Human & Social Sciences Journal

Volume 11, Issue 01, June 2025, P.P. 61-71

Ethical framework for the Responsible Use of Artificial Intelligence

الإطار الأخلاقي للاستخدام المسؤول للذكاء الاصطناعي

Dr.karima GHEDIRI[*1]

Date of submission 15/11/2024 - Date of acceptation 26/05/2025- Date of edition: 11 /06 /2026

ملخص

مع تزايد تأثير الذكاء الاصطناعي على مختلف القطاعات، برزت مخاوف أخلاقية بشأن تطبيقاته. فقد كشفت العديد من الحالات في مجال الرعاية الصحية والتكنولوجيا والسياسة الاجتماعية والإدارة البيئية وغيرها عن تحيزات داخل أنظمة الذكاء الاصطناعي، مما أدى في بعض الاحيان إلى استبعاد أو تهميش الأفراد والجماعات العرقية وحتى مجتمعات بأكملها. وتؤكد هذه القضايا على الحاجة الملحة إلى اتخاذ إجراءات تصحيحية إلى جانب الالتزام بالمبادئ الأخلاقية. لذلك، يستكشف هذا البحث بعض التأثيرات السلبية للذكاء الاصطناعي ويفحص المبادئ الأخلاقية كتدابير استباقية لمنع الذكاء الاصطناعي من تكريس الصور النمطية الضارة والتحيز. ومن خلال دراسات حالات عديدة، كما يدعو إلى إطار أخلاقي شامل لتوجيه التنمية المسؤولة والعادلة لتقنيات الذكاء الاصطناعي.

الكلمات المفتاحية: الذكاء الاصطناعي، اخلاقيات، التحيز، الذكاؤ الاصطناعي المسؤول.

Abstract

As Artificial Intelligence continue to influence diverse sectors, significant ethical concerns about its applications have emerged. Numerous cases in healthcare, technology, social policy, environmental management, and beyond have revealed biases within AI systems, or humain-driven AI applications, leading sometimes to the exclusion or marginalization of individuals, ethnic groups, and even entire communities, while in other instance they have caused notable environemental and economic harm. These issues underscore the urgent need for corrective actions alongside adherence to ethical guidelines. This paper, therefore, explores some negative impacts of AI and examines ethical principles as proactive measures to prevent AI from perpetuating harmful stereotypes and injustices. Through case studies, it advocates for a robust ethical framework to guide the responsible and equitable development of AI technologies.

Keywords : Artificial intelligence, Ethics, Bias, Responsible AI**.**

Résumé Alors que l'intelligence artificielle continue d'influencer divers secteurs, des préoccupations éthiques significatives concernant son application ont émergé. De nombreux cas dans les domaines de la santé, de la technologie, des politiques sociales, de la gestion environnementale, et au-delà, ont révélé des biais au sein des systèmes d'IA, ou des applications d'IA pilotées par des humains, conduisant parfois à l'exclusion ou à la marginalisation d'individus, de groupes ethniques, voire de communautés entières. Dans d'autres cas, elles ont causé des dommages environnementaux et économiques notables. Ces problématiques soulignent l'urgence de prendre des mesures correctives tout en respectant des lignes directrices éthiques. Cet article explore donc les impacts négatifs de l'IA et examine les principes éthiques en tant que mesures proactives pour empêcher l'IA de perpétuer des stéréotypes nuisibles et des injustices. À travers des études de cas, il plaide pour un cadre éthique solide afin de guider le développement responsable et équitable des technologies d'IA.

Mots clés: Intelligence artificielle, Éthique, Biais, IA responsable

*

[1] Karima ghediri: **National Higher school of Journalism and Information Sciences, Algeria, MUSC, Ghediri.karima@ensjsi.dz**

I. Introduction :

Artificial intelligence (AI) has revolutionized various industries and has achieved many great things, such as facial recognition, medical diagnosis, and self-driving cars. AI promises enormous benefits for economic growth, social development, as well as human well-being and safety improvement.

However, alongside these positives, AI brings some unintended side effects. One significant concern is that AI systems can unintentionally reinforce biases present in their training data. This means that, despite their advanced capabilities, they might perpetuate discrimination and unfairness, especially towards marginalized communities.

These challenges highlight a critical problem: without proper oversight, AI could deepen existing social inequalities. So, how can we address these harms? The answer lies in adopting practical measures and clear regulations that guide AI's development and use..

Establishing well-defined regulations ensures that AI systems adhere to standards that protect individuals and promote fairness.

Moreover, embracing an ethical framework is essential. By prioritizing principles like fairness, accountability, and transparency, we can embed ethical considerations directly into AI's design and implementation. This proactive approach doesn't just prevent negative outcomes—it actively promotes positive ones, helping to minimize AI's potential downsides.

In combining ethical principles with practical regulations, we pave the way for AI systems that are not only innovative but also socially responsible. This way, AI can become a tool that supports equality and justice, rather than undermining them.

II. Case Studies of AI Bias: Real-World Impacts and Consequences

In recent years, artificial intelligence (AI) has become deeply embedded in many aspects of daily life, influencing decision-making processes across critical domains. While its applications in healthcare, justice, and environmental sustainability demonstrate its potential, these same domains also expose the ethical vulnerabilities of AI. The widespread deployment of AI has revealed its dual capacity: it can improve outcomes, yet also amplify existing social disparities. However, alongside these positives, AI brings some unintended side effects. One significant concern is that AI systems can unintentionally reinforce biases present in their training data. This means that, despite their advanced capabilities, they might perpetuate discrimination and unfairness, especially towards marginalized populations. For instance, AI-powered recruitment tools have in some cases excluded qualified applicants due to the presence of historical bias in the data used for training.

AI systems have demonstrated bias in various sectors, such as healthcare and criminal justice. One prominent example is the « Correctional Offender Management Profiling for Alternative Sanctions » (COMPAS) tool utilized in the U.S. criminal justice system to predict the probability of a defendant reoffendingit . It has faced criticism for potential bias and lack of transparency. Critics argue that its proprietary algorithm, which is not fully disclosed, may perpetuate racial and socioeconomic disparities. In a 2016 investigation, ProPublica revealed that the COMPAS system disproportionately labeled Black defendants as high-risk compared to white defendants with comparable profiles. This hughlighted how biases embedded within machine learning algorithms can have real-life consequences for individuals' lives and perpetuate systemic inequalities within the criminal justice system. (Mensah, 2023)

Another telling example involves Joy Buolamwini, a leading researcher in AI ethics known for her pioneering work on exposing algorithmic bias. She is a computer scientist, activist, and the founder of the Algorithmic Justice League (AJL), an organization dedicated to advocating for ethical and accountable use of artificial intelligence. As a graduate student at MIT, she created a mirror that would project aspirational images onto her face, such as a lion or tennis star Serena Williams. But the facial-recognition software she installed wouldn’t work on her Black face, until she literally put on a white mask. (Wood, 2021) She later explained that many AI tools (especially facial recognition) depend on labeled training data, which can encode and replicate existing societal biases. It has recently been shown that algorithms trained with biased data have resulted in algorithmic discrimination (Buolamwini & Gebru, 2018).

The research gained widespread attention, highlighting the systemic biases embedded in AI systems and the potential for these biases to perpetuate discrimination. Buolamwini emphasizes the importance of diversity in AI design and development.and advocates for ethical AI practices, underscoring that technology should serve all of humanity equitably, without reinforcing existing social inequalities.

Another example of bias in In healthcare, an AI system used to predict patient mortality rates wich was found to be biased against African- American patients. A study conducted by Obermeyer et al. (2019) revealed that the system was more likely to assign higher-risk scores to African-American patients, even when other factors, such as age and health status, were identicle. This bias can result in African-American patients being denied access to necessary healthcare or receiving substandard treatment. (Ferrara, 2023). So the underlying Cause: is the algorithm relied on healthcare costs as a proxy for health risk. However, systemic inequities in healthcare access and socioeconomic disparities mean that African-American patients often incur lower healthcare costs than White patients with similar levels of need. As a result, the AI system underestimated the health risks of African-American patients.

Amazon's recruitment algorithm trained on resumes mainly from male applicants due to historical hiring trends. As a result, the algorithm developed a preference for resumes resembling those of existing Amazon employees, who were predominantly male. This implicit bias led the algorithm to systematically downgrade resumes that included terms or experiences associated with women, resulting in discrimination against female candidates and making it harder for qualified women to succeed in the recruitment process.

(Dastin, 2022).

This case illustrates how biased training data and implicit biases in algorithm development can lead to discriminatory AI outcomes. Amazon's recruitment algorithm demonstrated gender bias by favoring male candidates, emphasizing the critical importance of fairness and diversity in AI training data. (Min, 2023).

Bas in AI can perpetuate gender stereotypes and discrimination in various contexts, For exemple, facial recognition algorithms trained primerly on male data often struggle to recognize female faces, causing gender bias. Smilarly when generative AI (GenAI) models are prompted to create images of CEOs, they tend to reinforce stereotypes by depicting CEOs predominantly as men (Ferrara, 2023).

This case highlights the challenges of using AI in recruitment without addressing biases in training data. Amazon's experience highlights the necessity for transparency, fairness, and accountability in AI systems, particularly in employment contexts, where biased algorithms can reinforce existing disparities.

Bostrom conducted a study on superintelligence, examining the paths, dangers, and strategies associated with AI development. The study highlights critical concerns about the risks of creating superintelligent AI, emphasizing the potentially catastrophic consequences if such systems are not aligned with human values. Bostrom advocates for prioritizing safety and ethical considerations in AI development and proposes various strategies to achieve this goal. However, a notable limitation of Bostrom's study is its narrow focus on the risks of superintelligent AI, without offering a comprehensive framework for addressing broader ethical considerations in AI. Additionally, some critics have described Bostrom's arguments as overly pessimistic and speculative. (Huriye, 2023)

The growth of (AI) has raised concerns about the environmental costs of digital technologies. As AI models become increasingly large and energy-intensive, critics warn of their significant carbon footprint and disproportionate impact on vulnerable communities. A Forbes article captured this concern poignantly,stating, “Somewhere in America, a lump of coal is burned every time a book is ordered online,” (Castro, 2024). The authors estimated that, “half of the electric grid will be powering

the digital-Internet economy within the next decade, a claim that has been widely cited which became widely cited in debates about energy policy.

Some critics warn that AI's spread and the size of its models could raise energy use and harm the environment. For exemple: researchers estimated that training a model to improve English-German machine translation generated 626,155 pounds of CO2 emissions -roughly equivalent to300 roundtrip flights from the East Coast to the West Coast, (Castro, 2024).

This figure underscores growing concerns about the environmental sustainability of AI. While AI has transformative potential, its development comes with significant ecologicalcosts, raising questions about howto balance innovation with environmental responsibility.

Beyond environmental impacts, the ethical implications of AI development have come under scrutiny. Some researchers argue that the deployment of increasingly large AI models by wealthy Western nations constitutes a form of "environmental racism," as these systems have disproportionate burden placed on poorer communities in the Global South. They ask pointedly: «Is it fair or just to require, for instance, that the residents of the Maldives—likely to be submerged by 2100—or the 800,000 people in Sudan affected by devastating floods bear the environmental costs of training and deploying increasingly large English-language models, while no comparable models are being developed for languages like Dhivehi or Sudanese Arabic ?» (Castro, 2024).

This critique highlights a serious issue: training AI systems is both environmentally harmful and a form of racial injusti, and addressing these challenges requires a multifaceted approach, including diversifying training datasets, implementing fairness audits, and ensuring human oversight in AI decisions. There is a cruciel need to incorporate ethical considerations into the design and implementation of AI systems. This means that developers and policymakers need to be proactive in identifying and addressing potential ethical concerns, such as fairness, transparency, and accountability. This is particularly important to ensure that users are able to understand how decisions are made, and to identify and address potential biases or errors in the technology. ensuring that it is consistent with local values and expectations is essential to fostering trust and equitable use. By balancing innovation with ethical and environmental responsibility, the AI community can work toward a more sustainable and just technological future.

These challenges highlight a critical problem: without proper oversight, AI can exacerbate existing social inequalities. So, what can we do to address these harms? We need concrete, practical measures and robust regulations to guide the development and use of AI. Implementing clear guidelines ensures that AI technologies are held to standards that protect individuals and promote fairness.

Moreover, embracing an ethical framework is essential. By prioritizing principles like fairness, accountability, and transparency, we can embed ethical considerations directly into AI's design and

implementation. This proactive approach doesn't just prevent negative outcomes—it actively promotes positive ones, helping to minimize AI's potential downsides.

In combining ethical principles with practical regulations, we pave the way for AI systems that are not only innovative but also socially responsible. Together, we can work towards a technological landscape where AI contributes to equality and justice rather than detracting from them.

III. The Importance of Ethical Considerations in AI

People's well-being must come first when developing technology. Ethics should guide how we build and use AI (Floridi, 2019). Ethical guidelines and boundaries should serve as the foundation for technological innovation to ensure that the advancement of AI does not jeopardize humanity's welfare. By adhering to ethical principles, developers can help prevent potential harm and mitigate risks, thus protecting the well-being of all individuals.

One of the most significant roles ethics play is in building and maintaining public trust. Ethical behavior fosters trust, while unethical actions—such as dishonesty, manipulation, or injustice—can quickly erode it. In the realm of AI, where systems are designed to make important decisions that impact people's lives, trust is vital as seen earlier in real world cases. Fair and clear AI builds public trust. But if AI is unfair or invasive, it harms people and breaks that trust (raji & al, 2020).

Ethics in AI also support the principles of fairness and justice. Ethical AI ensures that every individual is treated equally and with dignity, regardless of their race, gender, socioeconomic status, or beliefs. Fair AI treats all people equally. It helps prevent discrimination and supports equal opportunity in areas like hiring, justice, and healthcare. In societies where ethical standards are neglected, there is a greater risk of discrimination, corruption, and exploitation of vulnerable populations. By embedding fairness into AI systems, we reduce the likelihood of injustice, foster mutual respect, and contribute to a more harmonious society.

The importance of ethics in AI has become even more critical as AI systems grow in power and influence. Today, AI is used to make decisions in a wide range of applications—from medical diagnoses to criminal justice assessments to personalized advertising. The potential for AI to shape lives on such a broad scale means that developers and users must be acutely aware of the ethical implications of their actions. Without ethical safeguards, AI has the potential to cause significant harm.

For developers, the responsibility to adhere to ethical standards is paramount. They are the architects of AI systems, and their choices directly impact the lives of individuals and communities. Developers must ensure that the data they use to train AI systems is accurate, representative, and free from bias. Biased data can lead to biased AI decisions, perpetuating existing inequalities.

Transparency is also essential—developers should explain how their AI systems work, the risks involved, and the limitations of the technology. An ethical developer recognizes the need for accountability and strives to create systems that are fair, inclusive, and reliable. Developers must think ahead and anticipate the societal consequences of their work, understanding that their innovations will have far-reaching impacts. Ignoring ethics in AI development could result in the creation of tools that inadvertently harm people, even if unintended.

Ethical users must also question AI decisions and report problems, helping improve systems (Morley & al, 2020). As the primary individuals interacting with AI tools, users must recognize the potential risks and limitations of these systems. AI is not infallible, and users must not blindly trust AI decisions. An ethical user understands the implications of using AI, questioning the outcomes and being aware of how the system was trained. Users should avoid employing AI in harmful or unethical ways, such as generating misleading content or infringing on someone's privacy. Ethical users also play a critical role in identifying issues with AI systems—reporting concerns, biases, or failures so that developers can address these problems and improve the technology.

An AI model built on ethical guidelines is more likely to inspire trust and confidence among users. This trust is nurtured through transparency, security, and respect for privacy. When AI systems are built with these ethical principles in mind, they are more likely to gain acceptance and be adopted globally. The long-term success and sustainability of AI technologies depend not only on their technical capabilities but also on their ethical framework. In a world that is becoming increasingly reliant on AI, establishing and adhering to strong ethical standards ensures that these systems contribute positively to society.

IV. Fundamental principles for responsible Ai

As AI systems increasingly shape decisions in healthcare, education, finance, and more, ensuring they are developed and used responsibly has become essential. Responsible AI is guided by a set of core principles that help align technological innovation with societal values. These principles—such as fairness, transparency, and accountability—are not abstract ideals; they are practical tools to ensure that AI benefits everyone and does not reinforce existing inequalities. Below, we explore one of the most crucial principles:

1. **Fairness:** When thinking about fairness in the design and deployment of AI systems, it is important to keep in mind that these technologies, no matter how neutral they may seem, are designed and produced by human beings, who are influenced by the limitations, contexts and biases of their environments, (Leslie, 2019). These factors inevitably find their way into AI systems, whether through the data used to train them, the objectives set by developers, or the

assumptions embedded in their design. As a result Fairness in AI is a complex and multifaceted concept that has been the widely debated in both the academic and industry communities. At its core, fairness refers to the absence of bias or discrimination in AI systems (Barocas & Selbst, 2019) However, it is not merely about avoiding explicit biases but ensuring that AI decisions do not disproportionately disadvantage certain groups or perpetuate societal inequalities. The definition of fairness can vary depending on the context and stakeholders, making it challenging to establish universally applicable standards (Ameyaw, Idemudia, & yelolu, 2024). To achieve fairness, we need more than theoretical definitions—we need action. That includes: using diverse, representative training data, designing algorithms that are explainable and transparent and running regular fairness audits to catch hidden biases. Without such efforts, the risk remains that AI systems may inadvertently harm those they are meant to serve, perpetuating inequalities instead of alleviating them.

2. **Accountability**: refers to a relationship of answerability that establishes essential conditions such as recognition of authority, mechanisms for interrogation, and limitations of power. In the context of AI systems, it involves the expectation that designers, developers, and deployers adhere to established standards and legislation to ensure proper functioning throughout the AI lifecycle cycle (Novelli, Taddeo, & Floridi, 2023) Accountability also requires holding developers and organizations responsible for harm caused by their AI technologies, ensuring they address mistakes and unintended consequences transparently and ethically. (Mensah, 2023) This is crucial to fostering responsible implementation and preventing unethical behavior.

   When AI systems cause harm, mechanisms must be in place to assign responsibility and liability, encouraging proactive risk mitigation and ethical practices. A clear and enforceable accountability framework helps ensure that organizations prioritize safety and consider the broader societal implications of deploying AI.

   A tragic example highlighting the importance of accountability is the 2018 A notable case is the fatal accident involving a self-driving car developed by Uber. The car hit and killed a pedestrian while operating in autonomous mode. Investigations revealed that there were flaws in the technology's design and inadequate safety measures. This shows why strong accountability is needed to ensure AI systems are tested and safe before use. so fostering trust and preventing harm. by establishing clear standards, liability mechanisms, and transparent practices, accountability promotes ethical AI deployment while safeguarding societal interests. (Mensah, 2023).

3. **Transparancy :** Ensuring transparency in AI development is essential for building trust, promoting accountability, and addressing ethical challenges. Explainable AI (XAI) aims to enhance the interpretability and understandability of AI models, especially in complex systems where real-time decision-making requires clear justification**.** Unlike traditional "black-box" models, explainable AI provides clear explanations of how AI systems arrive at specific decisions, making it easier to understand their reasoning.

   These approaches contribute to the developement of responsible and accountable AI systems, ensuring that their creation and deployment align with ethical principles and societal values. (Olatunji Akinrinola et al, 2024) For example, In healthcare, if an AI recommends a treatment, XAI can explain the reason, helping doctors understand and ensure it matches medical ethics and supports patient care.

4. **Sustainibility**: As mentioned earlier, the increased use of Artificial intelligence systems (AI systems) is associated with multifaceted social, environmental, and economic consequences. These include nontransparent decision-making processes, discrimination, increasing inequalities, but also rising energy consumption and greenhouse gas emissions in AI model development and application,. (Friederike Rohde et al, 2024). Also, the spread of AI has led to a few powerful companies controlling the tools and infrastructure for advanced AI, which increases the economic gap. It is recommended that policymakers address concerns about AI's energy consumption by taking the following steps:

- **Develop energy transparency standards for AI models** to ensure their energy consumption is clearly understood and monitored.
- **Encourage voluntary commitments on energy transparency** from organizations developing foundation models.
- **Evaluate the unintended consequences of AI regulations** to prevent policies that inadvertently increase energy use.
- **Leverage AI technologies to decarbonize government operations**, promoting sustainability in public sector practices.

5. **Privacy:** The European Union's General Data Protection Regulation (GDPR) has established guidelines on transparency by requiring organizations to provide individuals with clear information about how their data will be processed using automated decision-making systems like AI algorithms. This regulation aims to ensure that individuals understand how decisions affecting them are made, reducing opacity surrounding AI processes. (Mensah, 2023). However, while GDPR represents significant progress in protecting privacy, it illustrates a broader issue: regulations have not kept pace with the rapid advancements in AI technology

(Talati, 2024). As AI capabilities evolve at an unprecedented rate, legal frameworks often struggle to address the complex, nuanced challenges these systems pose. This regulatory lag creates a "gap" or "anarchy" in governance, as described by Talati, where existing rules are either insufficient or outdated. This vacuum enables influential players in the tech industry to exploit the lack of oversight. These actions can occur:

- "In Full View": When companies use loopholes in existing laws to justify practices that may not align with ethical standards.
- "Out of Sight": When organizations discreetly bypass accountability, often leveraging the technical complexity of AI systems to evade scrutiny.

These two tactics show why we urgently need new and stronger AI laws. Lawmakers must quickly close this gap by making rules that protect privacy and guide ethical AI use. Without these measures, the potential for privacy violations and unchecked power within the tech industry will escalate, undermining public trust and societal values.

V. Conclusion:

The integration of AI into daily life continues to reveal both the potential and pitfalls of this transformative technology. While AI systems hold promise in addressing complex problems, their effectiveness is undermined when they replicate societal biases and perpetuate systemic inequalities. the ethical challenges posed by biased algorithms in different sectors like healthcare, criminal justice, recruitment, and the environment, increase the need for transparency, accountability, fairness, and sustainability in AI development.

By examining these issues, we recognize that ethical AI is not solely a technical challenge but a societal one. It requires collaboration across disciplines and a commitment from developers, policymakers, and the public to implement solutions that mitigate bias and ensure equity. To foster a fair AI ecosystem, future work must integrate ethical frameworks from the outset, prioritizing inclusivity and environmental sustainability. Only through such proactive measures can AI serve as a true force for good, bridging rather than widening societal divides.

VI. Bibliographie

Ameyaw, M., Idemudia, C., & yelolu, T. (2024, jun). Financial compliance as a pillar of corporate integrity: A thorough analysis of fraud prevention. *Finance & Accounting Research Journal,, 6*(7), pp. 1157-1177.

Barocas, S., & Selbst, A. (2019, nov). Big data's disparate impact. California Law Review. *California Law Review, 104*(3), p. 671. doi:/doi.org/10.15779/Z38BG31

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in. *Conference on Fairness, Accountability, and Transparency (FAT*)*, (pp. 1-15).

Castro, D. (2024, jan 29). Rethinking Concerns About AI's Energy Use. *center for data innovation* .

Ferrara, E. (2023). FAIRNESS AND BIAS IN ARTIFICIAL INTELLIGENCE: A BRIEF SURVEY OF SOURCES, IMPACTS, AND MITIGATION strategies.

Friederike Rohde et al. (2024). Broadening the perspective for sustainable artificial intelligence: sustainability criteria and indicators for Artificial Intelligence systems. *sustainability criteria and indicators for Artificial Intelligence systems,, 66*. doi:doi.org/10.1016/j.cosust.2023.101411.

Graham Attwell et al. (2021, dec). *Artificial Intelligence & Vocational Education and Training: How to Shape the.* Consulté le 11 09, 2024, sur https://creativecommons.org/licenses/by-sa/4.0/

Huriye, A. (2023). The Ethics of Artificial Intelligence: Examining the Ethical Considerations Surrounding the Development and Use of AI. *American Journal of Technology, 2*, pp. 37-44. Récupéré sur https://doi.org/10.58425/ajt.v2i1.142

Leslie, D. (2019). *Understanding artificial intelligence, ethics and safety.* The Alan Turing Institute.

Mensah, G. B. (2023, nov). Artificial Intelligence and Ethics: A Comprehensive Review of Bias Mitigation, transparancy and accountability in AI systems.

Min, A. (2023, oct). Artifical Intelligence and Bias: Challenges, Implications, and Remedies. *Journal Of Social Research, 2*(11), pp. 3808-3817. doi:10.55324/josr.v2i11.1477

Novelli, C., Taddeo, M., & Floridi, L. (2023, fab). Accountability in artificial intelligence: what it is and how it works. *AI & SOCIETY, 39*(4). doi:10.1007/s00146-023-01635-y

Olatunji Akinrinola et al. (2024, mar). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews, 18*(3). doi:10.30574/gscarr.2024.18.3.0088

Simundic, A.-M. (2013, fev). Bias in research. *Biochemia*. doi:DOI: 10.11613/BM.2013.003

Smith, J., & Noble, H. (2014). Bias in research. Evidence-Based Nursing. *(4)*(17), pp. 100-101. Récupéré sur https://doi.org/10.1136/eb-2014-101946

Suresh , H., & Guttag, J. (2019, jan). A Framework for Understanding Unintended Consequences. *Association for the Advancement of Artificial*.

Talati, D. (2024, mar). Ethics of AI (Artificial Intelligence). Récupéré sur https://d197for5662m48.cloudfront.net/documents/publicationstatus/196997/preprint_pdf/492f93c4664c210596fa39b88e96865e.pdf

Wood, M. (2021, mar 22). Bias in facial recognition isn't hard to discover, but it's hard to get rid of. Consulté le 11 9, 2024, sur https://www.marketplace.org/shows/marketplace-tech/bias-in-facial-recognition-isnt-hard-to-discover-but-its-hard-to-get-rid-of/