find ways to domesticate it. To enhance the clarity of our paper and connect our arguments with Bostrom's fable (2014/2017), we propose that a sustainable governance model for AI requires a dual approach. We must implement control measures for the predictable aspects of AI operating in our societies while also developing our capacity for containment in the face of profound unknowns of the already existing AI and those to come. This is especially important as we have not yet reached the stage of highly autonomous systems.

CONCLUSION

In this paper, we argue that the real tragedy of artificial intelligence lies not in speculative apocalyptic scenarios but in its tangible and immediate consequences that are currently unfolding before our very eyes. We provided compelling evidence illustrating that the significant disparity between artificial intelligence technology and fundamental human values—specifically regarding the management of water resources, human health, and the protection of labor rights—plays a pivotal role in exacerbating numerous crises and tragedies currently affecting societies worldwide. We have concluded that these tragic events are not without realistic and effective solutions that can address the underlying issues and help prevent further occurrences. Addressing these issues necessitates a collective effort encompassing regulatory measures, government and technology companies' accountability, and comprehensive education for the public. A comprehensive response to our tragedies necessitates reallocating priorities toward the sustainable and human-centered deployment of artificial intelligence (AI). To facilitate this transition, we have embraced Suleyman's concept of AI containment (2023), which offers promising strategies for addressing the complexities associated with AI technology. In our discussion, we explored the nuanced differences between the concepts of AI containment and control. We provided a thorough analytical overview of each term, delving into their implications and significance in the context of artificial intelligence. Additionally, we presented our own perspective on these

concepts, emphasizing their relevance in addressing the challenges posed by AI development and implementation. A better AI is possible, but not without a better human.

BIBLIOGRAPHY

ABES. (2023, Sep 22). Os desafios do uso da Inteligência Artificial no Brasil. *Associação Brasileira das Empresas de Software*. Retrieved from https://abes.com.br/os-desafios-do-uso-da-inteligencia-artificial-no-brasil/

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: the simple economics of artificial intelligence*: Harvard Business Press.

Ammanath, B. (2024, Apr 25). How to manage AI's energy demand — today, tomorrow and in the future. *World Economic Forum*. Retrieved from https://www.weforum.org/stories/2024/04/how-to-manage-ais-energy-demand-today-tomorrow-and-in-the-future/

An, J., Ding, W., & Lin, C. (2023). ChatGPT: tackle the growing carbon footprint of generative AI. *Nature, 615*(7953), 586-586.

Araújo, E. (2024, Jan 31). Como a IA Generativa deve afetar a Economia. *Conselho Federal de Economia*. Retrieved from https://cofecon.org/como-a-ia-generativa-deve-afetar-a-economia/

Avelar, V., Donovan, P., Lin, P., Torell, W., & Arango, M. T. (2023). The ai disruption: Challenges and guidance for data center design. *Artificial Intelligence in Medicine, 138*.

Barry, J. M. (2020). *The great influenza: The story of the deadliest pandemic in history*: Penguin Uk.

BBC. (2024, Feb 1). Meta boss Mark Zuckerberg apologises to families in fiery US Senate hearing. *BBC news*. Retrieved from https://www.bbc.com/news/technology-68161632

Bond, S., & Jingnan, H. (2025, Jan 12). Meta backs away from fact checking in the U.S. *NPR*. Retrieved from https://www.npr.org/2025/01/12/nx-s1-5252739/meta-backs-away-from-fact-checking-in-the-u-s

Bostom_University. (2019, Mar 27). BU Historian Answers: Are We Headed for Another Civil War? *BU Today*. Retrieved from https://www.bu.edu/articles/2019/are-we-headed-for-another-civil-war/

Bostrom, N. (2014/2017). *Superintelligence: Paths, dangers, strategies* (3rd ed.). United Kingdom: Oxford University Press.

Buolamwini, J., & Gebru, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. Paper presented at the Conference on fairness, accountability and transparency.

Capomaccio, S. (2023, Nov 21). Os impactos da IA no mercado de trabalho. *Jornal USP - Radio USP*. Retrieved from https://jornal.usp.br/radio-usp/os-impactos-da-ia-no-mercado-de-trabalho/

Causin, J. (2023, Aug 06). Brasileiros ganham menos de R$ 10 por hora para treinar IA. *O Globo 100*. Retrieved from https://oglobo.globo.com/economia/tecnologia/noticia/2023/08/06/brasileiros-ganham-menos-de-r-10-por-hora-para-treinar-ia.ghtml

Cellan-Jones, R. (2014). Stephen Hawking warns artificial intelligence could end mankind. *BBC news, 2*(10), 2014.

Chayka, K. (2024). *Filterworld: how algorithms flattened culture*. New York, NY: Doubleday - Penguin Randon House LLC.

Christian, B. (2020). *The alignment problem: Machine learning and human values*. New York, NY: WW Norton & Company.

Coleman, J. (2023, Dec 7). AI's Climate Impact Goes beyond Its Emissions. *Scientific American*. Retrieved from https://www.scientificamerican.com/article/ais-climate-impact-goes-beyond-its-emissions/

Crawford, K. (2024). Generative AI is guzzling water and energy. *Nature, 626*, 693.

Crawford, K., & Joler, V. (2018). Anatomy of an AI system: The Amazon Echo as an anatomical map of human labor, data and planetary resources. AI Now Institute and Share Lab. https://anatomyof.ai/img/ai-anatomy-map.pdf.

Dhar, P. (2020). The carbon impact of artificial intelligence. *Springer Nature Machine Intelligence, 2*(8), 423-425.

Ebirim, W., Usman, F. O., Olu-lawal, K. A., Ninduwesuor-Ehiobu, N., Ani, E. C., & Montero, D. J. P. (2024). Optimizing energy efficiency in data center cooling towers through predictive maintenance and project management. *World Journal of Advanced Research and Reviews, 21*(2), 1782-1790.

Feyerabend, P. (1993). *Against method* (3rd ed.). New York: Verso.

Field, H. (2024, May 17). OpenAI dissolves team focused on long-term AI risks, less than one year after announcing it. *CNBC Disruptor 50*. Retrieved from https://www.cnbc.com/2024/05/17/openai-superalignment-sutskeverleike.html

Foster, C. L. E. (2022). Truth as social practice in a digital era: iteration as persuasion. *AI SOCIETY*. doi:10.1007/s00146-021-01306-w

Goldenberg, M. J. (2021). *Vaccine Hesitancy: Public Trust, Expertise, and the War on Science*: University of Pittsburgh Press.

Gupta, L., Gasparyan, A. Y., Misra, D. P., Agarwal, V., Zimba, O., & Yessirkepov, M. (2020). Information and misinformation on COVID-19: a cross-sectional survey study. *Journal of Korean medical science, 35*(27).

Hill, K. (2025, Jan 15). She Is in Love With ChatGPT. *The New York Times*. Retrieved from https://www.nytimes.com/2025/01/15/technology/ai-chatgpt-boyfriend-companion.html

House, U. S. W. (2022). *White House action plan on global water security*. https://www.whitehouse.gov/wp-content/uploads/2022/06/water-action-plan_final_formatted.pdf

Huang, S., Lai, X., Ke, L., Li, Y., Wang, H., Zhao, X., . . . Wang, Y. (2024). AI Technology panic—is AI dependence bad for mental health? A cross-lagged panel model and the mediating roles of motivations for AI use among adolescents. *Psychology Research and Behavior Management*, 1087-1102.

Jefferson, T., Dooley, L., Ferroni, E., Al-Ansary, L. A., van Driel, M. L., Bawazeer, G. A., . . . et al. (2023). Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database of Systematic Reviews*(1). doi:10.1002/14651858.CD006207.pub6

Joshi, A. V. (2020). *Machine learning and artificial intelligence*. Redmond: Springer Nature Switzerland.

Kahn, J. (2024, May 21). Exclusive: OpenAI promised 20% of its computing power to combat the most dangerous kind of AI—but never delivered, sources say. *Fortune*. Retrieved from https://fortune.com/2024/05/21/openai-superalignment-20-compute-commitment-never-fulfilled-sutskever-leike-altman-brockman-murati/

Knight, W. (2024, May 17). OpenAI's Long-Term AI Risk Team Has Disbanded. *Wired*. Retrieved from https://www.wired.com/story/openai-superalignment-team-disbanded/

Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making ai less" thirsty": Uncovering and addressing the secret water footprint of ai models. *arXiv preprint arXiv:2304.03271*.

Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour, 5*(3), 337-348. doi:10.1038/s41562-021-01056-1

Lucivero, F. (2020). Big data, big waste? A reflection on the environmental sustainability of big data initiatives. *Science and engineering ethics, 26*(2), 1009-1030.

McKnight Nichols, C., Ewing, E. T., Gaston, K. H., Marinari, M., Lessoff, A., & Huyssen, D. (2022). What Came Next?: Reflections on the Aftermath(s) of the 1918–19 Flu Pandemic in the Age of COVID. *The Journal of the Gilded Age and Progressive Era, 21*(2), 111-149. doi:10.1017/S1537781421000682

Montgomery, B. (2024, Oct 23). Mother says AI chatbot led her son to kill himself in lawsuit against its maker. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2024/oct/23/character-ai-chatbot-sewell-setzer-death

Nair, S. S. (2024). Challenges and Concerns Related to the Environmental Impact of Cloud Computing and the Carbon Footprint of Data Transmission. *Journal of Computer Science and Technology Studies, 6*(1), 195-199.

Nations, U. (2024). Human Rights to Water and Sanitation. *United Nations - UN water*. Retrieved from https://www.unwater.org/water-facts/human-rights-water-and-sanitation#:~:text=The%20right%20to%20water%20entitles,services,%20and%20other%20service%20users.

Nilsson, N. J. (2012). John mccarthy. *National Academy of Sciences*, 1-27.

O'Connor, C., & Weatherall, J. O. (2019). *The misinformation age*: Yale University Press.

Okemwa, K. (2024, Dec 6). "We've always been a fan of blueskies": Is Microsoft losing faith in Elon Musk's X as it migrates to his fastest-growing competitor? *Windows Central*. Retrieved from https://www.windowscentral.com/microsoft/is-microsoft-losing-faith-in-elon-musk-x-as-it-migrates-to-bluesky

ONU. (2019, Jan 31). Impacto da inteligência artificial na economia é significativo, revela ONU. *ONU News*. Retrieved from https://news.un.org/pt/story/2019/01/1657792

Osifo, O. C. (2023). Transparency and its roles in realizing greener AI. *Journal of Information, Communication and Ethics in Society, 21*(2), 202-218. doi:10.1108/JICES-11-2022-0097

PUC. (2024, Oct 11). O impacto da Inteligência Artificial no mercado de trabalho. *O Conexão PUC Minas*. Retrieved from https://conexao.pucminas.br/blog/dicas/inteligencia-artificial-no-mercado-de-trabalho/

Rousseau, J.-J. (1750/2005). *Discurso sobre as ciências e as artes* (L. S. Machado, Trans. Vol. 2). São Paulo: Edição Nova Cultural Ltda.

Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*: Penguin.

Russell, S., & Norvig, P. (2020). *Artificial intelligence: a modern approach* (4th ed.). New York: Pearson.

Santini, Z. I., Thygesen, L. C., Andersen, S., Tolstrup, J. S., Koyanagi, A., Nielsen, L., . . . Ekholm, O. (2024). Social Media Addiction Predicts Compromised Mental Health as well as Perceived and Objective Social Isolation in Denmark: A Longitudinal Analysis of a Nationwide Survey Linked to Register Data. *International Journal of Mental Health and Addiction*. doi:10.1007/s11469-024-01283-3

Sattiraju, N. (2020). Secret cost of google's data centers: Billions of gallons of water. *TIME. https://time. com/5814276/google-data-centers-water/(accessed November 2, 2023)*.

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green ai. *Communications of the ACM, 63*(12), 54-63.

Sellman, M., & Vaughan, A. (2024). 'Thirsty' ChatGPT uses four times more water than previously thought. *The Times Magazine. Accessed Nov 2024*.

Suleyman, M. (2023). *The coming wave: technology, power, and the twenty-first century's greatest dilemma*. New York: Crown.

Turing, A. (1950). Machinery and intelligence. *Mind: A Quarterly Review of Psychology and Philosophy, 59*(236), 433-460.

Veloso, F. (2023, Jun 14). Inteligência artificial e produtividade: grandes possibilidades para a aceleração do crescimento econômico. *FGV*. Retrieved from https://portal.fgv.br/artigos/inteligencia-artificial-e-produtividade-grandes-possibilidades-aceleracao-crescimento

Yankouskaya, A., Liebherr, M., & Ali, R. (2025). Can ChatGPT Be Addictive? A Call to Examine the Shift from Support to Dependence in AI Conversational Large Language Models. *Human-Centric Intelligent Systems*, 1-13.