University of Puerto Rico
Mayagüez Campus
Mayagüez, Puerto Rico

Department of Electrical and Computer Engineering

AI and Society: Ethical Impact Analysis

by:
Manuel Alejandro Umaña Rodriguez
Dahyna Gabrielle Martínez Pérez
Jan Luis Pérez de Jesús
Christopher Hans Mayens Matías

For: Profesor José Fernando Vega Riveros
Course: ICOM 5015, Section 001D
Date: May 16, 2025

Introduction:

Artificial Intelligence has rapidly changed the current state of our society, reshaping how people interact with technology as a whole. What started as a simple reasoning tool for simple goal based objectives has become a useful tool in medical diagnostics and autonomous systems. This huge lap in innovation has made incredible advancements, but has brought some ethical concerns, particularly regarding fairness, misuse, privacy, and accountability. This report represents an ethical analysis on different applications of AI and how they impact society in three major philosophical frameworks: consequentialism, deontological ethics, and virtue ethics. The work was divided among the group members to analyze different areas of the subject. Specifically Jan analyzed potential misuse, Dahyna focused on positive impacts in healthcare and education, Chris addressed privacy and surveillance, and Manuel covered employment and labor.

Overview:

Society has been heavily impacted by Artificial Intelligence in nearly every aspect of our daily lives. While this technology has brought significant advancements, it also has raised complex ethical questions. This report will feature four key areas of AI's societal impact to provide a fairly balanced perspective on its benefits and challenges. The studied areas include:

- **Potential for Misuse**, specifically in the realm of misinformation which threatens individual integrity.
- **Positive Impacts**, especially in healthcare and education, where AI serves as a tool for diagnostics and supports independent learning.
- **Effects on privacy and surveillance**, where AI systems can enhance security and erode civil liberties if left unchecked.
- **Impacts on employment and labor,** as automation and AI tools to reshape job markets.

Potential for Misuse

AI technology has become increasingly accessible in recent years. This revolution has taken the world by storm and altered our lives in fascinating and meaningful ways. However, not everyone has been using it for beneficial purposes. Concerns about the potential misuse of AI have become more relevant, especially in areas such as content personalization and media manipulation. Deepfake is a type of artificial intelligence used to create convincing fake images, videos, and audio recordings [1]. Often they can transform existing content in

which a person is replaced with another or fabricate entirely new scenarios in which individuals say or do things they never did. This represents a major threat given that they blur the line between reality and fiction, posing threats in privacy, consent, and reputational integrity. The technology has become so prevalent that prominent institutions like MIT have created tools to provide guidance on identifying deepfakes [2]. One example is DetectFakes, a website that challenges users to identify if an image is artificial or real [3]. From a consequentialist perspective, the harm caused by misinformation and the loss of public trust often outweighs the benefits deepfakes might bring. A deontologist would make the argument that deepfakes are unethical regardless of outcome because they violate the individual's right to autonomy and truthful representation. Virtue ethics would criticize the intentions of those creating harmful deepfakes asking whether such actions reflect the integrity and responsibility of the innovation [4]. Authorities across the world have begun leveraging similar technologies for surveillance, manipulation and even political propaganda. The University of Cambridge's report "The Malicious Use of Artificial Intelligence", describes how AI technologies, including deepfakes, can be exploited for criminal purposes such as political manipulation and social disruption [5]. Additionally, the Cloud Security Alliance highlights the severe emotional and reputational harm deepfakes can inflict on individuals emphasizing the need for comprehensive security measures [6]. There is great urgency for ethical guidelines to prevent misuse while promoting accountability in AI development.

Positive Impacts

With the aforementioned rise of Artificial Intelligence's popularity, most consumers are familiar with its applications in generative media. While many associate AI with fears of job replacement or academic dishonesty, fewer recognize the substantial benefits it brings, particularly in health care and education. In healthcare, AI is revolutionizing medical diagnostics, which is the process of evaluating medical conditions or diseases by analyzing symptoms, medical history, and test results [7]. In radiology, for example, the Institute of Radiology conducted a study using Aidoc, an FDA approved AI operating system algorithm, alongside two radiologists to identify incidental pulmonary embolisms (iPE) [8]. While the radiologists alone missed 50% of the iPE in the first batch of CT scans, the error rate dropped to 7.1% when the AI tool was implemented [8]. Even with these types of results, the World Health Organization cautions that AI should assist, not replace, the final clinical decisions of doctors as to avoid bias and preserve patient trust [9]. In regards to education, the general public has shown concern in AI potential as a means to fuel academic dishonesty. As

such a common method among educators has been to forbid the use of AI outright. This restricts learning opportunities for neurodiverse students who benefit from these adaptive tools [10]. In contrast, Institutes such as Harvard have leveraged AI by incorporating AI-powered chatboxes, implemented as a virtual rubber ducky [11]. It is meant to aid programming students without providing direct answers.The goal was to approximate 1:1 teacher-to-student ratio as the bot provides code clarification, syntax style improvements, and accurate adherence to curricular and administrative queries. In a Fall 2023 trial with approximately 500 students involved 47% of students rated the tool as “very helpful” as compared to the 6% that deemed it “unhelpful” [11]. Most students felt that the bot acted as a personal tutor that would listen to their concerns and explain in a non-judgmental manner. From a consequentialist perspective, these applications are ethically justified due to their clear benefits as earlier diagnosis and more inclusive education, despite concerns about misuse since legal requirements are being continuously updated. A deontological analysis highlights the importance of using AI in ways that respect individual rights, such as to uphold consent, privacy, and fairness by providing these tools to as many people as possible, but with still upholding regulation [4]. Meanwhile, virtue ethics would encourage the reflection of the developers and institutes intentions, asking whether these tools are designed and used with responsibility and compassion for the human’s well being. Taken together, these perspectives support the ethical use of AI when implemented thoroughly and transparently.

Effects on privacy, surveillance, and civil liberties

Artificial Intelligence (AI) is transforming how societies collect, process, and act on information, particularly in the realm of surveillance. It is reshaping security practices across public and private sectors, offering advanced tools to monitor, detect, and respond to threats. However, this growing reliance on AI-driven surveillance raises serious concerns about its impact on privacy, individual freedoms, and civil liberties. While AI can enhance security and efficiency, it often comes at a significant cost. Technologies like facial recognition allow quick identification, but collect large amounts of personal data without consent, violating fundamental rights. These systems also depend on the data they are trained with, which can reflect and amplify existing societal biases. As a result, tools meant to prevent crime may end up targeting specific groups unfairly, leading to discriminatory practices based on race, ethnicity, or socioeconomic status. Moreover, storing large volumes of personal data increases the risk of cyberattacks, putting sensitive information and basic freedoms at serious risk.

Ahead of the 2024 Summer Olympics, Paris deployed an AI-powered surveillance system to monitor potential threats using thousands of CCTV cameras. While intended to enhance security, it raised ethical and legal concerns over abuse of power, lack of transparency, and the risk of setting a dangerous precedent [12]. The system also faced cybersecurity risks, especially given the size of the crowds. Similarly, Clearview AI, a startup that launched an app in 2017, built a database of over 30 billion photos scraped from social media without consent, leading to fines in Europe and Australia for major privacy violations. As tech reporter Kashmir Hill noted, "Clearview made an ethical breakthrough, not a technological one" [13]. These cases reveal how unregulated AI surveillance can do more harm than good. If adopted, it must be governed by strict ethical guidelines and legal safeguards. To prevent further erosion of democracy, the international community must set clear limits and controls on these powerful tools [14].

The ethical concerns surrounding AI surveillance are profound and span across multiple frameworks. While it promises safety and efficiency, it often undermines justice, privacy, and democratic freedoms. From a consequentialist perspective, the harms caused by AI surveillance, such as discrimination, data misuse, and the erosion of public trust, may outweigh any intended benefits to security. At the same time, the violation of individual privacy raises serious deontological concerns because it ignores moral duties and disregards basic rights, even when those actions are justified by efficiency. This also reflects a failure in moral character and responsibility from those implementing the technology, falling short of the virtues that should guide institutions serving the public. Until there is a secure and transparent way to use AI without compromising rights or reinforcing harmful biases, this technology should not be widely adopted.

Impact on employment and labor markets

With the rise in artificial intelligence, new and improved technologies have been more accessible. However, due to the meteoric rise there have been some unintended consequences that have been drawn forward. These consequences are already problems that society was facing with some jobs becoming obsolete due to the improvements of technology such as cashiers and fast food workers to name a few examples [15]. In a study conducted by the McKinsey Global institute it speaks about the differences that AI will bring and is already bringing, saying that by 2030 up to 800 million jobs could be displaced by automation, but new jobs will also be created. With the biggest demand for new jobs being centered around STEM and healthcare, while the biggest worries for job loss are in manufacturing and retail.

As such inequality may worsen if transitions are not supported with social policies [16]. Another article by Forbes states that while some 300 million jobs will be lost, according to Goldman Sachs, some of the leading investors and experts have stated in the article that AI will not be replacing human workers but will be transforming to incorporate AI tools and requiring workers to adapt to the new tools [17]. Over the last couple of years with the introduction of Large language models such as ChatGPT,Gemini, or other models there have been looks into the potential future this technologies will have on the workforce where 80% of U.S. jobs will be affected, with 19% of workers seeing 50% or more of their tasks exposed to LLMs. The authors note that potential for augmented productivity is very possible and likely, but that there are real displacement risks if the models substitute rather than support workers[18]. To conclude, the integration of artificial intelligence into the workforce is not a question of *if*, but *how*, a natural consequence of ongoing technological innovation. As history shows, technological progress often displaces certain roles while creating new ones, but the ethical implications of how this transition is managed are critical. From a consequentialist standpoint, the benefits of increased productivity and the creation of new, more complex roles must be carefully weighed against the societal harms of unemployment and inequality. Deontological ethics demands that organizations uphold their moral duties protecting workers' rights, ensuring transparency, and resisting the impulse to treat employees as expendable means to profit. Meanwhile, virtue ethics reminds us that the character of those leading the AI transition matters deeply. Ethical leadership requires empathy, foresight, and a commitment to fairness. Thus, while AI may be an inevitable tool of progress, its deployment must be guided by the intentional choice to enhance human potential rather than replace it. The future of work depends not solely on the capabilities of machines, but on the moral compass of those who design and implement them.

Deliberation Process

During our collaborative exploration of AI's societal impact, our group discussed a wide range of topics including deepfakes, surveillance, healthcare applications, education, and labor market shifts, as shown in greater detail in Appendix A. Despite approaching these topics from different angles, a shared understanding emerged: there is an urgent need for ethical oversight in AI development and deployment. Whether discussing Jan's concerns over misinformation and media manipulation, Chris's emphasis on civil liberties and surveillance, Dahyna's exploration of benefits in healthcare and education, or Manuel's focus on employment, all members acknowledged that AI's transformative potential must be balanced

against its capacity for harm. We collectively agreed that AI can improve human life, through earlier medical diagnoses, inclusive education, and productivity enhancements, but only if it is used responsibly. There was also a consistent concern across the board about risks such as job displacement, bias, surveillance abuse, and misinformation. Jan and Chris emphasized the dangers of misuse and the importance of accountability and consent, while Dahyna and Manuel acknowledged that with proper regulation, AI can be a force for good. Although levels of optimism varied among members from more hopeful views on AI's benefits to cautionary perspectives on ethical pitfalls there was consensus that the future of AI must be shaped by values like fairness, transparency, and compassion. Ultimately, we concluded that AI should not replace human judgment, but augment it, and that strong ethical frameworks are essential to ensure its development serves society rather than undermines it.

Conclusion:

Artificial Intelligence holds immense potential to improve human life, from enhancing medical diagnostics and personalized learning to streamlining labor and boosting productivity. However, its rapid adoption also brings serious ethical challenges, particularly in surveillance, misinformation, and job displacement. The same tools that offer innovation can also undermine privacy, autonomy, and equality if left unchecked. As AI continues to shape modern society, it is critical that its development and use are guided by robust ethical standards, legal safeguards, and transparent governance. The future of AI must prioritize not just what is possible but what is responsible, fair, and aligned with human dignity. If implemented without oversight, AI could amplify inequality, weaken trust in public institutions, and compromise democratic values. But if guided ethically, it has the power to support a more just, inclusive, and empowered society. As such the group found that Deontological Ethics should be the leading principle for the creation of further AI tools and innovations as it protects the rights of types of groups and individuals instead of a single entity. Moreover it supports the creation of a future where restrictions in technology will serve as a shield to protect the integrity of individuals and promote a healthy integration of AI tools and users.

Independent Evaluation

Jan-

Prior to developing this report, I hadn't fully grasped how deeply AI can affect society beyond its technical capabilities. Through my research on generative AI, I explored its ethical implications in misinformation and impersonation especially in the use of deepfake. I learned

that it's not just the creation of a photo but can extend to video and even voices which represents an almost dystopian threat. After reviewing my classmates' findings, I came to understand that technological progress has to be weighed against its potential social impact.

Manuel-

Before developing this report I had some understanding of how AI was impacting society. However after deliberation with my classmates I have gotten a better understanding of the real impact AI has and will have on society. Throughout my research I have explored the impact on employment and labor markets and by increasing my scope of knowledge. It has become apparent that ethical oversight is necessary so that societal impacts such as unemployment and inequality do not increase as a result of this transition. Being carefully monitored so that new emerging technology is not used maliciously.

Chris-

Working on the effects of AI on privacy, surveillance, and civil liberties helped me better understand how powerful and potentially dangerous these technologies can be if not used under proper regulation. I learned how tools like facial recognition and large-scale data collection can violate ethical principles and compromise individual rights. Reading my classmates' work also showed me how their topics, such as misinformation in the media and its impact on jobs, connect with mine because they all raise similar concerns about the ethical risks of using AI without proper oversight.

Dahyna-

Working on this report has taught me that AI Ethics is not a simple question of right or wrong, it depends on how and why the technology is used. My primary focus was on the positive impacts of AI, particularly demonstrating how AI can aid in inclusive education, one-on-one tutoring, and medical diagnostics. However, I also learned from my companions research that even protective systems like surveillance can be misused raising critical ethical concerns about privacy. The project allowed me to see how different people approach ethics in general, and I found myself most aligned with the virtue and deontological frameworks, which emphasized intention, fairness, and respect for human dignity; even if it's not the majority.

References:

[1] K.Yasar, N.Barney, and I.Wigmore, “What is deepfake technology?,” *techtarget.com* , Aug. 2024. [Online]. Available: https://www.techtarget.com/whatis/definition/deepfake . [Accessed May. 11,2025].

[2] MIT Media Lab, “DetectFakes,” *MIT Media Lab*, 2024. [Online]. Available: https://detectfakes.media.mit.edu/ . [Accessed May. 11,2025].

[3] MIT Media Lab, “Project Overview: DetectFakes,” *media.mit.edu*, 2025. [Online]. Available: https://www.media.mit.edu/projects/detect-fakes/overview/ . [Accessed May. 11, 2025].

[4] J. F. Vega Riveros, “The Three Main Philosophical Schools of Ethics”, *University of Puerto Rico at Mayagüez*, Course Handout for ICOM 5015, May. 2025.

[5] S. Roberts, “The Malicious Use of Artificial Intelligence,” *cam.ac.uk*, 2018. [Online]. Available: https://www.cam.ac.uk/stories/malicious-ai-report . [Accessed May 11,2025].

[6] Cloud Security Alliance, “AI Deepfake Security Concerns,” *cloudsecurityalliance.org*, Jun. 25,2024. [Online]. Available: https://cloudsecurityalliance.org/blog/2024/06/25/ai-deepfake-security-concerns . [Accessed May 11,2025].

[7] M. Al-Antari, “Artificial Intelligence for Medical Diagnostics—Existing and Future AI Technology! ,” *pmc.ncbi.nlm.gov*, Feb. 12, 2023. [Online]. Available: Artificial Intelligence for Medical Diagnostics—Existing and Future AI Technology! - PMC .[Accessed May 12, 2025].

[8] V. Josephin et al., “Implementation of an AI Algorithm in Clinical Practice to Reduce Missed Incidental Pulmonary Embolisms on Chest CT and Its Impact on Short-Term Survival," *pubmed.ncbi.nlm.nih.gov*, Oct. 9,2024. [Online]. Available:Implementation of an AI Algorithm in Clinical Practice to Reduce Missed Incidental Pulmonary Embolisms on Chest CT and Its Impact on Short-Term Survival - PubMed .[Accessed May 12, 2025].

[9] World Health Organization, “Ethics and governance of artificial intelligence for health,” *www.who.int*, Mar. 25, 2025. [Online]. Available:https://iris.who.int/bitstream/handle/10665/375579/9789240084759-eng.pdf?sequence=1 .[Accessed May 12, 2025].

[10] M.Cardona, R.Rodríguez, and K.Ishmael, “Artificial Intelligence and the Future of Teaching and Learning,” *www.ed.gov*, May 2023. [Online]. Available:https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf .[Accessed May 12, 2025].

[11] R.Liu et al., “Teaching CS50 with AI,” *cs.harvard.edu*, Mar. 23, 2024. [Online]. Available:https://cs.harvard.edu/malan/publications/V1fp0567-liu.pdf .[Accessed May 12, 2025].

[12] My Privacy Blog, “AI and Surveillance: Balancing Security and Civil Liberties,” *myprivacy.blog*, Aug 27, 2024. [Online].Available:AI and Surveillance: Balancing Security and Civil Liberties .[Accessed May 14, 2025].

[13] L. Mineo, “How facial-recognition app poses threat to privacy, civil liberties,” *harvards.edu*, Oct. 26, 2023. [Online]. Available:How facial-recognition app poses threat to privacy, civil liberties — Harvard Gazette .[Accessed May 14, 2025].

[14] A. Olvera, “How AI surveillance threatens democracy everywhere,” *thebulletin.org*, Jun. 7, 2024. [Online]. Available: How AI surveillance threatens democracy everywhere - Bulletin of the Atomic Scientists .[Accessed May 14, 2025].

[15] Resumeble, “ The Power of Automation and AI: Jobs That Will Disappear by 2030,” *resumeble.com*, Sep. 2, 2024. [Online]. Available:https://www.resumeble.com/career-advice/jobs-that-will-be-gone-2030 .[Accessed May 13, 2025].

[16] J.Manyika et al., “ Job lost, jobs gained: What the future of work will mean for jobs, skills and wages ,” *mckinsey.com*, Nov. 28, 2017. [Online].

Available:https://www.mckinsey.com/featured-insights/future-of-work/jobs-lost-jobs-gained-what-the-future-of-work-will-mean-for-jobs-skills-and-wages .[Accessed May 13, 2025].

[17] M.T. Bannon, " How AI Is Changing The Future Of Work," *forbes.com*, Jun. 22, 2023. [Online]. Available:https://www.forbes.com/sites/marenbannon/2023/06/22/how-ai-is-changing-the-future-of-work/ .[Accessed May 13, 2025].

[18] T.Eloundou, S.Manning, P.Mishkin, and D.Rock, "GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models," *arxiv.org*, Aug. 21, 2023. [Online]. Available:https://arxiv.org/abs/2303.10130 .[Accessed May 13, 2025].

Appendix A:

Initial Minute Reunion

Location:	Google Meet
Date:	May 3, 2025
Time:	10:00 AM – 12:32 PM

Discussions:

1. The team gathered to review the entirety of the report guidelines/instructions.
2. Read and discuss the three main philosophical schools of ethics (consequentialism, deontological ethics, and virtue ethics) from the file provided by the Professor.
3. Each member presented their opinions of what they believed were the most relevant societal impacts of AI.
4. Each member chose an impact that reflected both positive and negative aspects of AI as to avoid bias.

Decisions:

1. Created the initial report draft that included a section for each deliverable.
2. Manuel Alejandro Umaña Rodriguez chose to research the Impacts on Employment and Labor Markets.
3. Dahyna G. Martínez Pérez chose to research the Positive impacts of AI with a focus on Medical Diagnostics and Education.
4. Jan Luis Pérez de Jesús chose to research about the Impacts of potential misuse with a focus on Deepfakes.
5. Christopher Hans Mayens Matías chose to research the Effects on Privacy, Surveillance, and Civil Liberties.

Pending actions:

Action	Responsable	Completion Date
Create the Written Report draft document	Jan Luis Pérez de Jesús	May 4,2025
Each member must research their selected topic and bring a cohesive summary of their findings with evidence.	All team members	May 7, 2025

Second Minute Reunion

Location: Google Meet

Date: May 7, 2025

Time: 7:30 PM – 8:45 PM

Discussions:

1. Beside the main ethical analysis that each member had been assigned at the last meeting, sections of the written report were distributed within this meeting.

2. Each member was to discuss their findings including their ethical analysis.

3. Jan and Dahyna's points were conflicting, as expected, by showing both the positive societal growths that AI has created within healthcare and education, but also some severe privacy issues that involve AI in regards to deepfakes.

4. Manuel's point demonstrated that the most jobs replaced by AI would incur the biggest demand for new jobs being centered around STEM and healthcare, while the greatest job loss would be in manufacturing and retail. His main point was not when

AI was going to replace people in labor intensive jobs, but how these need to be regulated as to affect the least amount of people possible.

5. Christopher's topic complemented part of Jan's topic as these pertain to the evaluation of AI and people's Personal Identifiable Information. His points also brought forth the undeniable fact that if AI is in charge of surveillance and collecting information, what happens when the model has inherent developer bias?

6. In his findings Christopher also presented how some models use a database of over billion photos scraped without consent, leading to major privacy violations. This information shed some light towards the team as most members thought that AI in surveillance would result in mainly positive impacts such as higher detection rates, longer monitoring life, and more precious observations.

Decisions:

1. Manuel Alejandro Umaña Rodriguez will work on the deliberation process of the report, independent evaluation, and his assigned ethical research and analysis.

2. Dahyna G. Martínez Pérez will work on the appendix, independent evaluation, modifying the reference page, and her assigned ethical research and analysis.

3. Jan Luis Pérez de Jesús will work on the introduction, overview, independent evaluation, and his assigned ethical research and analysis.

4. Christopher Hans Mayens Matías will work on the conclusion, independent evaluation, and his assigned ethical research and analysis.

5. To keep coherency within the report, every person is required to read the discussions of the previous member. If there are any concerns, separate meetings or calls should be organized.

Pending actions:

Action	Responsable	Completion Date

Write a draft and outline of each section of the report.	All team members	May 14, 2025
Organize and modify the reference pages.	Maunel Alejandro Umaña Rodriguez and Dahyna G. Martínez Pérez	May 15, 2025

Final Minute Reunion

Location: Google Meet

Date: May 14, 2025

Time: 5:00 PM – 5:40 PM

Discussions:

1. Once again all members were to explain the ethical findings of their most recent report writings.These were combined with the previous established report outline structure. This combination was refined within the report.

2. All members took a moment to read the entire report to verify if all deliverables were completed.

3. The ethical findings within the second round of deliberation proved not to be as different as the first round, this led to a quicker meeting as previous ethical differences were already discussed.

Decisions:

1. Each member is to read the entire report within the established deadline and notify any mistakes or inconsistencies.
2. Members with any conflicting analysis are required to discuss their differences and include their perspectives within the report as part of the deliberation processes.

Pending actions:

Action	Responsable	Completion Date
Finalize each section of the report.	All team members	May 16, 2025
Verification and Validation of the entire report.	All team members	May 16, 2025
Submission of the final report.	All team members	May 16, 2025