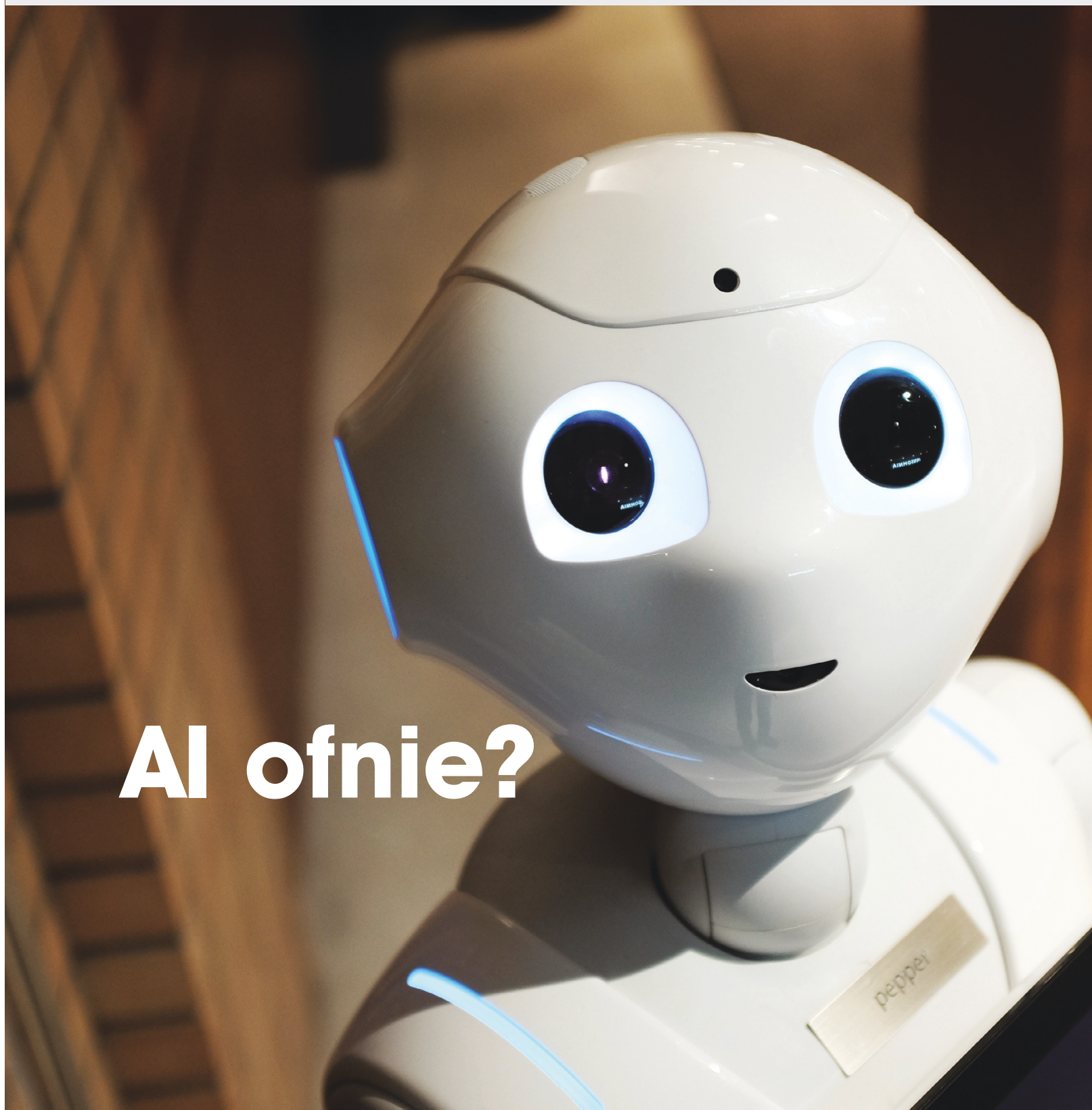




**Auteur:** Dré Lameir is CISO en CTO bij Enovation Group en is bereikbaar via [linkedin.com/in/lameir](https://www.linkedin.com/in/lameir).



# AI of nie?



Sinds medio 2022 valt iedereen over elkaar om over de impact van Artificial Intelligence (AI) te praten. Artikelen, podcasts, mainstream media en vooral Twitter zijn er vol van. Zelf roep ik al een tijdje dat we over een paar jaar de ene AI tegen de andere AI gaan zien 'battelen' in cybersecurity. Ik stel me zo voor dat een CISO van een groot bedrijf aan het eind van de maand een mooi PowerBI dashboardje heeft waarop staat:

#### Dec 2025

- Cyber attacks:
  - o Lost: 4
  - o Won: 36
  - o Paid: 17 (Total amount BTC 4,5643)
- Excuted by:
  - o Human: 13 (22,8%)
  - o AI: 41 (71,9%)
  - o Unknown: 3 (5,3%)

Het gaat nu echt hard vooruit met kunstmatige intelligentie. Ik denk niet dat ik het nog hoeft uit te leggen, maar toch ... we hebben het hier over beelden creëren met DALL-E 2, Midjourney, et cetera Teksten schrijven met Lex en chatten met OpenAI.com waar de engine GTP-3 onder ligt. Ik heb zelf een beetje gespeeld met ChatGTP (op chat.openai.com) en net als miljoenen anderen voel ik dat de resultaten verbluffend zijn, er staat iets te gebeuren.

Vroeger wilde ik altijd naar mijn ome Piet om daar Pong te kunnen spelen: "Wow! Ik kan de televisie besturen!" Nu voel ik dezelfde 'wow!' bij het gebruiken van deze tools. ChatGTP is minder leuk dan een potje Pong, maar leidt zeker tot verwondering.

#### Een voorbeeld:

**Invoer:** What is AI? How could it be used in a cyberattack? Can we write a poem about it?

**Respons:**

Artificial Intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think and act like humans. These intelligent machines can be trained to perform a variety of tasks, such as recognizing patterns, making decisions, and solving problems. AI has the potential to revolutionize many industries and has already begun to transform fields such as healthcare, transportation, and manufacturing.

In the context of a cyberattack, AI could be used to automate various aspects of the attack, such as identifying and exploiting vulnerabilities, evading detection, and adapting to countermeasures. For example, an AI-powered malware might be able to modify its own code to bypass security measures or generate new tactics to evade detection.

#### Here is a poem about AI:

"AI, a curious mind,  
A digital being, of a different kind,  
With algorithms for thought, and data for sense,  
Its intelligence grows, with every experience.  
..."

Misschien niet mijn beste stukje proza... maar het kostte me minder dan vijf seconden om te fabriceren. Stel je voor dat we iets invoeren dat we eenvoudig zouden kunnen vinden op openbare bronnen als Facebook? Een beetje AI meets Open Source Intelligence (OSINT) zullen we maar zeggen.

**Invoer:** Stel een vraag aan Ome Piet die glaskunst maakt en een vintage gameconsole heeft met Pong erop.

**Respons:** Ome Piet, hoe ben je begonnen met het maken van glaskunst en wat trekt je het meest aan in dit medium? Heb je ook ervaring met andere soorten kunst, zoals schilderen of tekenen? Wat vond je het leukste aan het spelen van Pong op je vintage gameconsole? Zou je ooit een andere oude game willen proberen?

Met wat gestoei kan je een vlot gesprek simuleren. Kijk hier eens naar...

**Invoer:** Ome Piet antwoord: Die oude Pong game werkt niet meer hoor!

**Respons:** Sorry dat te horen. Het kan frustrerend zijn als een apparaat of spel niet langer werkt zoals verwacht. Zou je interesse hebben in het proberen om de Pong game te repareren of misschien zelfs te modificeren? Of ben je meer geïnteresseerd in het spelen van andere vintage games op je console? Ik ben benieuwd naar je ervaringen met het verzamelen en spelen van oude games.

Dit is natuurlijk geen conversatie waar Ome Piet in zou trappen. Je hoort bijna een computerstem in je hoofd als je het leest. Ik stuur een beetje bij...

**Invoer:** Nog veel informeler, amicaler zoals een puber zou doen.

**Respons:** Hé Ome Piet, jammer dat die oude Pong game niet meer werkt. Ben je geïnteresseerd in het proberen om het te repareren of heb je misschien zin om andere oude games te spelen op je console? Dat lijkt me zo cool! Is er een game in het bijzonder waar je naar uitkijkt om te spelen?

Hmmm, veel beter. Nu is het zo dat je eenvoudig met wat scripting tegen deze ChatGTP aan kunt programmeren, er is gewoon een API

## AI ofnie?

beschikbaar. Het is dus goed denkbaar dat je op grote schaal 'nep gesprekjes' kunt voeren. Eenmaal 'beet' kan de uiteindelijke phisher het gesprek overnemen en de buit binnenhalen. ChatGPT gaat zelfs door op voorgaande responses en kan daar slim op variëren. Voor nu vergezocht, maar met de juiste context en wat extra informatie trappen er misschien wel mensen in deze vorm van communiceren.

Echt van nep onderscheiden wordt steeds moeilijker. De eerdergenoemde AI's die beelden genereren op basis van prompts (invoer) kunnen daarbij ook invloeden van kunstenaars meenemen. Er is een Zweedse kunstschilder (Simon Stålenhag) die al bezwaar heeft gemaakt tegen de output, het leek gewoon sprekend op zijn werk. Met tekst werkt het net zo, je kunt ChatGPT vragen iets in de stijl van Shakespeare te schrijven.

Als kunstenaars en schrijvers moeten gaan vechten om de rechten op intellectueel eigendom met een AI dan weet je dat we op onze hoede moeten zijn. Tot nu toe was een van onze grootste wapens tegen cybercriminaliteit dat je met wat training, oppassen en controleren snel 'fake' kon herkennen. Vaak mist gegenereerd materiaal een ziel. Die verdediging lijkt nu voorgoed van tafel.

En we zijn niet de enige beroepsgroep die gaat worstelen. De eerste afstudeerverslagen die door een computer zijn geschreven, zijn al gesignaleerd. "Nakijken is onzin vanaf nu!", roepen sommige leraren. In Amerika heeft ChatGPT het advocaat-examen met goed gevolg afgelegd en Getty Images verbiedt AI gegenereerde kunst. Bij het schrijven van software krijgen engineers hulp in Github van CoPilot. Ook hier schrijft ChatGPT mee. Wie is er aansprakelijk als dat tot fouten leidt? Van wie is de code? Moeten we dit juist meenemen in onderwijs en gaan omarmen, zoals we dat ook met industriële robots en 3D printers doen? Moeten we meer nadruk gaan leggen op correct, veilig en ethisch gebruik van de tools?

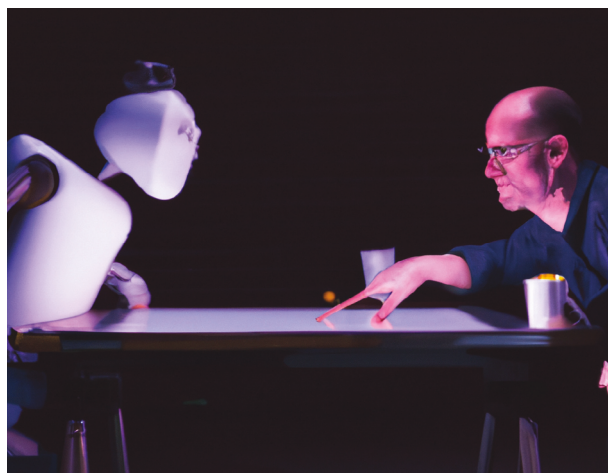
Maar goed, AI gedreven phishing bots. Mijn ome Piet is helaas al een paar jaar dood en trapt er dus niet meer in. Voor de rest van ons hier nog de tweede strofe van het gedicht waarop de chatbot me trakteerde:

"...  
*It can learn and adapt, at an alarming pace,  
 A formidable force, in the digital space,  
 But with great power comes great responsibility,  
 For AI, a potential for both good and malevolency."*

Een heel goed nieuwjaar en ... sterkte allemaal in 2023.

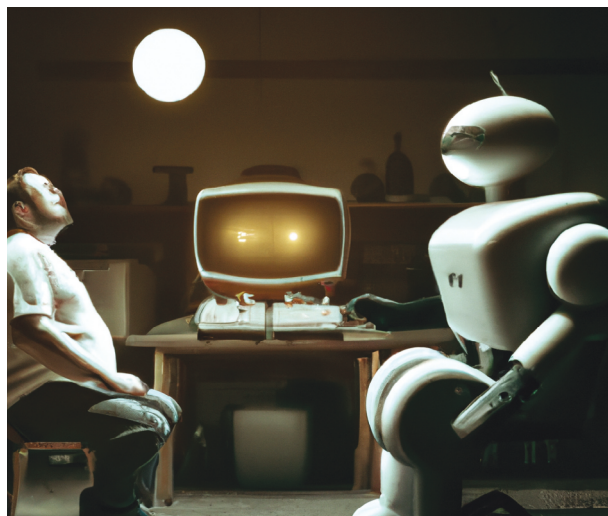
**Invoer:** A game of pong played between a robot and a older man in a dark lit room. The style is photo realistic. They are looking at a screen where the game is visible.

### Respons DALL-E 2:



**Invoer:** A futuristic robot that resembles a man is playing a computer game of classic pong against a male older bold chubby human. They look directly at the screen where the game of PONG is visible. The room is dark lit and the style of the image is like a photo by Annie Leibovitz. Both are holding a joystick.

### Respons DALL-E 2:



### Referenties

CoPilot: <https://github.com/features/copilot>

DALL-E 2: <https://openai.com/dall-e-2/>

ChatGPT: <https://chat.openai.com/chat#>

Midjourney: <https://midjourney.com>

Lex: <https://lex.page/>

OSINT: Open Source Intelligence ([https://en.wikipedia.org/wiki/Open-source\\_intelligence](https://en.wikipedia.org/wiki/Open-source_intelligence))

Simon Stålenhag geïmiteerd: <https://www.wired.com/story/artists-rage-against-machines-that-mimic-their-work/>