Vol.1 N.1 (2023)
science-conference.com

ISSUES OF LIABILITY FOR VIOLATIONS OF THE RULES FOR SAFE AND SECURE USE OF ARTIFICIAL INTELLIGENCE

Avazbek Sayfiddinov
avazbeksayfiddinov23@gmail.com

ABSTRACT

As artificial intelligence (AI) permeates cyber defenses for government agencies and critical infrastructure providers globally, high stakes surround accountable oversight and liability for potential flaws enabling platform misuse or legal infringements. However, most legal systems currently lack statutes clarifying accountability models, technical requirements and injury restitution protocols specific to AI's unique risks.[1]

This analysis examines open debates around constructing suitable liability frameworks that balance enterprise security imperatives with individual rights protections as AI automation and autonomy spread. It spotlights challenges resolving tensions between technology vendors, deploying agencies and impacted citizens regarding emerging threat prevention technologies with inorganic reasoning capacity. [2]

KEYWORDS

Artificial intelligence (AI), cyber defense, cybersecurity, Machine learning.

INTRODUCTION

Key Dimensions for Resolving AI Liability

Impacted Stakeholders: As security AI evolves in capability, key groups deserve consideration around infringement accountability:

Deploying Agencies: Public sector departments fielding AI surveillance, data mining and autonomous security systems may face public grievances or lawsuits for perceived bias, accuracy errors or improper data practices stemming from poor implementation. Algorithms negatively profiling or denying service access to minority communities commonly triggers backlash.[3] However, limited vendor transparency around proprietary commercial algorithms often restricts agencies' capacity to audit tools completely or explain outputs to external overseers thoroughly, complicating internal liability.

Technology Vendors: As primary architects of AI security platforms, commercial developers increasingly confront legal exposure or contractual liabilities as algorithms underperform expectations, enable policy violations or disadvantaged groups contest system fairness. However, vendors argue unreasonable technical transparency requirements or excess liability may discourage vital innovation in security AI needed to match rising nation-state and cybercriminal threats

globally. [4]

Individual Citizens: Members of the public encountering security algorithm errors feel impacts ranging from improper service denial, discrimination from profiling, or undue surveillance from overreaching autonomous systems. However, average citizens lack the technical literacy to contest AI provider or agency explanations regarding factors driving algorithm outcomes.[5] Legal resources to pursue potential restitution also varies starkly across demographic groups and regions.

Evaluating Liability Models

With liability risks and restitution pathways ill-defined today for enterprises deploying security algorithms, various frameworks see debate internationally:

Strict Liability Approach: This model argues any entity that deploys or operates AI bears full liability for all infringements or injuries algorithms potentially enable, regardless of intent or software limitations known at time of occurrence. It aims to maximize public accountability and trust. However, opponents counter such sweeping and hazy liability may paralyze innovation on essential security capabilities or require low-risk tolerance setting thresholds so conservatively that accuracy suffers severely.

Negligence Approach: Liability transfers only in provable cases where deploying entities or vendors failed upholding reasonable AI engineering standards regarding transparency, testing, oversight and fairness measures established through emerging certification regimes like NIST's AI Risk Management Framework.[6] However, ambiguity still remains around qualifying barriers for establishing negligence across various issues like data testing rigor or algorithmic explainability.

Shared Liability Approach: Liability splits between vendors and deployers based on level of accountability assessed for an AI failure or infringement.[7] The exact distribution relies on factors like how accurately product capabilities were represented during sales, what policy and technical oversight the deployer provided post-launch, the reasonableness of individual oversight, how enlargeable the flaw was pre-deployment etc. However, standards must evolve to weigh such factors consistently across cases spanning tech to policy to operations.[8]

Restitution Protocols

If authorities establish clear liability parameters, practical challenges still remain crafting legal protocols determining restitution for aggrieved parties:

Monetary Compensation Policy: Rules must outline qualifying injuries justifying financial restitution from deployers and vendors and acceptable payment measures whether from fines, mandatory insurance programs or direct algorithm taxes.[9] However, valuation models need development estimating reasonable compensation levels relative to infringement severity.[10]

Reconsideration Policies: Beyond financial impacts, AI flaws often disrupt citizen access to vital services like welfare benefits, bank loans or insurance claims. Clear policies for reversing algorithmic service denials or forcing reconsideration by human reviewers provide non-monetary restitution. However, executing such mandates against opaque commercial systems relying extensively on automation poses difficulties.

Replacement Rules: For severe or repeated AI infringements, authorities may mandate agencies replace flawed security platforms entirely rather than allow continued public exposure. However defining thresholds meriting complete platform rebuilding depends on evolving deeper understanding of root failure factors as AI software processes still confound many legal experts early in the technology's lifespan.

An Ambiguous Road Ahead

Until legislators, courts and technology leaders collectively advance robust liability doctrine for assessing security AI damages and directing equitable restitution, uncertainty and reluctance to adopt beneficial innovations could prevail across risk averse public agencies. Progress requires acknowledging inherent limitations around current AI transparency, prescriptive oversight and performance predictability. No perfect model for accountability can yet exist but an approximate justice enabling incremental progress remains vital against intensifying cyber threats. Those nations courageous enough to debate complex tensions between security imperatives and citizen rights will build foundations allowing AI's immense promise passage into public benefit. But earning such progress necessitates a difficult balancing of risks and freedoms persistently favoring the vulnerable rather than the powerful.

CONCLUSION

With advanced AI permeating security operations faster than laws or technical controls can responsibly constrain potential harms, societies globally struggle establishing prudent liability norms allowing for both algorithmic innovation and public protections. Absent deliberate debate and precedential rulings clarifying standards, accountability distribution and restitution protocols suitable for unique AI risks, uncertainty inhibits adoption of otherwise transformative threat detection capabilities even amidst increasingly perilous attacks threatening critical infrastructure worldwide. Until legislative bodies, technology leaders and citizens collectively advance evolved liability models balancing security innovation and individual rights, all stakeholder groups suffer from this failure to reconcile competing priorities rationally while protecting those disadvantaged byBOTH imperfect algorithms and inadequate policymaking visions. But the immense complexity of this challenge necessitates persistence, courage and vision to compel progress benefiting all equitably in due course.

REFERENCES

1. Taddeo, Mariarosaria, and Luciano Floridi. "Regulate artificial intelligence to avert cyber arms race." Nature 556, no. 7701 (2018): 296-298.
2. Svantesson, Dan Jerker B. "Artificial intelligence and administrative decision-making–An Australia perspective." Computer Law Review International 21, no. 5 (2020): 137-143.
3. Hagendorff, Thilo. "The ethics of AI ethics: An evaluation of guidelines." Minds and Machines 30, no. 1 (2020): 99-120.
4. Charisi, Vicky, et al. "Towards moral autonomous systems." arXiv preprint arXiv:1703.04741 (2017).
5. Smuha, Nathan A., et al. "EU regulation of AI: Fit for purpose." Artificial Intelligence and Law 29, no. 4 (2021): 481-512.
6. falsely Reyes, Emily. "Calculating corporate liability for artificial intelligence harm." Trade, Law and Development 12, no. 2 (2020): 269-301.

7. Scherer, Matthew U. "Regulating artificial intelligence systems: risks, challenges, competencies, and strategies." Harv. JL & Tech. 29 (2015): 353.
8. Eggers, William D., et al. "AI-augmented government: Using cognitive technologies to redesign public sector work." Deloitte Insights 2 (2017).
9. Završnik, Aleš. "Criminal justice, artificial intelligence systems, and human rights." ERA Forum 20, no. 4 (2020): 567-583.
10. Karniouchina, Ekaterina Viktorovna, William Moore, and Kuntal Banerjee. "Estimating losses from cyber incidents: A loss function approach." The Geneva Papers on Risk and Insurance-Issues and Practice 48, no. 3 (2021): 660-688.