Codacy

AI and Cyber Security

Business Survival in the Age of Accelerating Cyber Threats

April 2024

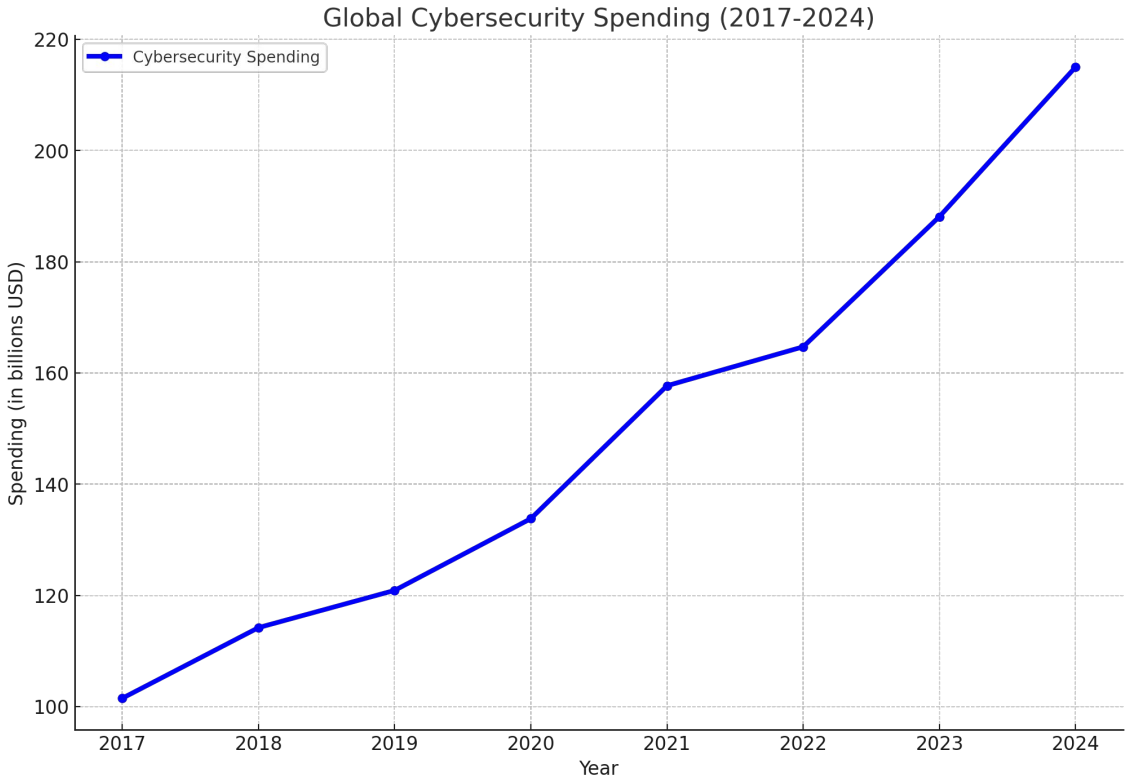
Jaime Jorge
Co-founder & CEO of Codacy

The world cares about Cyber Security

$215b

Total world spending in Cybersecurity according to Gartner.

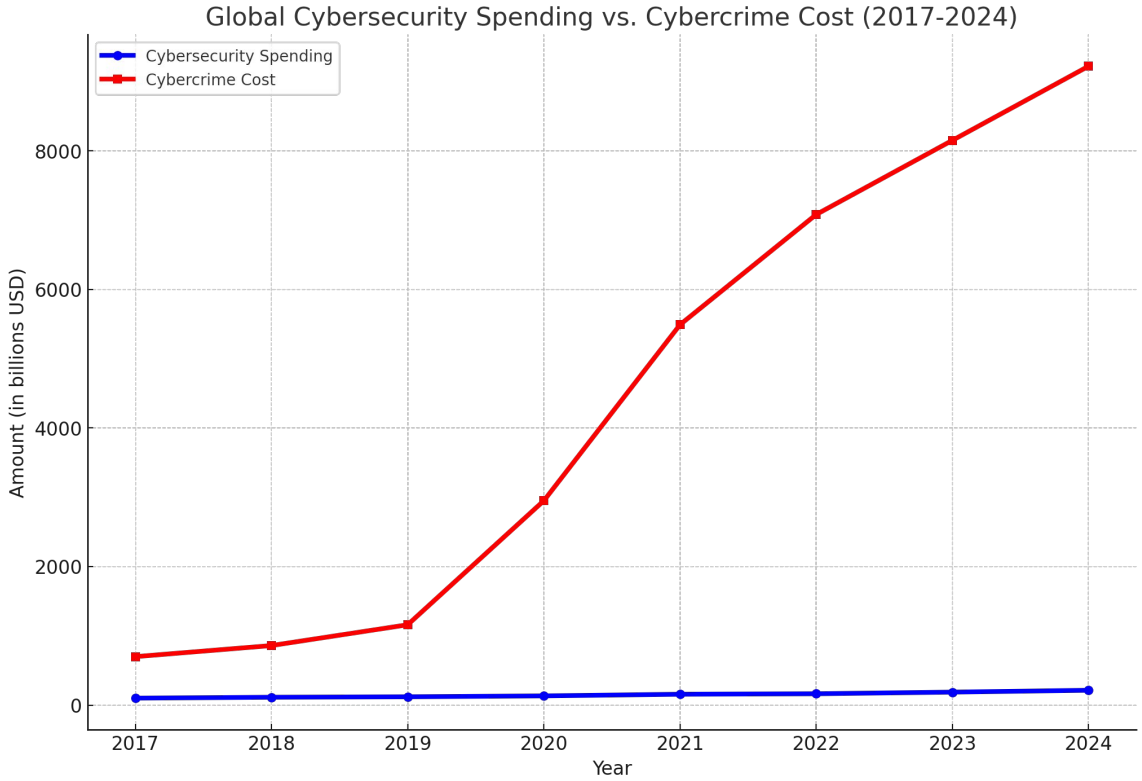
This investment has been increasing over the years.



This investment has been increasing over the years.

But it’s eclipsed by the cost of the impact.

Cybercrime is growing faster than cybersecurity spending by about 39.36% annually.



Today:

1.

Why costs are going to accelerate

2.

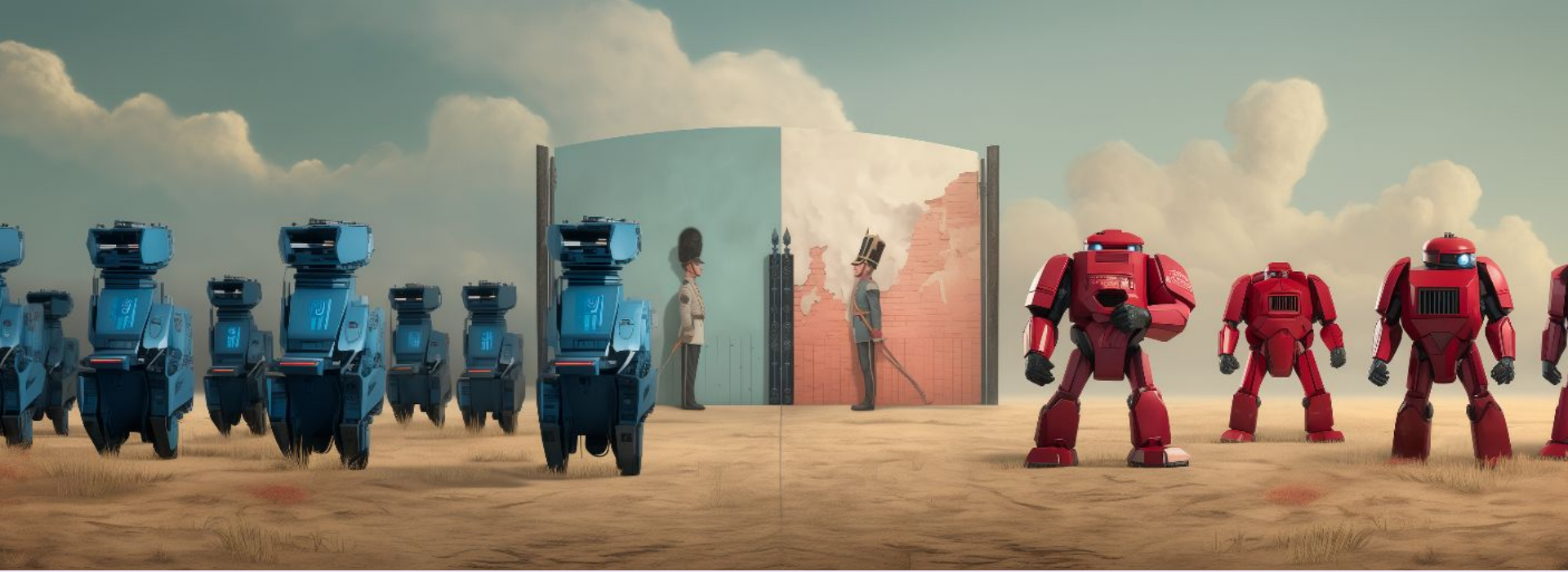
Some companies will suffer more than others

Part 1

Accelerating costs

Cybersecurity and cybercrime is always a standoff

That is limited by humans and some automation



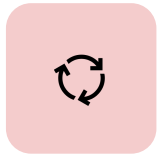
AI scales that conflict.

AI is changing offensive / defensive security

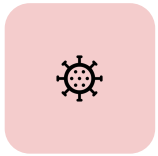
Offensive

Phishing

Ransomware



Supply Chain



Malware

Code Injection Attacks

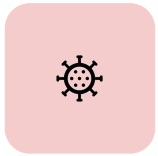
DDoS

Offensive

Phishing

Ransomware

Supply Chain



Malware

Code Injection Attacks

DDoS

Media Contact:
MediaRelations@fcc.gov

For Immediate Release

FCC MAKES AI-GENERATED VOICES IN ROBOCALLS ILLEGAL

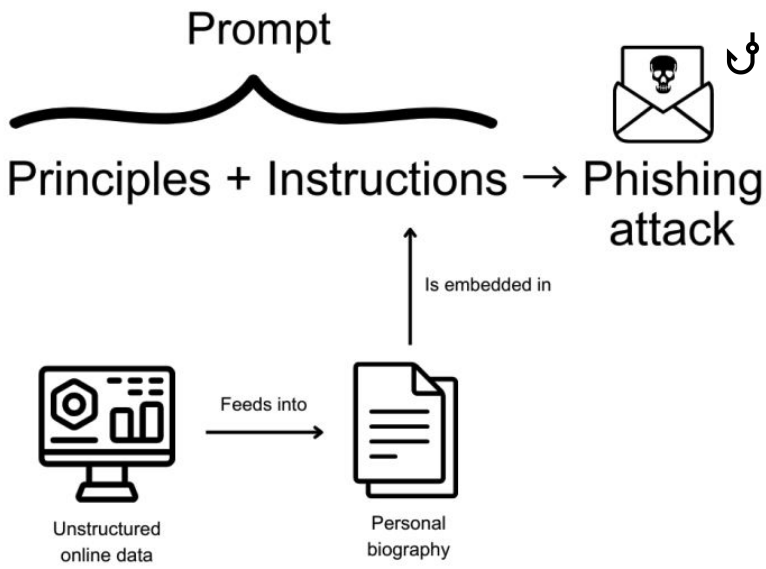
State AGs Will Now Have New Tools to Go After Voice Cloning Scams

Higher conversion rate of success for attackers

On a **small scale**, phishing isn't very successful, but on a large scale, the one or two victims that fall for the scheme make it worth it.

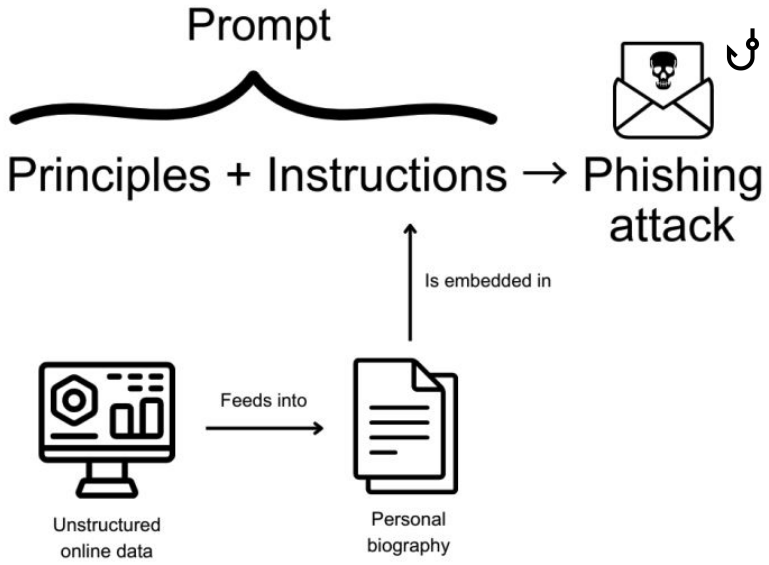
AI enables **spear phishing** attacks at scale.

Phishing



"Spear Phishing With Large Language Models", J Hazell 2023

Every hour, Multiple times



"Spear Phishing With Large Language Models", J Hazell 2023

Online Identification is fragile

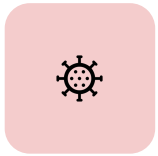
*“Just from analyzing a **small clip from an online video,** scammers can **replicate a voice** to a chilling degree of accuracy and use it to call your loved ones pretending to be you.“*

Offensive

Phishing

Ransomware

Supply Chain

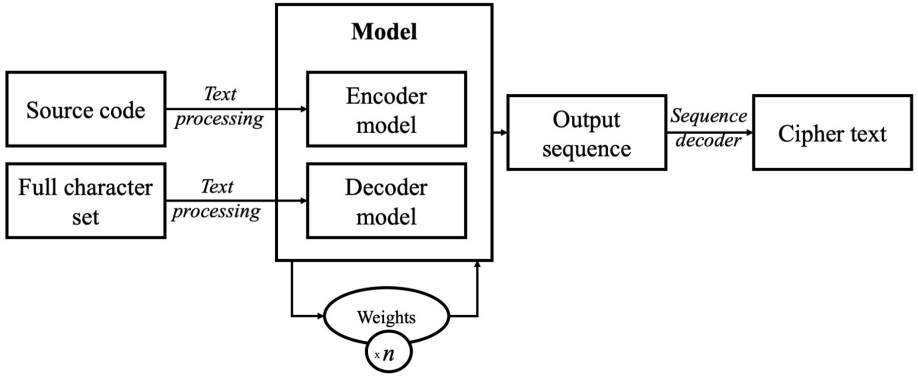


Malware

Code Injection Attacks

DDoS

GAN to obfuscate Malware

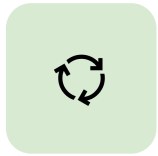


S. Datta. 2020. DeepObfusCode: Source Code Obfuscation Through Sequence-to-Sequence Networks. In Advances in Intelligent Systems and Computing

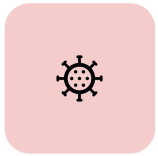
Offensive

Phishing

Ransomware



Supply Chain



Malware

Code Injection Attacks

DDoS

Reverse Engineering and Vulnerability detection is enabled by AI

- Tiffany Bao et al. 2014. {BYTEWEIGHT}: Learning to recognize functions in binary code. In 23rd {USENIX} Security Symposium ({USENIX} Security 14). 845–860.
- Steven HH Ding et al. 2019. Asm2vec: Boosting static representation robustness for binary clone search against code obfuscation and compiler optimization. In 2019 IEEE Symposium on Security and Privacy (SP). IEEE, 472–489.
- Yue Duan et al. 2020. DEEPBINDIFF: Learning Program-Wide Code Representations for Binary Diffing. In Proceedings of the 27th Annual Network and Distributed System Security Symposium (NDSS'20).
- Qian Feng et al. 2016. Scalable graph-based bug search for firmware images. In Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. 480–491.
- Mikolás Janota. 2018. Towards Generalization in QBF Solving via Machine Learning.. In AAAI. 6607–6614.
- Jian Jiang et al. 2019. A Survey of the Software Vulnerability Discovery Using Machine Learning Techniques. In International Conference on Artificial Intelligence and Security. Springer, 308–317.
- Vitaly Kurin et al. 2019. Improving SAT solver heuristics with graph networks and reinforcement learning. arXiv preprint arXiv:1909.11830 (2019).
- Zhen Li et al. 2019. A comparative study of deep learning-based vulnerability detection system. IEEE Access 7 (2019), 103184–103197.
- Zhen Li et al. 2018. Vuldeepecker: A deep learning-based system for vulnerability detection. arXiv preprint arXiv:1801.01681 (2018).
- Jia Hui Liang et al. 2018. Machine learning-based restart policy for CDCL SAT solvers. In International Conference on Theory and Applications of Satisfiability Testing. Springer, 94–110.
- Bingchang Liu et al. 2018. αdiff: cross-version binary code similarity detection with dnn. In Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering. 667–678.
- Horst Samulowitz and Roland Memisevic. 2007. Learning to solve QBF. In AAAI, Vol. 7. 255–260.
- Eui Chul Richard Shin et al. 2015. Recognizing functions in binaries with neural networks. In 24th {USENIX} Security Symposium ({USENIX} Security 15). 611–626.
- Yan Wang et al. 2020. A systematic review of fuzzing based on machine learning techniques. PloS one 15, 8 (2020), e0237749.
- Xiaojun Xu et al. 2017. Neural Network-based Graph Embedding for Cross-Platform Binary Code Similarity Detection. Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (Oct 2017). https://doi.org/10.1145/3133956.3134018
- Fangke Ye et al. 2020. MISIM: An End-to-End Neural Code Similarity System. arXiv preprint arXiv:2006.05265 (2020).
- Seongjun Yun et al. 2019. Graph transformer networks. arXiv preprint arXiv:1911.06455 (2019).

Information from Mirsky, Yisroel, et al. "The threat of offensive ai to organizations." *Computers & Security* 124 (2023)

Any vulnerability in software will be more likely to be found

Thoughts

AI will enable automation of human behavior in cybercrime. Targets that would otherwise be economically unreasonable will become viable and profitable

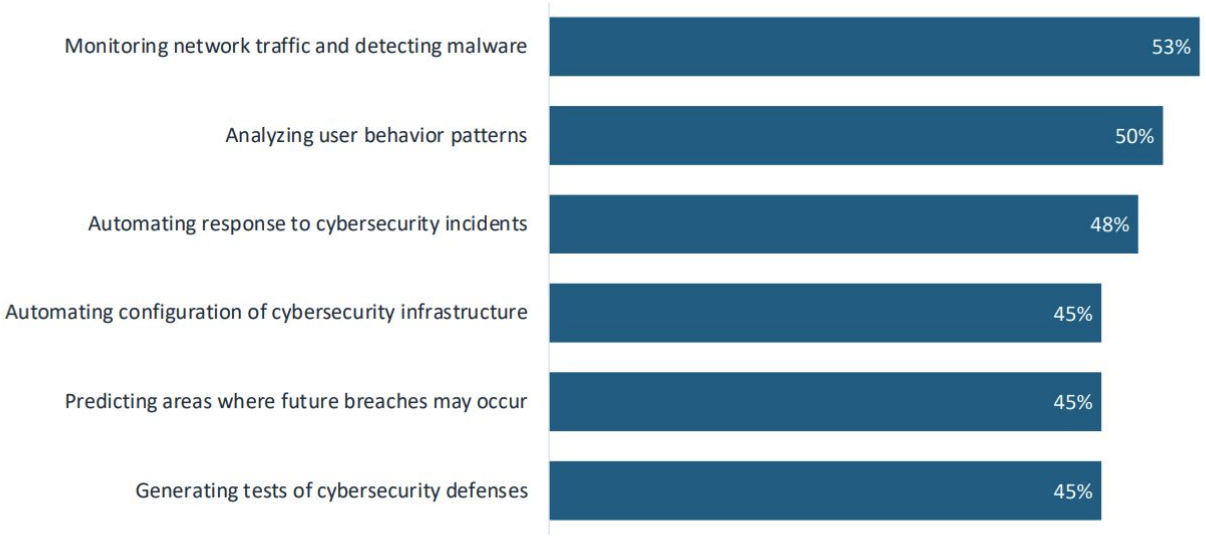
Lowered bar for cybercrime → more attackers

AI creates uncertainty in online human protocols.

We'll have **AIs battling AIs** where we control them. We'll live in digital walled gardens

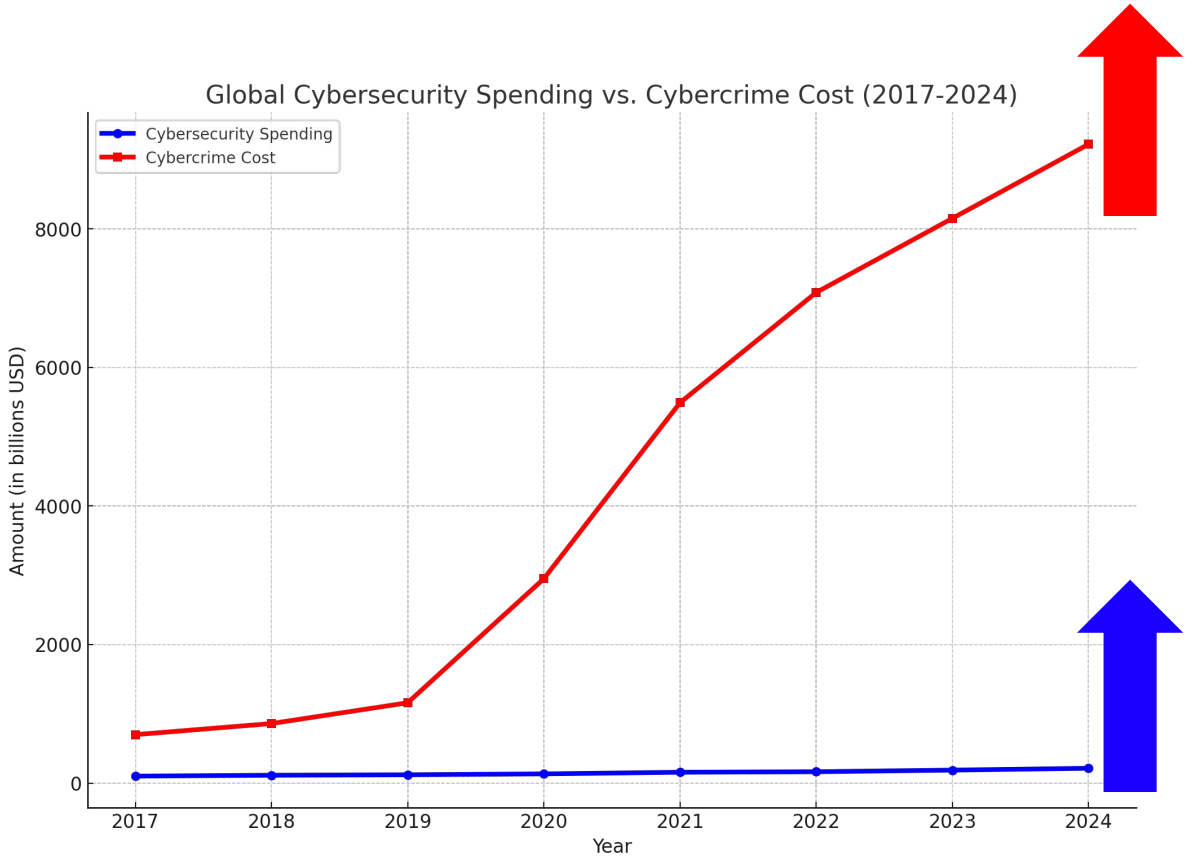
AI is changing offensive / defensive security

Potential Uses of AI in Cybersecurity



Source: CompTIA 2024 State of Cybersecurity | n=511 U.S. technical and business professionals

AI brings cybersecurity arms race.



Part 2

SMBs will suffer more

43%

attacks are aimed at SMBs

14%

of SMBs are prepared for attacks

47%

SMBs have fallen victim to a cyberattack in 2022

Why?

- **Lack of resources**
- **Not the target of security vendors**
- **Not enough time**
- **Protected by the herd**

Before, SMBs would be safe in a herd

The cost of opportunity for attackers was too high

However, AI scales attackers.
Humans don’t limit reach.

So the herd is no longer protection. It’s profit.

At Codacy we believe that Security is akin to a fundamental right.

To make every line of code trustworthy.

Codacy

@codacy
@jaimefjorge
Jaime at codacy.com