DOI: https://doi.org/10.12700/btsz.2025.7.1.SI.87

The corruption risks of artificial intelligence, what to consider in the golden age of artificial intelligence?

A mesterséges intelligencia korrupciós kockázatai, avagy mire érdemes figyelni az AI aranykorában?

SOÓS Georgina[1]

Abstract

The study examines public perception of artificial intelligence (AI), which presents a mixture of caution and optimism. Many view AI as a valuable tool for automation, boosting efficiency and simplifying tasks. However, concerns persist about its potential impact on employment and the risks of misuse. The research identifies three key areas that need improvement: accuracy, ethical considerations, and the integration of human judgment. Enhancing accuracy is critical, requiring ongoing development of data and algorithms to ensure AI systems are more reliable and minimize errors and so are ethical considerations, urging the adoption of robust ethical frameworks to guide the development and deployment of AI technologies. Additionally, the integration of human judgment is seen as vital for the effective use of AI, as it facilitates collaboration between humans and AI, ensuring that human oversight is maintained, especially in critical decision-making processes. The study underscores that to fully realize AI's potential, continuous development and refinement are essential.

Keywords

Artificial intelligence, corruption, algorithms, data manipulation, transparency, accountability

Absztrakt

A tanulmány a mesterséges intelligencia (AI) közvéleményről alkotott képet elemzi, amely vegyes, de óvatos optimizmus jellemzi. Sokan értékes automatizálási eszközként tekintenek rá, amely növeli a hatékonyságot és racionalizálja a feladatokat, ugyanakkor aggályok is felmerülnek a munkahelyek megszűnésével és a visszaélések lehetőségével kapcsolatban. A kutatás három fő területet emel ki, ahol fejlesztések szükségesek: pontosság, etikai megfontolások és az emberi ítélőképesség integrálása. A pontosság javítása érdekében a mesterséges intelligencia megbízhatóságát növelő adat- és algoritmusfejlesztés szükséges. Az etikai kérdések prioritást élveznek, szigorú etikai irányelvek bevezetését igényelve az AI fejlesztésében. Az emberi ítélőképesség bevonása segíthet a mesterséges intelligencia és az emberek közötti hatékony együttműködésben, biztosítva az emberi felügyeletet a kritikus döntések során. A tanulmány hangsúlyozza, hogy az AI hatékony alkalmazásához folyamatos fejlesztés szükséges.

Kulcsszavak

Mesterséges intelligencia, korrupció, algoritmusok, adatmanipuláció, átláthatóság, elszámoltathatóság

[1] soosgeorgina01@gmail.com I ORCID: 0009-0007-7282-7591 I University student, Cybersecurity Masters, Faculty of Public Governance and International Studies, Ludovika University of Public Service | Egyetemi hallgató, Kiberbiztonsági mesterképzési szak, Nemzeti Közszolgálati Egyetem Államtudományi és Nemzetközi Tanulmányok Kar

INTRODUCTION

Artificial Intelligence (AI) has the potential to revolutionize industries and improve lives through data analysis and automation. However, it also poses a significant corruption risk due to its reliance on vast datasets, which can be manipulated through data poisoning or algorithmic bias. The complexity of AI systems also makes their decision-making processes opaque, creating potential loopholes for exploitation. To combat this, a proactive approach is necessary, including Explainable AI (XAI), human oversight, ethical frameworks, public awareness, a culture of integrity, public discourse, and education, and building a trustworthy AI future.

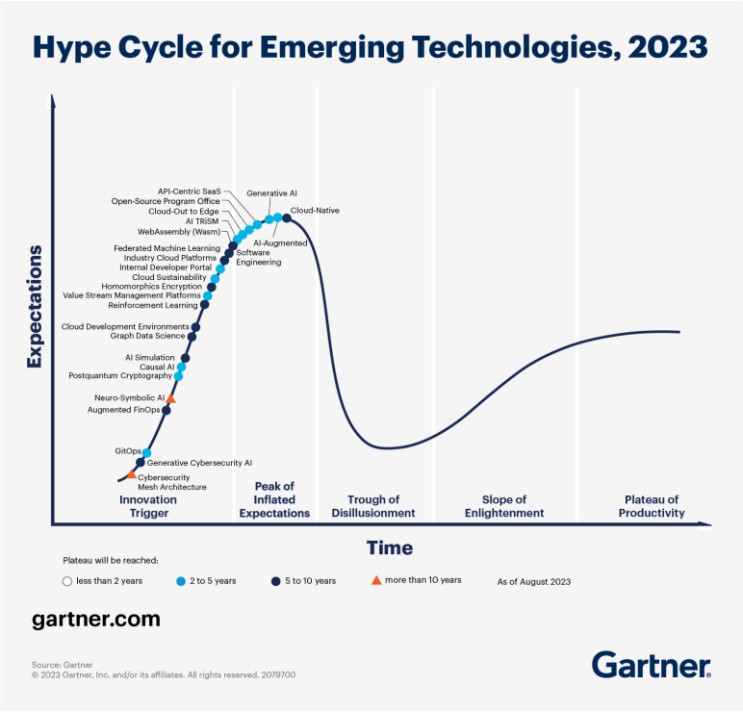
Explainable AI can demystify decision-making processes, while human-in-the-loop design ensures human involvement in critical decisions. Open discussions involving experts, educators, and policymakers can lead to robust policies and regulations that mitigate AI-driven corruption. Public education empowers individuals to identify and report potential AI misuse, fostering ethical practices. By implementing these proactive measures, we can harness the transformative potential of AI while safeguarding against its vulnerabilities.

Further research is needed to explore the specific mechanisms of AI manipulation for corruption and develop targeted mitigation strategies. Legal and regulatory implications of AI in the context of corruption should also be considered, with governments and international organizations developing clear laws and regulations that govern the use of AI in the public sector and private enterprises.

GARTNER'S 2024 MEGATRENDS

Gartner's 2024 Megatrends outlines ten key emerging technologies that are expected to significantly influence business and technology decisions over the next three years. These trends include AI Trust, Risk & Security Management (AI TRiSM), Continuous Threat Exposure Management (CTEM), Sustainable Technologies, Platform Engineering, AI-Augmented Development, Industry Cloud Platforms, Intelligent Applications, Democratized Generative AI, Augmented Connected Workforce, and Machine Customers.

AI TRiSM focuses on building trust by mitigating risks and ensuring ethical considerations throughout the AI lifecycle. CTEM goes beyond reactive measures by identifying and managing potential threats before they exploit vulnerabilities. Sustainable Technologies offer a win-win proposition, enhancing business efficiency and minimizing environmental impact. Platform Engineering empowers developers by providing pre-built tools, services, and APIs, streamlining development processes, and reducing time-to-market. AI-Augmented Development automates repetitive tasks, generates code, and suggests best practices, allowing developers to focus on creative problem-solving and strategic thinking. Industry Cloud Platforms offer pre-built, industry-specific solutions, often including security measures and compliance features.



1. Figure: Ava McCartney, 'Gartner Top 10 Strategic Technology Trends 2024'. Accessed: Apr. 07, 2024. 23:14:12. Available: https://www.gartner.com/en/articles/gartner-top-10-strategic-technology-trends-for-2024

TYPES OF THREATS POSED BY AI IN CORRUPTION

1. Automation of Corrupt Practices

The automation of tasks previously handled manually, such as contract review, payment approvals, and procurement process management, can inadvertently facilitate corrupt practices. AI algorithms can manipulate bidding processes without human intervention, inflate invoices, or divert funds to personal accounts. This automation can make it easier for corrupt actors to operate undetected.

2. Data Manipulation and Fraud

AI's ability to analyze and process vast amounts of data makes it a powerful tool for manipulating and concealing corrupt activities. AI-based systems can generate fake documents, modify transaction records, or create false identities to mask illicit transactions. This capability poses a significant threat to the integrity of financial and administrative systems.

3. Targeted Bribery and Influence Peddling

AI can be used to identify and target individuals susceptible to bribery and influence peddling, as well as to facilitate influence peddling activities. AI algorithms can analyze social media profiles, professional networks, and personal interests to identify individuals

who may be vulnerable to financial incentives or personal favors. This capability can be exploited to corrupt decision-making processes and undermine ethical behavior.

4. Weaponizing AI for Surveillance and Control

AI can be used to develop sophisticated surveillance systems that track individuals' activities, movements, and communications. This surveillance can be used to intimidate and silence whistleblowers and hinder anti-corruption investigations. Such surveillance can create an atmosphere of fear and intimidation, stifling efforts to expose and address corruption.

5. Lack of Transparency and Accountability

The complexity and opacity of AI systems can make it difficult to track and scrutinize their decision-making processes. The lack of transparency can hinder accountability and make it easier for corrupt actors to exploit AI for their own gain. Without clear transparency mechanisms, AI systems can become black boxes, allowing corruption to flourish undetected.

To mitigate the pervasive threats of AI-based corruption, it is essential to establish comprehensive safeguards and ethical frameworks governing the development and deployment of AI. These safeguards should include the following key elements:

- Embracing transparency and verifiable processes: AI systems should be inherently transparent and verifiable, allowing for open scrutiny of the underlying algorithms and data sources. This openness facilitates independent audits and reviews, ensuring accountability and responsible adoption of AI.
- Strengthening human oversight and governance: AI systems should not operate autonomously but should be subject to continuous human oversight and governance. Allowing individuals to override AI-generated decisions and holding responsible parties accountable for AI-driven outcomes is essential to ensure ethical AI operations.
- Establish a comprehensive ethical framework: the AI sector needs to adopt a robust ethical framework and standards that guide the development and application of AI in a responsible and ethical manner. These frameworks should address privacy, reduce bias, and prevent corruption, ensuring consistency with ethical principles.
- Public Awareness and Education: Raising public awareness and promoting education about the potential risks and benefits of AI is essential to building a society that can harness the transformative power of AI while proactively addressing its misuse. This includes educating individuals about the ethical implications of AI and enabling them to identify and report potential corruption.
- Mitigating corruption threats and promoting ethical AI: By implementing these safeguards and promoting the responsible development of AI, we can minimize the corruption threats posed by AI and harness its potential to promote transparency, accountability and a more just society.

THE ROLE OF ARTIFICIAL INTELLIGENCE IN FACILITATING CORRUPTION

AI's remarkable automation, data analytics and adaptability make it a formidable tool for facilitating and hiding corrupt practices. AI-based chatbots can be deployed to impersonate government officials or real companies, paving the way for fraudulent transactions and phishing schemes. AI algorithms can also be used to manipulate data, hide illegal activities, and evade regulatory scrutiny. In addition, AI can be used to target specific individuals or groups, enabling discrimination and abuse of power. The Covid19 pandemic has significantly increased the threat of AI-based corruption. The shift to teleworking and heavy reliance on digital platforms has opened new avenues for deception and exploitation. During the pandemic, AI-driven tools were used to spread misinformation, impersonate healthcare providers, and target susceptible individuals. As AI becomes more deeply embedded in our daily lives, these threats will only become more prevalent.

Examples of AI-enabled corruption

1) Political corruption:

a) Microtargeting: AI can analyze vast amounts of voter data to identify specific demographics and tailor political messages, accordingly, influencing voting behavior.

b) Social media manipulation.

c) Deepfake: Creating deepfake videos of political actors can foment discord, undermine trust in democratic institutions and influence elections.

2) Financial corruption:

a) Automated Fraud: AI algorithms can automate fraudulent transactions, bypassing traditional detection methods, making it more difficult to identify and prevent financial crimes.

b) Exploiting legal loopholes: AI can identify and exploit loopholes in financial systems, facilitating money laundering and other illegal activities.

c) Transaction masking: AI-driven systems can mask the true nature of financial transactions, obscuring the origin and destination of illicit funds, making it difficult to track and trace illegal financial flows.

3) Environmental corruption:

a) Hiding illegal activities: AI can be used to analyze satellite imagery and manipulate geospatial data, hiding illegal logging, mining and pollution activities from regulatory scrutiny.

b) Manipulating environmental data: AI can be used to manipulate environmental data, presenting a false picture of environmental compliance, allowing polluters to operate without consequences.

c) Regulatory evasion: AI can be used to automate and streamline environmental permitting processes, allowing corrupt officials to approve harmful projects without proper scrutiny.

4) Labor violations:

a) Surveillance and monitoring: AI-based surveillance systems can track employee movements, monitor communications, and identify potential labor organizers, allowing employers to suppress labor protests and union activities.

b) Automated repression.

c) Discrimination: AI algorithms can be biased in their hiring, promotion, and performance appraisal practices, leading to discrimination and unfair treatment of workers.

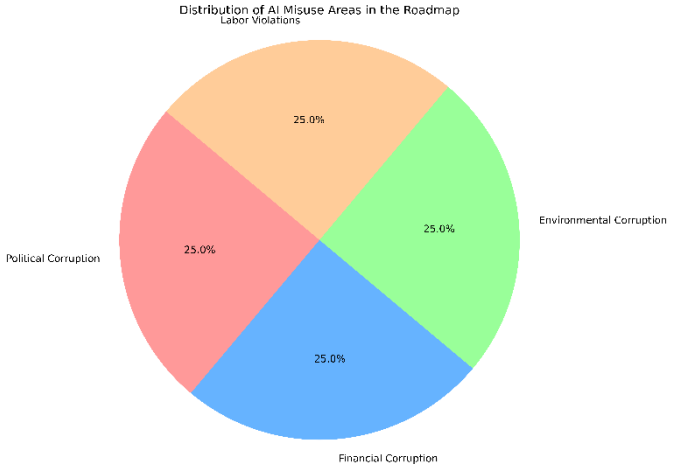


Figure 2. Self-made chart. 'Distribution of AI Misuse Areas in the Roadmap.'. (2024)

EUROPEAN LEGISLATION TO MITIGATE AI-BASED CORRUPTION

Current legal framework:

- Regulation (EC) No 2019/2161 [2] of the European Parliament and of the Council on the ethical development and use of AI: sets out ethical requirements for the development and use of AI systems, including transparency, accountability, fairness, and non-interference.
- Regulation (EC) No 2016/680 [3] of the European Parliament and of the Council on the protection of natural persons regarding the processing of personal data and on the free movement of such data and repealing Regulation (EC) No 95/46/EC (General Data Protection Regulation, GDPR) [12].
- The Commission Communication 2020 [4] outlines a strategy for Artificial Intelligence in Europe, including the Artificial Intelligence Regulation and the Digital Services Act (DSA) [5]. The AI Act will regulate AI systems' market entry, classifying them into risk categories and imposing requirements on high-risk systems. The DSA aims to improve accountability and transparency of digital platforms, including AI-based ones. Strategic initiatives include the European AI Partnership, which invests in AI research and development, and the European AI Observatory, which monitors AI developments and advises the EU on AI policy. The aim is to minimize corruption risks, protect privacy, protect consumers, and promote EU competitiveness in AI. The European Union is actively addressing potential risks associated with Artificial Intelligence, particularly those related to corruption. This

proactive stance is reflected in a multi-pronged legislative approach being developed and implemented. The proposed Artificial Intelligence Act [6] seeks to establish a comprehensive framework governing the development, deployment, and use of AI within the EU, with high-risk systems facing stricter regulations. These measures could include mandatory human oversight, algorithmic transparency requirements, and robust testing and validation procedures.

- The existing General Data Protection Regulation (GDPR) plays a crucial role in

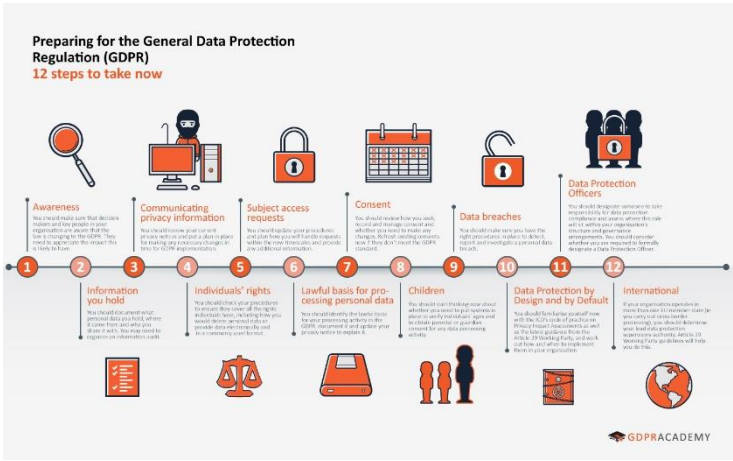


Figure 3. 'Checklist', GDPR Academy. Accessed: April 20, 2024. 23:21:43. Available: https://www.gdpracademy.org/gdpr-checklist-are-you-ready/

mitigating AI-based corruption by emphasizing data privacy and security [7] . The EU is considering expanding the scope of GDPR [12] to specifically address AI-related data use and potential biases within algorithms. The EU's existing anti-corruption frameworks, such as the Convention on the Fight against Corruption involving Parties to the Council of Europe (CETS No. 191) [8], provide a foundation for adapting regulations to encompass new risks associated with AI use in specific areas. For example, existing anti-bribery and undue influence legislation can be strengthened to address potential corruption risks in AI-driven public procurement processes. Transparency requirements within anti-corruption frameworks can ensure the explainability and fairness of algorithmic decision-making used by public authorities.

The EU's commitment to tackling AI-based corruption extends beyond existing legislation. It is expected to continue refining its approach through strategies such as developing specific guidance, introducing targeted regulations, and fostering international cooperation.

- International legislation to mitigate corruption.
- The lack of a unified international law specifically addressing AI-based corruption is a growing concern. The OECD has released recommendations on responsible AI development, emphasizing fairness, transparency, and accountability. The UN Office on Drugs and Crime is working on developing a global framework for preventing corruption in the context of AI, focusing on areas like public procurement, algorithmic bias, and data security. The UN Convention Against Corruption, which focuses on bribery and money laundering, can be adapted to address new corruption risks arising from AI use by member states. The European Union is at the forefront of tackling AI-based corruption with initiatives like the proposed Artificial Intelligence Act (AIA) and leveraging existing frameworks like the GDPR [19], [20], [21], [22], [23]. Other regional blocs are developing their own approaches, such as the African Union. Challenges in developing a truly international framework for AI regulation include harmonization, technical expertise, and global cooperation.

A multifaceted strategy, including legislative measures, technological advances, and ethical considerations, is essential to effectively combat the growing threat of AI-enabled corruption:

- A robust legislative framework: establishing clear and comprehensive legislation to regulate the development, deployment, and use of AI to prevent the abuse of corrupt practices. Put in place data protection and security standards to prevent the exploitation of sensitive information for corrupt purposes. Enforce strict sanctions and penalties for individuals and organizations involved in AI-based corruption.
- Technological innovation: developing AI-based tools and techniques to detect, investigate and prevent corruption, such as AI-based tracking systems and fraud analytics. The use of AI for increased transparency and accountability, including the disclosure of AI-based decisions and control mechanisms. Encourage research and development on ethical AI to ensure that AI systems are designed and implemented in a responsible and accountable manner.
- Ethical Guidelines and Human Oversight: Establish clear ethical frameworks and guidelines for the development and application of AI to prevent bias, discrimination, and abuse. Ensure human oversight and decision-making processes without human intervention to prevent AI from acting autonomously and making decisions that may have ethical consequences. Promote digital literacy and awareness among citizens to enable them to identify and report corruption schemes involving AI.
- Cultivating transparency and ethics in AI: Enforce accountability of AI by providing transparent decision-making processes and ethical guidelines for AI algorithms. Prohibit bias and discriminatory practices in the development and application of AI to promote equitable outcomes. Encourage public scrutiny of AI systems to maintain trust and prevent abuse for corrupt purposes.

- Data security and privacy: adopt robust data governance practices to protect sensitive information from unauthorized access and misuse. Implement robust encryption mechanisms to protect data security and privacy. Robust security and privacy safeguards are in place to ensure robust security and privacy safeguards are in place.
- Establish a clear regulatory framework: formulate comprehensive and enforceable rules for the development, deployment, and monitoring of AI. Establish clear rules and sanctions for the misuse of AI in relation to corruption and unethical practices. Establish robust enforcement and accountability mechanisms to deter and punish AI-based corruption.
- Promote public awareness and education: raise public awareness of the potential risks of AI-based corruption to promote responsible decision-making. Educate individuals about the ethical implications of AI and prepare them to identify and resist AI-facilitated fraud. Encourage individuals to report suspicious AI-based activities to the competent authorities for investigation.

Examples from the past

- The Cambridge Analytica scandal: in 2018, the data analytics firm Cambridge Analytica was embroiled in a huge scandal over the misuse of Facebook user data. The company collected and analyzed data from millions of Facebook users without their consent, and then used that data to target voters with personalized political ads during the 2016 US presidential election [13], [14]. This scandal has raised concerns about whether artificial intelligence can be used for political manipulation and voter suppression.
- The Russian troll farms. The Internet Research Agency (IRA), a St. Petersburg-based organization, is accused of using social media platforms like Twitter [11], Facebook, and TikTok to spread misinformation and influence public opinion in favor of Russian interests. This can involve creating fake accounts, posting biased content, and manipulating online conversations. The IRA has been linked to attempts to interfere in global elections, particularly the 2016 US presidential election. They have also been accused of discrediting critics of the Russian government and its policies. The widespread use of disinformation tactics can erode public trust in democratic institutions, exacerbate social divisions, and pose challenges for social media platforms in identifying and removing malicious content.
- Examples of arming AI for surveillance: Skynet, a network of cameras with facial recognition technology, is used for public surveillance and tracking individuals in China. The Social Credit System assigns scores based on social behavior, financial history, and online activity. Palantir Gotham is a data analysis platform used by law enforcement to track people, places, and events. Domain Awareness System (DAS) collects data from CCTV cameras and other sources to track people and vehicles in cities like Chicago. UK uses CCTV with facial recognition software, but privacy concerns remain. India uses an Integrated Network Security System (INSS) for citywide surveillance and monitoring.
- Creating a deepfake: Deepfakes are videos or audio recordings that have been manipulated to spread misinformation and damage reputations and can undermine

public trust in institutions and individuals. Deepfakes have become a significant issue in various fields, including politics, entertainment, and social media.

In the 2019 Belgian Election [9] a video allegedly showing offensive remarks by a political candidate was debunked, highlighting the potential for misinformation to disrupt elections. In the 2020 US Election [10] several deepfakes targeting political candidates circulated online, raising concerns about misinformation and its influence on voters. Celebrities like Scarlett Johansson and Gal Gadot have also been targeted by deepfakes, highlighting their potential for harassment and exploitation. In the entertainment industry, Tom Cruise's hyper realistic videos went viral on TikTok, blurring the line between reality and fiction. The American sketch comedy show, Saturday Night Live, has also used deepfakes to create satirical content impersonating celebrities and politicians. These scandals highlight the spread of misinformation, privacy erosion, and potential for harm, such as harassment and blackmail.

- AI chatbots can impersonate trusted individuals, leading to fraud and phishing scams [16],[17],[18]. They can manipulate public opinion through social media manipulation, suppression of dissent, and social engineering. Social media manipulation involves analyzing data to tailor messaging, while suppression involves censorship and surveillance. AI can also generate fake news articles and social media posts, undermining trust in institutions [24], [25], [26], [27], [28]. For example, a government might target young people with messages that downplay the importance of voting. This includes China's Social Credit System, which uses AI to track and rate citizens based on their online activity, financial history, and social behavior, and Russia's interference in elections through social media manipulation tactics and disinformation campaigns.

THE RESULTS OF THE RESEARCH

In my research, I used combined methods to examine the connection between artificial intelligence and human individuals based not only on narrative literature review, but also on a survey created solely for this purpose. Given the answers of the research participants, we can conclude that Artificial Intelligence (AI) is a powerful tool that uses computational techniques to learn from data and perform tasks that require human intelligence. It excels in pattern recognition and extracting insights from vast datasets, revolutionizing fields like automation, machine learning [29], and data science.

Artificial Intelligence (AI) offers transformative potential but also presents significant risks, particularly in the context of corruption and governance. AI systems, especially in law enforcement, hiring, and criminal justice, can perpetuate societal biases. Eubanks [30] highlights how predictive policing algorithms disproportionately affect marginalized communities, while Binns et al. [28] show that AI in hiring or credit scoring can exacerbate inequalities if not carefully designed.

The "black-box" nature of many AI models complicates accountability and transparency, potentially enabling corrupt practices. Lipton [31] emphasizes the challenge of

making AI models interpretable while maintaining their predictive power, an issue critical for reducing corruption in sectors like public administration.

AI's use in surveillance and information manipulation also presents corruption risks. Tufekci [35] warns of AI's weaponization in authoritarian regimes, where it can undermine democracy and facilitate political manipulation.

To address these concerns, researchers focus on Explainable AI (XAI) to improve transparency. Carvalho et al. [29] explore methods for developing interpretable models, and Ribeiro et al. [33] propose counterfactual explanations to provide alternative decision scenarios, mitigating bias.

Efforts to detect and mitigate algorithmic bias have intensified. Mehrabi et al. [32] review techniques for bias detection, such as data preprocessing and fairness constraints, while Sculley et al. [34] introduce fairness algorithms to combat systemic biases, especially in sectors vulnerable to corruption.

To ensure accountability, Zhao et al. [36] propose fairness-aware auditing tools to evaluate AI models for bias and transparency, essential in areas like judicial decision-making and government surveillance. These advancements are critical for preventing AI systems from fostering corruption or injustice.

-	Robots and AI in everyday life	AI vs. Real workforce	Positive impact of AI
count	15.0	15.0	15.0
mean	3.467	1.466	3.345
std	1.0.68	0.7433	0.8165
min	2.0	1.0	2.0
25%	3.0	1.0	3.0
50%	3.0	1.0	4.0
75%	4.0	2.0	4.0
max	5.0	3.0	4.0
mode	3.0	1.0	4.0

Figure 4. Self-made chart. Situations regarding artificial intelligence, chatbots and programmed empathy. (2024).

The chart above presents three crucial point in the social acceptance of artificial intelligence among the research participants that serves as a reflection on the usage of AI in everyday life:

- Robots and AI in everyday life: The average rating is approximately 3.47, with a mode of 3. This suggests a generally positive attitude towards the integration of robots and AI in everyday life.
- AI vs. real workforce: The average rating is about 1.47, with a mode of 1, indicating a strong disagreement with the idea that AI is more useful than the real workforce.
- Positive impact of AI: The average rating is around 3.33, with a mode of 4, showing a positive perception of AI's impact.

CONCLUSION AND SUGGESTIONS

The research reveals a mixed public perception of Artificial Intelligence (AI), with cautious optimism. Some see AI as a valuable automation tool, enhancing efficiency and streamlining tasks. However, concerns remain about AI's potential dangers, such as job displacement [15] and potential misuse for malicious purposes. The high usage of AI features suggests a complex interplay between convenience and apprehension, highlighting the need for continued development. The study identifies three key areas for improvement: accuracy, ethical considerations, and human judgment integration.

Accuracy is a desire for more reliable AI systems, which could involve advancements in training data and algorithms to minimize errors. Ethical considerations are also a priority, requiring robust ethical frameworks for AI development and deployment. Human judgment integration is a preference for a collaborative approach between humans and AI, incorporating human oversight into critical decision-making processes. Overall, the study highlights the need for continued development to address these concerns and ensure the effective use of AI in daily life.

I would like to emphasize the importance of transparency and human oversight in AI development: explainable AI (XAI) tools that are user-friendly, fostering trust and accountability. In my opinion, AI is to be a powerful tool, not a replacement for human judgment. There is a growing demand for a culture of ethical AI, involving shared responsibility between developers, policymakers, and the public in creating robust ethics frameworks. The need to prioritize fairness throughout the AI development process, including data collection and algorithm design remains a slowly expanding territory, therefore widespread AI education campaigns should be organized to help people understand the capabilities and limitations of AI, and whistleblowers should be able to report suspected AI misuse without fear of retaliation. This field of academic research requires further work as Artificial Intelligence develops.

SUMMARY

The research explores public perception of Artificial Intelligence (AI), highlighting concerns about job displacement and misuse. It suggests three areas for improvement: accuracy, ethical considerations, and human judgment integration. The study advocates for explainable AI tools, fostering trust and accountability. It calls for a culture of ethical AI, where responsibility is shared between developers, policymakers, and the public. It also calls for AI education campaigns to help people understand AI's capabilities and encourage whistleblowers to report misuse. AI remains in its developmental stages, requiring further academic exploration.

BIBLIOGRAPHY

Treaties and acts

[1] Council of Europe, ‘Council of Europe – Additional Protocol to the Criminal Law Convention on Corruption (ETS No. 191) – Translations - Treaty Office - www.coe.int’, Treaty Office. Accessed: March 26, 2024. [Online]. Available:

https://www.coe.int/en/web/conventions/-/council-of-europe-additional-protocol-to-the-criminal-law-convention-on-corruption-ets-no-191-translations
[2] European Commission, 'White Paper on Artificial Intelligence: A European approach to excellence and trust'. Accessed: Apr. 22, 2024. [Online]. Available: https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en
[3] European Parliament, 'Artificial intelligence act', 0 2023.
[4] International Telecommunication Union, 'United Nations Activities on Artificial Intelligence (AI)', ITU Hub. Accessed: Apr. 22, 2024. [Online]. Available: https://www.itu.int/hub/publication/s-gen-unact-2022/
[5] O. Radley-Gardner, H. Beale, and R. Zimmermann, Eds., Fundamental Texts On European Private Law. Hart Publishing, 2016. doi: 10.5040/9781782258674.
[6] 'Artificial intelligence as an anti-corruption tool (AI-ACT): Potentials and pitfalls for top-down and bottom-up approaches', arXiv: Computers and Society, Feb. 2021, Accessed: Mar. 19, 2024. [Online]. Available: https://typeset.io/papers/artificial-intelligence-as-an-anti-corruption-tool-ai-act-1csi4pbaj7

Articles and books

[7] Margarita Leib, Nils C. Köbis, Rainer Michael Rilke, Marloes Hagens, 'Corrupted by Algorithms? How AI-generated and Human-written Advice Shape (Dis)honesty', Social Science Research Network, vol. abs/2301.01954, Jan. 2023, doi: 10.48550/arXiv.2301.01954.
[8] Nils Köbis, Christopher Starke & Iyad Rahwan, 'Artificial intelligence as an anti-corruption tool (AI-ACT): Potentials and pitfalls for top-down and bottom-up approaches', arXiv: Computers and Society, Feb. 2021, Accessed: Mar. 05, 2024. [Online]. Available: https://typeset.io/papers/artificial-intelligence-as-an-anti-corruption-tool-ai-act-1csi4pbaj7
[9] Nils Köbis, Christopher Starke& Iyad Rahwan, 'Bad machines corrupt good morals.', Nature Human Behaviour, vol. 5, no. 6, pp. 679–685, Jun. 2021, doi: 10.1038/S41562-021-01128-2.
[10] A. Ray, 'Disinformation, Deepfakes and Democracies: The Need for Legislative Reform', UNSWLJ, vol. 44, no. 3, Sep. 2021, doi: 10.53637/DELS2700.
[11] D. Wiessner and D. Wiessner, 'Disabled employee sues Twitter over Musk's ban on remote work', Reuters, Nov. 17, 2022. Accessed: March 26, 2024. [Online]. Available: https://www.reuters.com/business/disabled-employee-sues-twitter-over-musks-ban-remote-work-2022-11-17/
[12] 'Checklist', GDPR Academy. Accessed: March 26, 2024. [Online]. Available: https://www.gdpracademy.org/gdpr-checklist-are-you-ready/
[13] 'The deepfake 2020 election threat is real, but containable | TechTarget', Enterprise AI. Accessed: March 26, 2024. [Online]. Available: https://www.techtarget.com/searchenterpriseai/feature/The-deepfake-2020-election-threat-is-real-but-containable
[14] 'Corrupted by Algorithms? How AI-generated and Human-written Advice Shape (Dis)honesty', Jan. 2023, doi: 10.48550/arxiv.2301.01954.

[15] ‘Artificial Intelligence Decision Making and the infringement of human rights: Focusing on the recruitment process’, Han’gug bu’pae haghoebo, vol. 28, no. 1, pp. 5–30, Mar. 2023, doi: 10.52663/kcsr.2023.28.1.5.
[16] Kollár Csaba. A mindennapok mesterséges intelligenciája. (2024)’. Accessed: Jun. 20, 2024. [Online]. Available: https://m2.mtmt.hu/api/publication/34914783
[17] A. Kiss and C. Kollár, ‘Az információbiztonság időszerű kérdései a magyarországi kkv-k körében’, ScientSec, vol. 4, no. 2, pp. 98–107, Apr. 2024, doi: 10.1556/112.2023.00166.
[18] I. Jagodics and C. Kollár, ‘21. századi social engineering támadások, védekezés és szervezeti hatások Európában’, BSZ, vol. 71, no. 1, pp. 111–126, Jan. 2023, doi: 10.38146/BSZ.2023.1.6.

Directives

[19] ‘Directive - 2019/2161 - EN - omnibus directive - EUR-Lex’. Accessed: March 26, 2024. [Online]. Available: https://eur-lex.europa.eu/eli/dir/2019/2161/oj
[20] ‘Strategic plan 2020-2024 – Communication - European Commission’. Accessed: March 26, 2024. [Online]. Available: https://commission.europa.eu/publications/strategic-plan-2020-2024-communication_en
[21] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, vol. 281. 1995. Accessed: March 26, 2024. [Online]. Available: http://data.europa.eu/eli/dir/1995/46/oj/eng
[22] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, vol. 119. 2016. Accessed: March 26, 2024. [Online]. Available: http://data.europa.eu/eli/dir/2016/680/oj/eng
[23] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance), vol. 277. 2022. Accessed: March 26, 2024. [Online]. Available: http://data.europa.eu/eli/reg/2022/2065/oj/eng

Conference papers

[24] ‘Criminological risks and legal aspects of artificial intelligence implementation’, presented at the International Conference on Artificial Intelligence, ACM Press, Dec. 2019. doi: 10.1145/3371425.3371476.
[25] ‘The corruptive force of AI-generated advice’, arXiv: Artificial Intelligence, Feb. 2021, Accessed: Mar. 19, 2024. [Online]. Available: https://typeset.io/papers/the-corruptive-force-of-ai-generated-advice-4j3sdufmbb

[26] 'A study on the constitutional suitability of artificial intelligence-based anti-corruption tools(AI-ACT)', Han'gug bu'pae haghoebo, vol. 27, no. 1, pp. 5–30, Mar. 2022, doi: 10.52663/kcsr.2022.27.1.5.
[27] 'Digtial technologies in combating global corruption', Вісник Харківського національного університету імені В.Н. Каразіна, no. 41, pp. 30–39, Jul. 2022, doi: 10.26565/2220-8089-2022-41-04.
[28] R. Binns, H. Prakken, and A. Stent, "Algorithmic Fairness: Insights from Data Science and Ethics," Proc. of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1-14, 2021, [Online]. Available: https://doi.org/10.1145/3411764.3445683.
[29] D. Carvalho, F. Pereira, and M. Silva, "Explaining and Interpreting Deep Learning Models: A Survey of Recent Approaches," IEEE Trans. Neural Networks Learn. Syst., vol. 32, no. 8, pp. 3359-3374, 2021, [Online]. Available: https://doi.org/10.1109/TNNLS.2020.2983477.
[30] V. Eubanks, Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor, St. Martin's Press, 2020.
[31] Z. C. Lipton, "The Mythos of Model Interpretability," Commun. ACM, vol. 64, no. 3, pp. 34-37, 2021, [Online]. Available: https://doi.org/10.1145/3341574.
[32] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A Survey on Bias and Fairness in Machine Learning: Towards a Systematic Literature Review," ACM Comput. Surv., vol. 54, no. 6, pp. 1-35, 2021, [Online]. Available: https://doi.org/10.1145/3285029.3285037.
[33] M. T. Ribeiro, S. Singh, and C. Guestrin, "Anchors: High-Precision Model-Agnostic Explanations," Proc. 32nd Conf. Neural Inf. Process. Syst. (NeurIPS 2020), pp. 1-10, 2020, [Online]. Available: https://arxiv.org/abs/2008.02625.
[34] D. Sculley, M. Murnane, and C. Tan, "Fairness in Machine Learning: A Survey of Methods and Applications," ACM Comput. Surv., vol. 54, no. 5, pp. 1-35, 2021, [Online]. Available: https://doi.org/10.1145/3419765.
[35] Z. Tufekci, The Global Surveillance State: Technology, Privacy, and Governance in the Digital Age, Routledge, 2021.
[36] H. Zhao, J. Wang, and X. Yu, "Fairness-Aware Auditing of AI Systems: Principles and Applications," J. Artif. Intell. Res., vol. 70, no. 1, pp. 431-456, 2021, [Online]. Available: https://doi.org/10.1613/jair.1.11820.