The Benefits and Challenges of AI in Cybersecurity

Kai Fleming
December, 14, 2024

Pēdējos gados, mākslīgais intelekts ir kļuvis svarīgāks daudzās jomās. No transporta līdz izklaidei, un kiberdrošība nav izņēmums. Kiberdrošība ir veids, kā aizsargāt tīklus, ierīces un datus no nepiederošām personām. Tā ir svarīga, lai pasargātu cilvēku privāto informāciju un intelektuālo īpašumu no krāpniekiem un zagļiem. Mūsdienās ļaundari ir sākuši izmantot MI tehnoloģijas, lai ātrāk uzlauztu kontus un uzņēmumus, tāpēc organizācijām ir jāmeklē jauni kiberdrošības risinājumi, izmantojot MI. Ar MI palīdzību organizācijas var iegūt daudz priekšrocību, piemēram, labāku draudu atklāšanu un ātrāku reaģēšanu, bet MI lietošanai ir arī savas problēmas.

Vispirms, MI sniedz vairākas priekšrocības tīklu un ierīču aizsardzībā salīdzinājumā ar tradicionālo kiberdrošību. Vecākas aizsardzības metodes galvenokārt izmanto zināmu draudu parakstus, lai atklātu uzbrukumus. Šī manuālā datu salīdzināšana ir viegli ievainojama pret jauniem nezināmiem draudiem, un tai ir nepieciešama cilvēka iejaukšanās. Turklāt vecās sistēmas bieži kļūdaini ziņo par draudiem, kas patiesībā nav īsti. Automatizējot šos procesus ar MI, organizācijas var efektīvāk cīnīties ar uzbrukumiem.

Pēc Cisco (liela kiberdrošības uzņēmuma) datiem 2018. gadā tika nobloķēti 7 triljoni draudu. 2020. gadā uzņēmumi, kas izmantoja MI automatizāciju, ietaupīja 3,58 miljonus dolāru visā pasaulē, salīdzinot ar uzņēmumiem bez MI. 2024. gadā MI tirgus kiberdrošībā ir aptuveni 25 miljardi dolāru, un tiek prognozēts, ka līdz 2034. gadam tas pieaugs līdz aptuveni 147 miljardiem dolāru.

MI galvenās priekšrocības ir tā spēja atklāt, paredzēt un reaģēt. MI visbiežāk izmanto, lai atklātu aizdomīgas darbības un kaitīgu kodu ienākošajos datos. Tas var arī analizēt lietotāju uzvedību, salīdzinot to ar parasto, lai atklātu iekšējos draudus un kontu parņemšanu. MI spēj vienlaicīgi pārbaudīt milzīgus datu apjomus, ļaujot uzņēmumiem aktīvi meklēt ļaundarus, nevis tikai pasīvi aizsargāt informāciju.

Turklāt MI var izmantot, lai modelētu kiberuzbrukumus, atklājot vājās vietas, kuras var nostiprināt, pirms kibernoziedznieki tās izmanto. MI var arī atpazīt parastus tīkla satiksmes modeļus, lai uzlabotu darba slodzi un pārvaldītu visas tīkla ierīces.

Beidzot, MI reaģēšanas spējas tiek izmantotas vismazāk, bet tas joprojām ir ļoti svarīgi efektīvai tīkla aizsardzībai. MI var automātiski sūtīt brīdinājumus, bloķēt kaitīgus datus un izolēt drošības pārkāpumus. Turklāt MI var palīdzēt kiberuzbrukumiem reālajā laikā, uzlabojot reaģēšanas laiku un veidojot spēcīgāku aizsardzību, balstoties uz iepriekšējiem gadījumiem. Piemēram, MI var brīdināt lietotājus par vājām, viegli uzminamām parolēm.

MI spēja atklāt draudus, paredzēt uzbrukumus un ātri reaģēt, kā arī automatizēt ikdienas uzdevumus, sniedz daudz pozitīvu ieguvumu. Piemēram, IT speciālisti var veltīt vairāk laika sarežģītākiem uzdevumiem un var ietaupīt daudz laika un resursu, kas citādi tiktu tērēti, izmeklējot viltus

trauksmes. Neskatoties uz šīm daudzajām priekšrocībām, uzņēmumi saskaras arī ar vairākiem trūkumiem, izmantojot MI.

MI lietošanai kiberdrošībā ir arī dažādi trūkumi. Viens piemērs ir aizspriedumu iespēja. Ja MI drošības programmatūra saņem nepareizus vai neobjektīvus apmācības datus, tas var novest pie diskriminējošiem rezultātiem vai viltus trauksmēm, kas vājina organizācijas drošību. Turklāt nepilnīgi vai nepareizi apmācības dati var likt MI kļūdaini interpretēt situāciju, bloķējot likumīgas darbības un atļautus lietotājus.

Vēl viens piemērs ir ētiskas problēmas, kas rodas, izmantojot MI kiberdrošībā. Algoritmi, kas analizē uzvedību, var pārkāpt cilvēku privātumu, kad tie vāc aizvien vairāk datu. Turklāt cilvēku mazāka iesaistīšanās samazina caurskatāmību un rada problēmas ar atbildību par iespējamiem kaitējumiem. Arī bezdarbs ir faktors, par ko jādomā. Automatizētie procesi var aizstāt IT darbiniekus, piemēram, klientu apkalpošanā, un tas var atturēt no MI izmantošanas.

Turklāt MI var radīt grūtības pat pirms tā ieviešanas, galvenokārt saistībā ar zināšanu trūkumu par MI, datu sarežģītību un programmatūras rīkiem, kas nepieciešami organizācijas atbalstam. Uzņēmumi, kas plāno ieviest MI, uzskata, ka būtiska problēma ir zināšanu trūkums par MI programmēšanu un pārvaldību kiberdrošībā. Organizācijas, kas jau izmanto MI, par lielāku problēmu uzskata lielos datu apjomus un pareizu MI rīku izvēli. Datu sarežģītība nozīmē grūtības pārvaldīt milzīgos ienākošos datu apjomus un iegūt kodu MI apmācībai. Kopā ar laiku un resursiem, kas nepieciešami MI izmantošanai, kā arī daudziem valdības noteikumiem, piemēram, Mākslīgā intelekta pamatkonvencijai, kas regulē MI lietošanu, mākslīgā intelekta ieviešana kiberdrošībā var būt sarežģīta lielākajai daļai uzņēmumu.

MI izmantošana kiberdrošībā neaprobežojas tikai ar organizācijām, kas aizsargā savus datus – hakeri un kibernoziedznieki izmanto MI, lai palielinātu savu darbību efektivitāti. Ļaundari var izmantot MI, lai automatizētu un personalizētu krāpšanas un pikšķerēšanas uzbrukumus sociālās inženierijas shēmās. Hakeri var apzināti manipulēt ar organizāciju MI ievadītajiem datiem, procesu, ko sauc par "fāzingu", lai atklātu vājās vietas. Visbeidzot, kibernoziedznieki var izmantot dziļviltotus attēlus un audio, lai radītu materiālus identitātes zādzībām, un mūsdienu algoritmiem ir viegli atšifrēt paroles. MI ir devis hakeriem daudzas iespējas izvairīties no aizsardzības un uzbruktu uzņēmumiem, radot daudz briesmu tīklu un sensitīvas informācijas aizsardzībai.

Kopumā MI izmantošana kiberdrošībā ir daudzpusīga problēma ar daudzām priekšrocībām un trūkumiem tās ieviešanā un izmantošanā. Modernas kiberdrošības metodes, kas izmanto MI, var izmantot arī kibernoziedznieki, lai uzbruktu biežāk un postošāk. 2025. gadā 93% drošības vadītāju sagaida MI uzbrukumus katru dienu, un vēl lielāks skaits – 95% drošības speciālistu uzskata, ka MI programmatūra uzlabos tīklu drošību, tā kā vadošie tehnoloģiju uzņēmumi, piemēram, Cisco, PaloAltoNetworks un CrowdStrike, strauji attīsta MI programmatūras, un mākslīgais intelekts turpina kļūt aizvien nozīmīgāks kā efektīvs kiberdrošības risinājums mūsdienu sabiedrībā.

Research
Cybersecurity
Source: What is cybersecurity? (US cybersecurity and infrastructure agency) I - What is cybersecurity? 1. Art of protecting___from unauthorized access/criminals a. Networks b. Devices c. data 2. Practice of ensuring: a. Confidentiality b. Integrity c. Availability of information 3. Many aspects of one's life relies on technology: a. Transportation b. Entertainment

c. Communication
d. Shopping
e. Medicine
f. ECT.

A - What are the risks to having poor cybersecurity?

1. Malware erasing systems/data
2. Alteration of files
3. Unauthorized purchases
4. Important info stolen
5. Your computer being used to attack others

B - What can you do to improve your cybersecurity?

RECOGNIZE THE RISKS

1. hacker/attacker/intruder:
    a. Exploit weakness of system for own gain
    b. Motivation may be curiosity/mischief to malicious activity
2. Malicious code
    a. Unwanted code/files that cause harm
    b. Types: worm, trojan horse, virus

MINIMIZING RISK:

3. Keep software up to date
4. Run (up to date) antivirus software
    a. detects/quarantines/removes malicious code
5. Use strong passwords
6. Change default username/password
7. Implement multi factor authentication
8. Install a firewall
9. Be cautious of certain emails/websites

AI

Source: What is AI? University of Illinois Chicago

I - What is AI?

1. Branch of computer science
2. Aim: machines capable of performing tasks that require human intelligence
3. Tasks:
    a. Learning from experience -> machine learning
    b. Understanding natural language
    c. Recognizing patterns
    d. Solving problems
    e. Making decisions
4. Can evolve/adapt using data -> improve performance over time

A - How does Artificial Intelligence work?

1. AI subsets: (subfields in AI that replicate different parts of intelligence)
    a. Machine learning: development of algorithm/models -> allows decision making + pattern recognition without programming explicitly
    b. Neural networks: replicates human brain structure. (layers of connected nodes)

-> image recognition + natural language processing + playing games
    c. Natural language processing: (NLP) understand/interpet/generate language
    d. Playing games: algorithms to play games like chess

B - Weak AI vs Strong AI:

1. Weak AI: specific task/ narrow application
2. Strong AI: understand/learn/apply data for multiple tasks

Cybersecurity and AI

Source: AI and Cyber Security (Morgan Stanley)

- Cyber security products: 15 billion (2021) 135 billion (2030)
- AI for cybersecurity: can analyze enormous data sets
    a. Detecting actual attacks accurately
    b. identifying/flagging suspicious emails
    c. Simulating attacks -> identify weaknesses
    d. Analyzing incident-related data -> respond quickly
- AI for hackers:
    a. Social engineering schemes: automating/personalizing scams/phishing
    b. Password hacking: algorithm to decipher passwords quickly
    c. Deepfakes: creating/manipulating video/audio to mimic people (can be used with other types)
    d. Data poisoning: alter training data of AI (bad inputs -> bad outputs) difficult to detect

Source: AI in cybersecurity. (excelsior university)

- Benefits:
    a. Improved threat detection
    b. Responses: (sending alerts, isolating compromised devices, blocking malicious traffic
    c. Machine learning: improving defenses through experience (past incidents -> threat detection) EX: weak/guessable password alerting
    d. Less human error (typos/formatting mistakes)
    e. Automated processes: continuous monitoring, patch management
- Challenges:
    a. No human judgment: or creativity: (ethical issues: privacy invasion, misused sensitive info, bias -> discriminatory outcomes (complacency)
    b. Struggle to keep up with threats
    c. False positives: alert incorrectly indicating vulnerability -> AI gets overwhelmed -> IT teams don't know real/fake threats
    d. Cost: special hardware/software and training. (especially for small business)
    e. Unemployment: Ai replaces IT professionals (EX: customer service)

Source: CISA.gov

CISA's roadmap: (cybersecurity and infrastructure security agency)

- GOALS:
    a. Cyber defense: AI tools to defend cyberspace
    b. Risk reduction/resilience: CISA supports risk-aware/responsible adoption of AI

    c. Operational collaboration: CISA communicates important threat info to public and organization
    d. Agency unification: integrate AI software into agency + create workforce for handling AI software
- LINES OF EFFORT:
    a. Responsibly use AI: ethical and safe use: in accordance with constitution
    b. Assure ai systems: assist in AI software adoption for government agencies and private organizations through guidance + best practices developed
    c. Protect important infrastructure (with AI): assist/recommend mitigation of AI threat
    d. Collaborate with interagency, international partners, the public: develop approaches/policies/ national strategy on AI (with partners abroad)
    e. Educate workforce on AI: education on software/techniques + recruiting of workforce + training has legal/ethical consideration

Source: top AI cybersecurity companies
- Crowdstrike: AI-native defense: lots of services/products (endpoint protection/threat intelligence)
- Zscaler: forefront of zero-trust security: new AI security controls
- Fortinet: forefront of cybersecurity (merge: network and security)
- Palo Alto networks: many services: cortex: AI security platform

Source: AI in cybersecurity

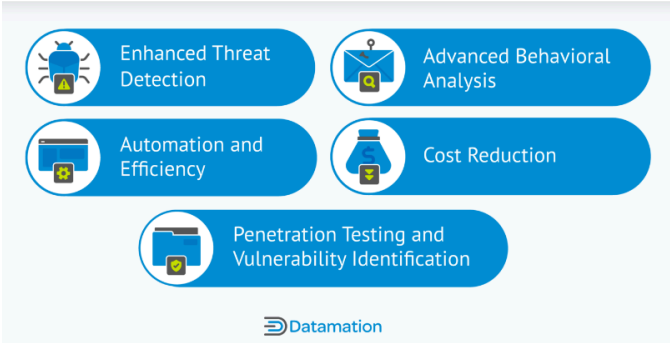
	Traditional cybersecurity	AI cybersecurity
Methodology	Rule-based approaches	Machine learning/AI algorithms
Threat detection	Signature-based detection	Anomaly detection/behavioral analysis
Adaptability	Limited adaptability to threat	Real time evolving
Response time	Manual response -> slow	Automatic -> fast
Human involvement	High	Low (automation)
False positives	Higher rates	Lower rates
Predictive capabilities	Limited	Advanced proactive defense

Traditional cybersecurity:
- Rule based method: known signatures/pre defined rules used to identify/block threats -> detection method: signatured-based: matching data to known attack signatures -> no adaption -> slower response (intervention) + possible human error
- Requires regular updates: if not, solution don't reflect current threats -> false positive
- No predictive ability

Impact of AI in cybersecurity
- Less recovery time
- Automation Frees up time for more complex tasks:
- Development of AI: benefits cybercriminals (evasion/attacks) and organizations

Benefits of AI:

- Behavioral analysis: monitor user behavior and deviations from normal patterns (insider threats/ account takeovers) -> analyzing login attempts + verifying users through behavioral data
- Cost reduction: automation -> time/resource saved. Less resource wasted (investigating false positives, overlooking genuine incidents)

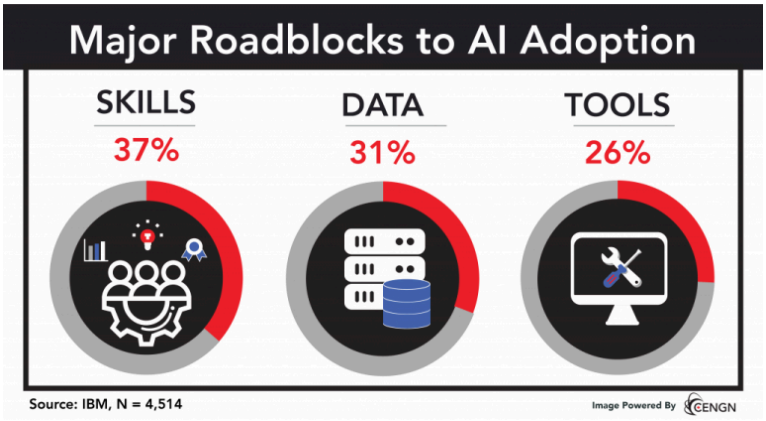


Challenges of AI:

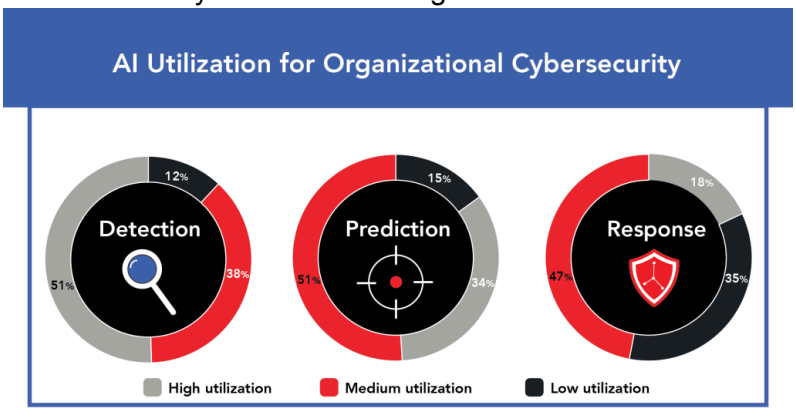
- Bias and ethics: bad training data -> bias (ETHICS: privacy, dual-use of AI, human autonomy)
- Data manipulation: bad data is put in on purpose for bad results (AI can be manipulated and used against companies) hackers can use functions of software against
- Misinterpretation: false/incomplete data -> AI gives false positives/threats overlooked -> block genuine operation/authorized users
- Regulations: AI in cybersecurity has many regulations from government -> requires thorough understanding (regulations change often too) + limits what AI/companies can do to protect

Source: AI in cyber security

- 3.86 million: global/average company's total cost per data breach (2020)
- Security automation (fully deployed) saved 3.58 million (compared to no security automation) security automation 2.45 mil < no automation 6 mil
- Cisco: 7 trillion threats blocked for customers (2018)
- +60% companies: AI threats from cyber criminals -> most relevant
- Fuzzing method by hackers: inject bad inputs -> bad outputs -> take advantage of outputs (exposes vulnerability)
- BENEFITS TO AI:
    a. Threat hunting: actively searching for possible threats
    b. Vulnerability management:
    c. Network security: AI creates groupings of workloads/policies based on specific needs



- Considerable time and money investment
- Hard to find people with expertise in AI (usually with companies still exploring AI)
- AI development people: data complexity, tools available = main issues
- Data complexity: keeping up with data flow -> sourcing malware/code forAI training
- Tools: cyber solutions integrate into network



- +50% companies: AI highly used for detection
- 35% companies: AI highly used for prediction (second highest)
- 18% companies: AI highly used for response

PLANNING

Introduction: 100 words

a. what is cybersecurity? Why is it important?
b. Intro to AI in cybersecurity
c. Why is AI needed in cybersecurity?

Body 1: 300 words

1. Traditional cybersecurity VS AI cybersecurity

A-benefits

2. AI detection capability
3. AI prediction capability
4. AI response capability
5. Impacts of AI

B-disadvantage

6. Bias and ethical implication
7. Misinterpretation and unemployment
8. Roadblocks to AI

Body 2: 300 words

1. AI for hackers (methods hackers use)
2. CISA goals
3. CISA lines of effort

Conclusion: 100 words

1. How can you protect yourself (what to watch out for) ect.
2. Real examples: events, AI software, AI cybersecurity companies,
3. summary/conclusion

SOURCES

Costa, Estevao. "Artificial Intelligence in Cybersecurity: The Benefits and Challenges." *CENGN*, 18 Aug. 2021, www.cengn.ca/information-centre/innovation/artificial-intelligence-in-cybersecurity-the-benefits-and-challenges/. Accessed 19 Dec. 2024.

Council of Europe. "The Framework Convention on Artificial Intelligence." Council of Europe, 2024, www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence. Accessed 4 Jan. 2025.

---. "What Is Cybersecurity?" *Cybersecurity and Infrastructure Security Agency*, 1 Feb. 2021, www.cisa.gov/news-events/news/what-cybersecurity. Accessed 19 Dec. 2024.

Fox, Jacob. "Top 40 AI Cybersecurity Statistics." *Cobalt*, 10 Oct. 2024, www.cobalt.io/blog/top-40-ai-cybersecurity-statistics. Accessed 4 Jan. 2025.

Grace, Meeba. "Top 8 AI Cybersecurity Companies – Tips to Choose the Best Company." *Sprinto*, 10 Oct. 2024, sprinto.com/blog/ai-cybersecurity-companies/. Accessed 19 Dec. 2024.

Kerwin, Jenna. "What Is the Role of AI in Cybersecurity?" *Excelsior University*, 1 Aug. 2024, www.excelsior.edu/article/ai-in-cybersecurity/. Accessed 19 Dec. 2024.

Morgan Stanley. "AI and Cybersecurity: A New Era." *Morgan Stanley*, 11 Sept. 2024, www.morganstanley.com/articles/ai-cybersecurity-new-era. Accessed 19 Dec. 2024.

Precedence Research. "Artificial Intelligence (AI) in Cybersecurity Market Size, Share, and Trends 2024 to 2034." *Precedence Research*, 13 Nov. 2024, www.precedenceresearch.com/artificial-intelligence-in-cybersecurity-market. Accessed 19 Dec. 2024.

Ticong, Liz. "AI in Cybersecurity: The Comprehensive Guide to Modern Security." *Datamation*, TechnologyAdvice, 29 Apr. 2024, www.datamation.com/security/ai-in-cybersecurity/. Accessed 19 Dec. 2024.

University of Illinois Chicago. "What Is (AI) Artificial Intelligence?" *University of Illinois Chicago*, Board of Trustees of the U of Illinois, 7 May 2024, meng.uic.edu/news-stories/ai-artificial-intelligence-what-is-the-definition-of-ai-and-how-does-ai-work/. Accessed 19 Dec. 2024.

Wheeler, Kitty. "Top 10: Cybersecurity Companies." *Technology Magazine*, 16 Oct. 2024, technologymagazine.com/articles/top-10-cybersecurity-companies. Accessed 19 Dec. 2024.