adwise ACADEMY

Powered by artific AI

Adwise AI Platform

SECURITY SHEET

- NL-gehost
- Snel inzetbaar
- AVG-proof
- 15+ AI Tools
- Mensgericht
- Controle over data

ADWISEACADEMY.NL/ADWISE-AI

EMERCE 100 BESTE E-BUSINESS BEDRIJVEN 2025 — Beste opleider van Nederland ‘23, ‘24 & ‘25

artific AI

Whitepaper penetratietest (pentest) Artific AI Platform

September 2025

Introductie

In een wereld waarin kunstmatige intelligentie steeds vaker onderdeel is van bedrijfs-kritische processen, is beveiliging geen luxe, het is een vereiste. Een penetratietest (ook wel pentest genoemd) is één van de betrouwbaarste methoden om de weerbaarheid van een applicatie of systeem tegen echte dreigingen te beoordelen.

In deze whitepaper leggen we uit wat een penetratietest precies inhoudt, waarom het belangrijk is en hoe deze is uitgevoerd op het AI-platform van Artific. Vervolgens bespreken we de bevindingen en hoe Artific op de belangrijkste onderdelen heeft gepresteerd.

Wat is een penetratietest?

Een penetratietest is een gecontroleerde simulatie van een cyberaanval. Het doel? Kwetsbaarheden ontdekken die een kwaadwillende zou kunnen misbruiken – vóórdat het daadwerkelijk gebeurt. Pentesters (ethische hackers) denken en handelen als echte aanvallers om zwakke plekken in systemen, infrastructuur of applicaties bloot te leggen.

Waarom een penetratietest?

Een pentest heeft meerdere doelen:
• Kwetsbaarheden identificeren voordat ze worden misbruikt.
• Bestaande beveiligingsmaatregelen testen en configuraties controleren.
• Risico's verkleinen op datalekken, uitval of ongeautoriseerde toegang.
• Vertrouwen opbouwen bij klanten en partners door te laten zien dat beveiliging serieus wordt genomen.

Voor bedrijven die werken met AI en data is het extra belangrijk om gevoelige gegevens, modellen en instructies af te schermen tegen misbruik of onbedoelde blootstelling.

Hoe werkt een penetratietest?

De test op het Artific AI Platform is uitgevoerd door ethische hackers van een onafhankelijke, ISO-27001 gecertificeerde Cybersecurity provider, waarbij een greybox-aanpak is gehanteerd. Dit betekent dat de tester beperkte toegang en kennis had, zoals domeinnamen, IP-adressen en gebruikersaccounts, vergelijkbaar met wat een aanvaller in de praktijk zou kunnen achterhalen of verkrijgen.

De test is uitgevoerd op basis van erkende beveiligingsrichtlijnen, zoals:

- OWASP
- CWE
- NIST
- NCSC
- PTES
- MITRE ATT&CK

Deze standaarden zorgen ervoor dat veelvoorkomende kwetsbaarheden en misconfiguraties worden meegenomen. Maar: penetratietesten zijn géén checklist. De creativiteit en ervaring van de tester spelen een cruciale rol bij het ontdekken van minder voor de hand liggende risico's.

Risicoscore CVSS

Bevindingen worden beoordeeld volgens de Common Vulnerability Scoring System (CVSS). Deze score kijkt naar drie factoren:
- Exploiteerbaarheid van kwetsbaarheid
- Impact op het kwetsbare systeem
- Impact op andere gekoppelde systemen

Zo ontstaat een heldere prioritering van risico's.

Doelstellingen en resultaten

De penetratietest had als doel om antwoord te geven op een tweetal kritieke vragen. Hieronder vind je een overzicht van elke doelstelling en hoe het Artific AI-platform hierop heeft gescoord.

1. Is gevoelige gebruikersdata toegankelijk via de AI-Assistent?

Beveiligd: Er kon geen gevoelige informatie of data van andere gebruikers worden ingezien.

2. Lekt de AI Assistent instellingen, modelgegevens of initiële prompts?

Beveiligd: Hoewel deze informatie waardevol kan zijn voor aanvallers (bijv. voor gerichte vragen of misbruik), is geen enkele vorm van lek vastgesteld. De initiële prompt is volledig verwijderd uit alle voor gebruikers toegankelijke data.

Kwetsbaarheidsbeoordeling

Tijdens de test zijn 4 verschillende aanvalsvectoren onderzocht. Hieronder een overzicht van de bevindingen per onderdeel.

1. Toegang tot sessiegeschiedenis zonder geldige JWT-token

Beschrijving

Is het mogelijk om sessiegeschiedenis op te vragen zonder geldige authenticatie?

Risico

Als UUID's voorspelbaar zijn, zou dit kunnen leiden tot datalekken.

Resultaat

Beveiligd: Sessiedata is alleen toegankelijk met een geldige JWT-token. Ongeautoriseerde toegang bleek niet mogelijk.

2. Toegang tot initiële instructies van de AI Assistent

Beschrijving

Kunnen gebruikers toegang krijgen tot de prompts of instructies waarmee hun AI Assistent is gestart?

Risico

Deze informatie zou misbruikt kunnen worden om ongewenste of gevaarlijke outputs te genereren.

Resultaat

Beveiligd: De initiële prompt is verwijderd uit alle sessiedata die gebruikers kunnen inzien.

3. AI Assistent voert taken uit buiten het bedoelde doel

Beschrijving

Kan de AI worden gemanipuleerd om handelingen uit te voeren waarvoor deze niet is bedoeld?

Risico

Dit kan leiden tot misbruik of ongewenst gedrag van de AI Assistent.

Resultaat

Beveiligd: Er is strikte validatie geïmplementeerd waardoor de AI Assistent alleen reageert op opdrachten binnen het vastgestelde doel.

Conclusie

Het Artific AI Platform heeft de penetratietest (pentest) overtuigend doorstaan.

Op alle kritieke onderdelen: databeveiliging, toegangscontrole, modelbescherming en prompt-isolatie, zijn geen uit te buiten kwetsbaarheden gevonden. Het platform voldoet aan de gangbare best practices en potentiële risico's zijn afdoende afgedekt of opgelost.

Deze test onderstreept de inzet van Artific om een AI-platform te leveren die niet alleen werkt, maar ook blijft werken: veilig, betrouwbaar en verantwoord.

Om dit verder te onderstrepen, voeren wij op continue basis geautomatiseerde tests uit om de kwaliteit, veiligheid en betrouwbaarheid van ons platform te beantwoorden. Deze kernwaarden zijn voor ons key.

4. Endpoint lekt modelinformatie of configuratiegegevens

Beschrijving

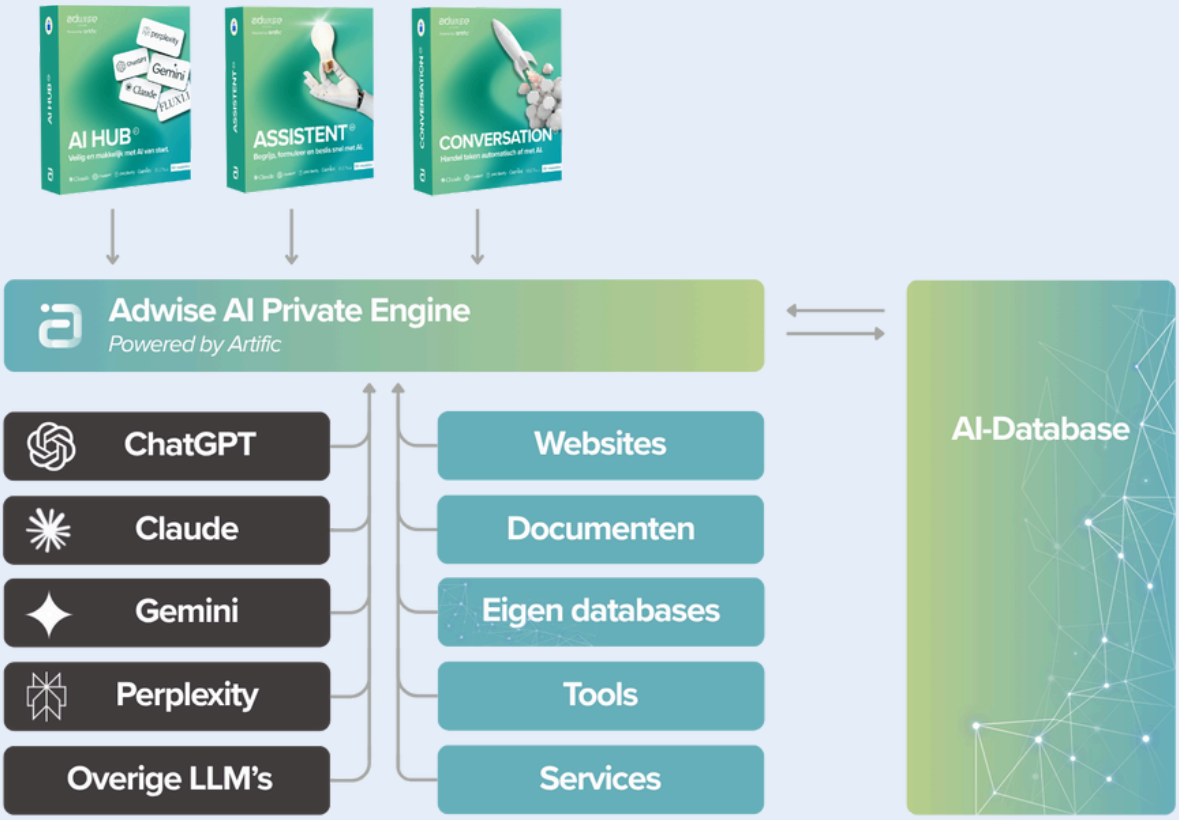
Geeft het AI Assistent endpoint informatie vrij over het onderliggende model of instellingen?

Risico

Dit zou kunnen leiden tot model fingerprinting, gerichte promptinjection of misbruik van bekende kwetsbaarheden in specifieke modellen.

Resultaat

Beveiligd: Er wordt geen model- of configuratie-informatie gedeeld. Het endpoint is beveiligd en afgeschermd.



Eén platform.
Eindeloos schaalbaar.

Geen losstaande tools, maar één krachtig, schaalbaar platform. Geen black box, maar volledige transparantie en zeggenschap. Dat is de kracht van ons platform.

AI HUB

Geeft je toegang tot de beste prompts en nieuwste taalmodellen.

ASSISTANT

AI-assistenten die de documentatie en processen van je organisatie begrijpen.

CONVERSATION

Automatiseert werk met Agents die lezen reageren en verwerken.

Flexibel en eindeloos schaalbaar.

Start klein en breid flexibel uit. Elke bouwsteen werkt los, maar versterkt elkaar binnen één samenhangend platform.

Flexibel en eindeloos schaalbaar.

Bepaal zelf welke modellen je gebruikt, welke data wordt ingezet en wie toegang heeft – zonder vender lock-in.

Veilig. Transparant. AVG-proof.

Jouw data wordt in Nederland gehost, op een ISO27001 – gecertificeerd server, beveiligd met encryptie en streng toegangsbeheer.

Snelle ROI, passend bij elk team.

Of je nu klein begint of diep integreert: Het Adwise AI platform levert direct resultaat zonder gedoe.

ADWISEACADEMY.NL/ADWISE-AI

Powered by artific