35TH USENIX SECURITY SYMPOSIUM

"*Abuse risks are often inherent to product features*": Exploring AI Vendors' Bug Bounty and Responsible Disclosure Policies

Yangheran Piao (lawrence.piao@ed.ac.uk), Jingjie Li (jingjie.li@ed.ac.uk), Daniel W. Woods (daniel.woods@ed.ac.uk)

Introduction

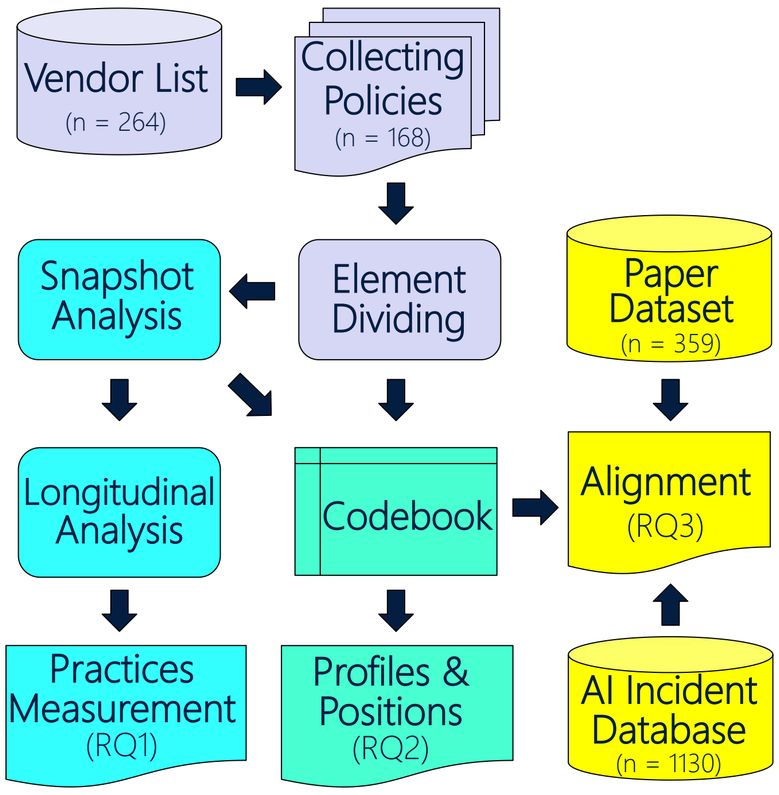
- Alongside the rapid growth of AI vulnerabilities, we are now seeing the first wave of companies beginning to adapt their disclosure practices.
- However, how AI vendors define, structure, and communicate disclosure policies remains largely undocumented.
- To address this gap, we conducted a mixed-methods study. We ask:

RQ1. What is the state of vulnerability disclosure in the AI industry?

RQ2. How do vendors approach AI vulnerabilities?

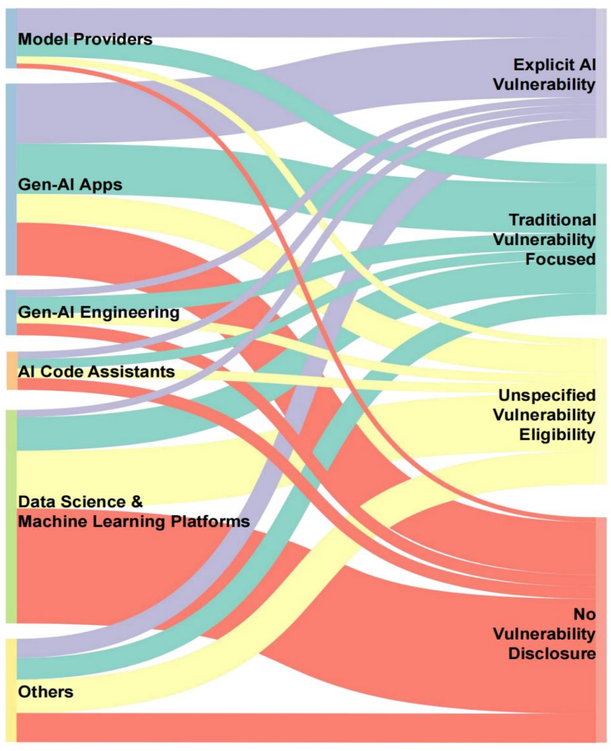
RQ3. What is the alignment with AI incidents and research?

Approach



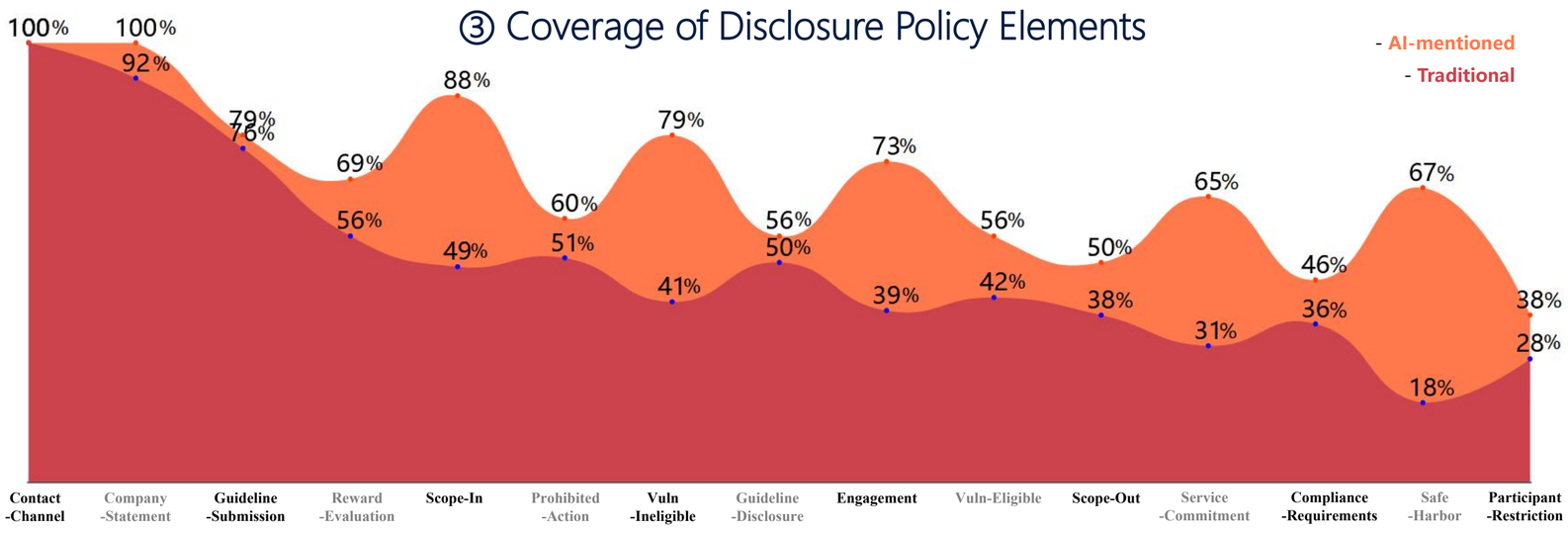
AI Vulnerability Reporting Practices Measurement RQ1

① AI companies vs Disclosure Approaches

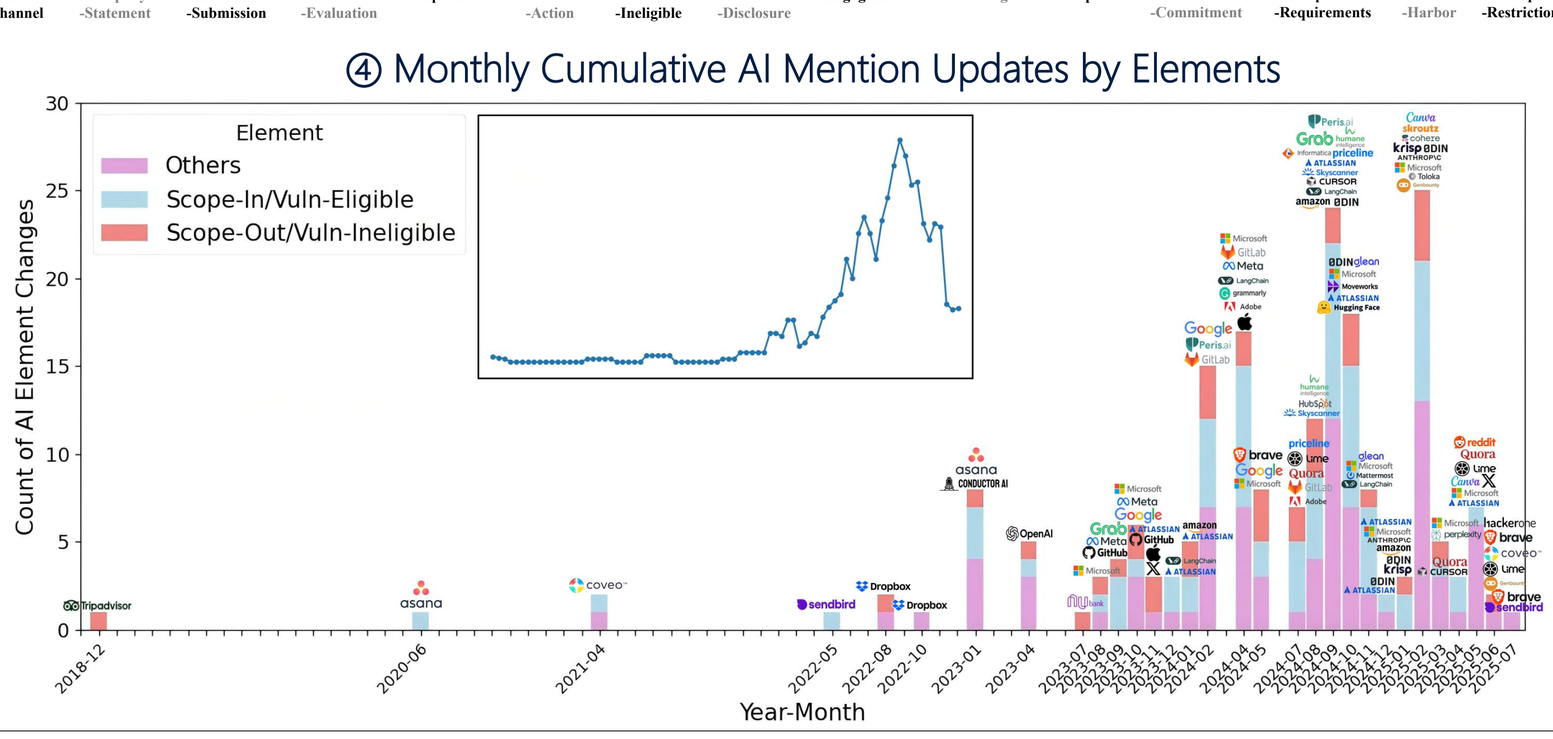


② AI Vulnerability Types with Eligibility

③ Coverage of Disclosure Policy Elements



④ Monthly Cumulative AI Mention Updates by Elements



Profiles and Positions RQ2

① Approaches for Addressing AI Vulnerabilities

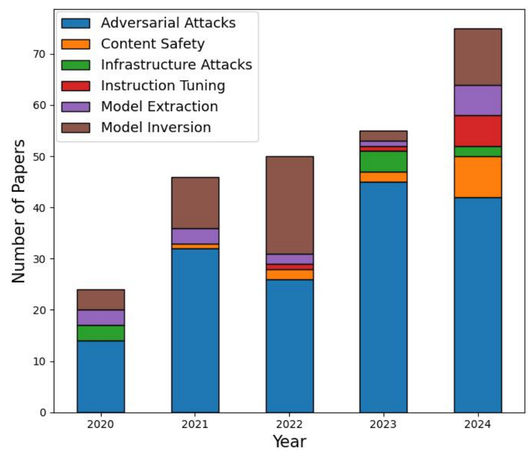
Proactive Clarification: Vendors explicitly define AI vulnerabilities as in scope and welcome reports

Silent: Policies make no mention of AI vulns, leaving ambiguity for researchers

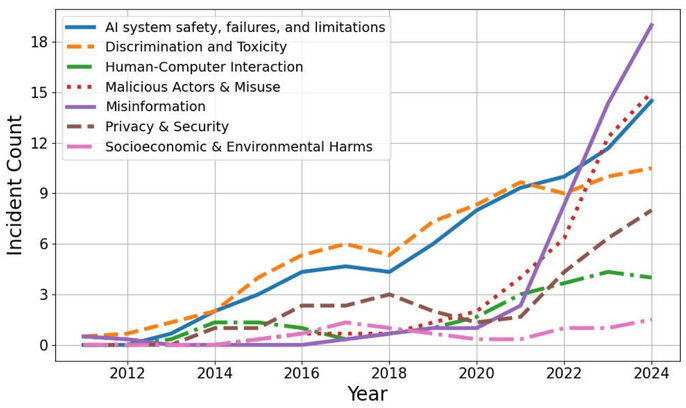
Restrictive: Vendor policies exclude AI vulns from scope or provide no reporting channel

AI Incident & Research Alignment RQ3

① Yearly distribution of AI security papers



② Monthly distribution of AI incidents



Conclusion

- Only **64%** of AI companies maintain a disclosure channel, and explicit recognition of AI vulnerabilities is limited (**18%**), with approaches varying widely.
- Disclosure policies and academia have more focus on upstream technical issues in isolation, relative to the AI incidents that are dominated by content safety issues.

Piao, Y., Li, J., Woods, D. W. "Abuse risks are often inherent to product features": Exploring AI vendors' bug bounty and responsible disclosure policies. In 2026 USENIX Security Symposium (USENIX SEC'26).

THE UNIVERSITY of EDINBURGH | SECURITY, PRIVACY & TRUST GROUP