

Do cybersecurity professionals have an ethical obligation to disclose 'morally opaque' features in AI-driven security tools, even if it reveals vulnerabilities or proprietary technology?

According to Brey's (2000) disclosive method discussed in our lectures, many technological features are "morally opaque", meaning users aren't aware of their moral implications. This is particularly relevant in cybersecurity, where AI-driven malware detection systems and automated incident response tools may make decisions that affect privacy, security, and individual rights without users understanding how or why.

As cybersecurity professionals, we face a dilemma: should we disclose these hidden features to stakeholders and the public to ensure transparency and informed consent? Or does protecting system integrity and proprietary technology take precedence?

Our lectures presented two perspectives that might guide this decision. Gotterbarn (1991) argues from a professional ethics standpoint that computer ethics is fundamentally about professional responsibility to clients and employers. Meanwhile, Moor (1985) takes a philosophical ethics approach, emphasizing the broader social impact of technology and the need for policies that serve the public good.

Microsoft's "The Future Computed" (2018) reinforces this tension by identifying transparency and accountability as core AI principles, yet acknowledging the complexity of implementing these in practice.

I'm curious whether the answer changes depending on context, for example, should critical infrastructure systems (power grids, healthcare) have different disclosure requirements than commercial security products? Does the potential harm from non-disclosure outweigh the risks of revealing vulnerabilities?

References

- Brey, P. (2000). Disclosive computer ethics. *ACM SIGCAS Computers and Society*, 30(4), 10–16.
<https://doi.org/10.1145/572260.572264>

Moor, J. H. (1985). *What is computer ethics?* **Metaphilosophy**, 16(4), 266–275.
<https://doi.org/10.1111/j.1467-9973.1985.tb00173.x>

Gotterbarn, D. (1991). Computer ethics: Responsibility regained. *National Forum*, 71(3), 26.

Microsoft Corporation. (2018). *The future computed: Artificial intelligence and its role in society* (B. Smith & H. Shum, Forewords).

https://news.microsoft.com/cloudforgood/_media/downloads/the-future-computed-english.pdf