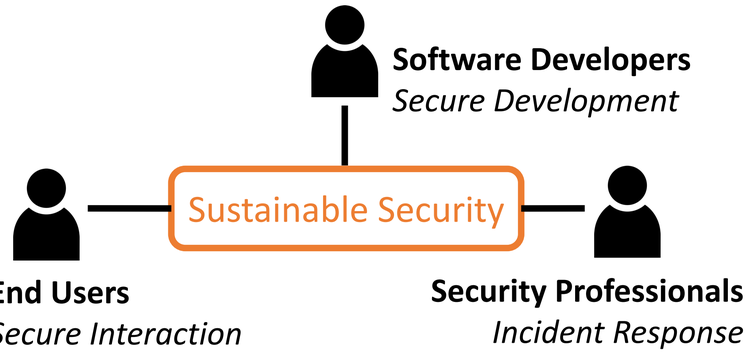
Human-AI Collaboration for Sustainable Security: Opportunities and Challenges

Wanling Cai, Liliana Pasquale, Kushal Ramkumar, John McCarthy, Bashar Nuseibeh, and Gavin Doherty

Background

- Cyber-threats and attacks increase in both quantity and complexity in the digital age.
- AI-powered autonomous and adaptive security approaches aim to enhance system security with *minimal human intervention*.

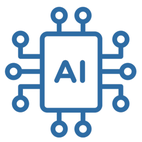
However, to establish a *resilient and sustainable security ecosystem*, we rely on *different stakeholders* to contribute their efforts at different stages of securing systems.



Research Question

"human as problem"

"human as solution" [1]



How can we use AI to augment human capability in implementing *sustainable security practices*?

Human-AI Collaboration for Sustainable Security

Opportunities: AI Can Act as a Collaborator or an Educator [2]

Secure Development

- Assess and regenerate secure code
- Prioritize test cases

Secure Interaction

- Monitor user activities and context
- Identify potential security risks
- Take proactive security measures

Incident Response

- Perform analysis of system logs
- Identify anomalies
- Respond to security threats

Collaboration Design Considerations

- Transparency
- Contextual Recommendation
- Explanation (XAI)
- Learner-centered Explanation

- Level of automation
- Level of human control
- Adaptation
- Conversational AI

- Effective communication
- Contextual Explanations
- Real-time Feedback

Challenges: Enabling Effective Collaboration for Sustainable Security Practices

Appropriate Reliance [3]

Encouraging human developers to think critically when interacting with automated code generation

Long-term Engagement

Engaging users in adopting security measures and maintaining their security awareness for a long-term

Transparent Collaboration

Engineering transparent collaboration in such high-risk and time-critical conditions

Conclusion and Future Work

- Human capabilities can be augmented by AI collaboration to establish a more sustainable security ecosystem.
- Future work will examine the pros and cons of existing AI tools for security practices and investigate multiple stakeholders' perception of human-AI collaboration in security-critical contexts, e.g., in smart home.

[1] Zimmermann, Verena, and Karen Renaud. "Moving from a 'human-as-problem" to a 'human-as-solution" cybersecurity mindset." *International Journal of Human-Computer Studies* 131 (2019): 169-187.
[2] Wang, Dakuo, et al. "Human-AI collaboration in data science: Exploring data scientists' perceptions of automated AI." *Proceedings of the ACM on human-computer interaction* 3.CSCW (2019): 1-24.
[3] Lee, John D., and Katrina A. See. "Trust in automation: Designing for appropriate reliance." *Human factors* 46.1 (2004): 50-80.