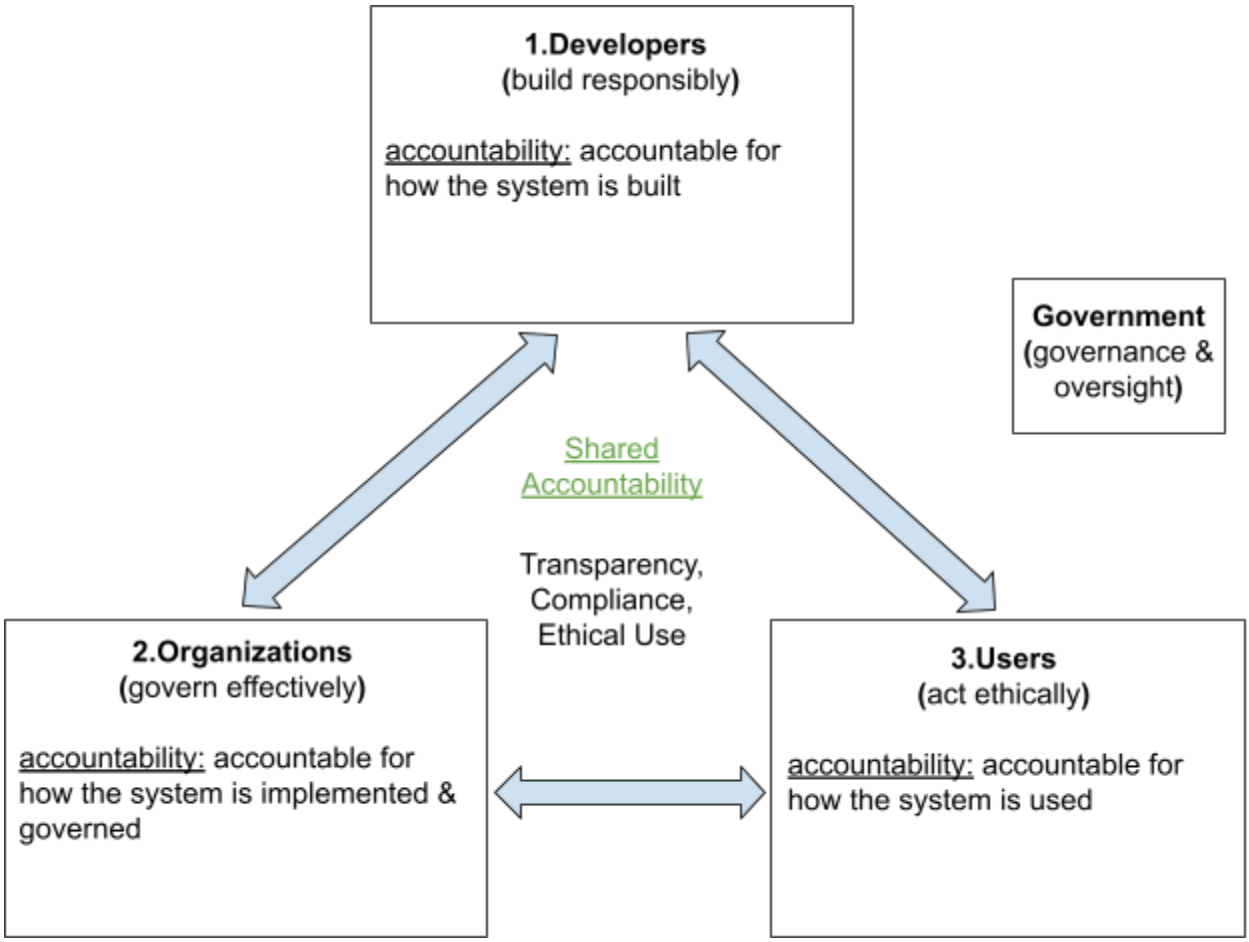
The Accountability Model:



Contributions:

- 1) who should be accountable for the misuse of AI in cybersecurity -- between the users, developers, and organizations?
- 2) comparison of different AI regulations -- between the US, EU, and Saudi Arabia
- 3) what are the core ethics/principles to developing a safe/secure AI regulations

Accountability Model: With this model, we compare and contrast what kind of responsibilities do users, developers, and organizations have in regard to the use of AI. As well as, exploring where the accountability lies as AI is misused by malicious actors in cybersecurity.

Analysis of the accountability model:

In this model, each party—users, developers, and organizations—has its own distinct responsibilities that collectively contribute to the overall safety and security of AI use. However, accountability must first start with the AI creators in developing and maintaining an effective yet secure AI system. They are liable for how their AI systems are built, for implementing measures to prevent its misuse, and for testing for any vulnerabilities and backdoors that could be exploited. Additionally, they are required to establish clear policies and usage guidelines, warning users about acceptable and prohibited uses of AI [15].

The next layer of responsibility relies on the organizations and individual companies to deploy defensive strategies to protect against vulnerabilities. As well as deploying deterrents, such as explicit policies, rules, and penalties for violations. Even though organizations cannot safeguard their systems from each person individually, they can deter them from having unethical and illegal thoughts through these measures. Making it clear that if anyone carries out such behaviors, they are expected to be caught and that penalties will be administered [3], [6].

Ultimately, the true prevention of AI misuse lies with users themselves, for not misusing it in the first place. It is up to the users to apply their ethical principles and morals to decide what is right and wrong and what not to do [2]. In fact, both the AI developers and organizations can greatly influence user behavior. Through guidelines

and deterrence, they can help define which actions are illegal or unethical and encourage users to act responsibly.

This leads to another aspect of the accountability model. The accountability model employs a system of checks and balances among its three parities, much like the three branches of the U.S. government. Thus, each party can actually help keep the others in check while supporting one another's role in achieving secure AI use [10]. For example, AI developers can restrict or regulate certain features to limit misuse by users and organizations; both developers and organizations can influence user behavior through policy enforcement; and by monitoring user activities, developers and organizations can strengthen their defenses and develop responsive policies. Between these three parties, they really have an interdependent relationship and that safeguarding AI use is ultimately a collective effort from all three parties. Thus, accountability does not—and should not—fall solely on any one party, but rather be shared among all of them [9]. Each shares similar responsibilities and can mutually enhance one another's role through transparency, compliance, and ethics.

Finally, the accountability model emphasizes the need for a larger governing body or authority—specifically, the federal government—to regulate these three parties. Only through the government oversight and governance, can we ensure that these parties truly comply and carry out their roles and responsibilities. Moreover, with the government establishing official laws, regulations, and penalties, it enforces a universal standard of accountability. Thus, allowing all parties to operate under consistent ethical and legal expectations [9], [14]. In conclusion, this accountability model not only specifies the individual responsibilities of the users, developers, organizations, and the government, it also demonstrates that they are interconnected, sharing the accountability of safeguarding AI use, and that they operate in a system of checks and balances where each party can regulate one another.

References

[1]

E. Tabassi, “Artificial Intelligence Risk Management Framework (AI RMF 1.0),” National Institute of Standards and Technology (U.S.), Gaithersburg, MD, NIST AI 100-1, Jan. 2023. doi: 10.6028/NIST.AI.100-1.

[2]

N. Polemi, I. Praça, K. Kioskli, and A. Bécue, “Challenges and efforts in managing AI trustworthiness risks: a state of knowledge,” *Front Big Data*, vol. 7, p. 1381163, May 2024, doi: 10.3389/fdata.2024.1381163.

[3]

D. K. A. Abdelaziz, “Criminal liability for the misuse and crimes committed by AI: A comparative analysis of legislation and international conventions,” *Journal of Infrastructure, Policy and Development*, vol. 9, no. 1, p. 10722, Jan. 2025, doi: 10.24294/jipd10722.

[4]

C. Velasco, “Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments,” *ERA Forum*, vol. 23, no. 1, pp. 109–126, 2022, doi: 10.1007/s12027-022-00702-z.

[5]

C. Novelli, F. Casolari, P. Hacker, G. Spedicato, and L. Floridi, “Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity,” *Computer Law & Security Review*, vol. 55, p. 106066, Nov. 2024, doi: 10.1016/j.clsr.2024.106066.

[6]

R. Rodrigues, “Legal and human rights issues of AI: Gaps, challenges and vulnerabilities,” *Journal of Responsible Technology*, vol. 4, p. 100005, Dec. 2020, doi: 10.1016/j.jrt.2020.100005.

[7]

T. Krishnamani, “LEGAL CHALLENGES IN REGULATING AI-POWERED HACKING, PHISHING, AND IDENTITY THEFT,” *Lex localis - Journal of Local Self-Government*, vol. 23, no. S4, pp. 1889–1900, Aug. 2025.

[8]

S. Wachter, “Limitations and Loopholes in the EU AI Act and AI Liability Directives: What This Means for the European Union, the United States, and Beyond,” *SSRN Journal*, 2024, doi: 10.2139/ssrn.4924553.

[9]

M. V. Zucca and G. Fiorinelli, “Regulating AI to Combat Tech-Crimes: Fighting the Misuse of Generative AI for Cyber Attacks and Digital Offenses,” *Technology and Regulation*, vol. 2025, pp. 247–262, July 2025, doi: 10.71265/23nqtq40.

[10]

E. Papagiannidis, P. Mikalef, and K. Conboy, “Responsible artificial intelligence governance: A review and research framework,” *The Journal of Strategic Information Systems*, vol. 34, no. 2, p. 101885, June 2025, doi: 10.1016/j.jsis.2024.101885.

[11]

M. L. Montagnani, M.-C. Najjar, and A. Davola, “The EU Regulatory approach(es) to AI liability, and its Application to the financial services market,” *Computer Law & Security Review*, vol. 53, p. 105984, July 2024, doi: 10.1016/j.clsr.2024.105984.

[12]

L. Coppolino, S. D’Antonio, G. Mazzeo, and F. Uccello, “The good, the bad, and the algorithm: The impact of generative AI on cybersecurity,” *Neurocomputing*, vol. 623, p. 129406, Mar. 2025, doi: 10.1016/j.neucom.2025.129406.

[13]

I. Jada and T. O. Mayayise, “The impact of artificial intelligence on organisational cyber security: An outcome of a systematic literature review,” *Data and Information Management*, vol. 8, no. 2, p. 100063, June 2024, doi: 10.1016/j.dim.2023.100063.

[14]

M. Buiten, A. de Streel, and M. Peitz, “The law and economics of AI liability,” *Computer Law & Security Review*, vol. 48, p. 105794, Apr. 2023, doi: 10.1016/j.clsr.2023.105794.

[15]

B. C. Cheong, “Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making,” *Front. Hum. Dyn.*, vol. 6, July 2024, doi: 10.3389/fhumd.2024.1421273.