SECURITY DAYS

KAI NESUVALDYTAS DI VEIKIA PRIEŠ ĮMONĘ: NUO LENOVO „CHATBOTO“ IKI DARBUOTOJŲ GUDRAVIMO

Jurgita Lapienytė,
Cybernews vyriausioji redaktorė

cybernews®

Šiandien sužinosite:

1. Dėl skuboto DI įrankių diegimo kylančios rizikos
2. Spraga „Lenovo“ pokalbių robote atvėrė duris kibernetinėms atakoms
3. Darbuotojų slapta DI naudojimo praktika – rizika verslui
4. Rekomendacijos verslui – integruotas DI saugumo valdymas

cybernews®

Rizikos:

- Kibernetiniai incidentai
- Duomenų vagystės ir nutekėjimai
- Verslo reputacijos praradimas
- Finansiniai nuostoliai
- Teisinės problemos
- Operaciniai trikdžiai
- Darbuotojų pasitikėjimo ir kultūros prastėjimas

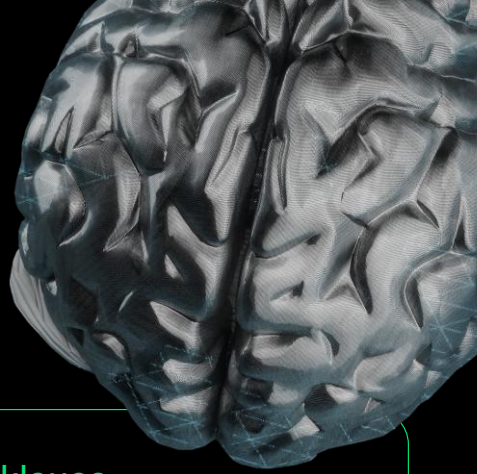
cybernews®

Kai nesuvaldytas DI veikia prieš įmonę:

„Lenovo“ pokalbių robotas „Lena“

cybernews®

Kas nutiko „Lenai”?



Kritinė saugumo spraga – XSS (Cross-Site Scripting) pažeidžiamumas

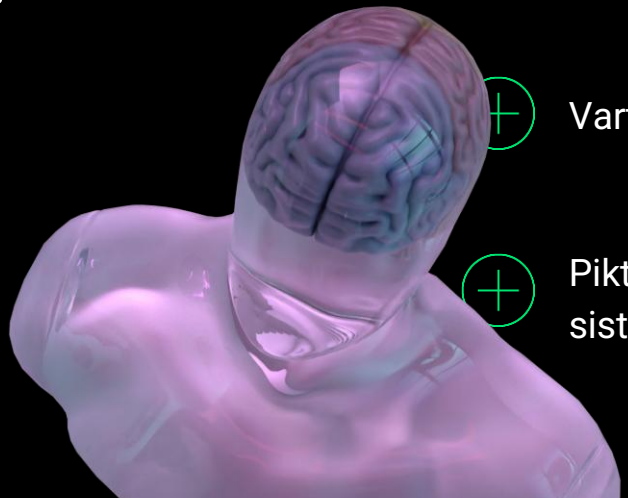
Vagystė įvykdoma įterpiant kenksmingą kodą kaip HTML ir JavaScript

Viena 400 simbolių užklausa leido pavogti aktyvius sesijos slapukus

HTML kodas su kenkėjiškais elementais saugomas „chatboto” pokalbių istorijoje „Lenovo” serveryje ir aktyvuojamas kiekvieną kartą jį atidarius

cybernews®

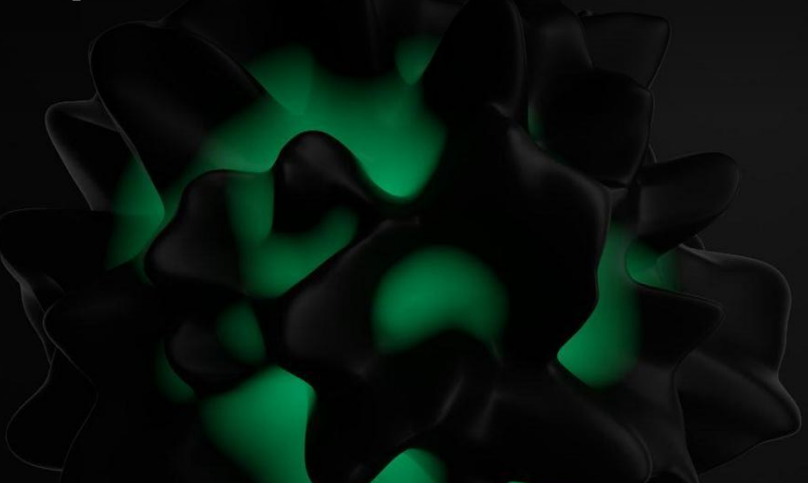
Ką tai reiškia?



- LLM’ai neturi „saugumo instinkto“ – vykdo viską, ko paprašo vartotojas
- „Lena“ neturėjo griežtų įvesties ir išvesties saugumo kontrolės mechanizmų ir monitoringo
- Vartotojų sesijos slapukai galėjo būti pavogti
- Piktavaliai galėjo įsilaužti į vidines „Lenovo“ sistemas

cybernews®

„Lenovo” pamokos ateičiai

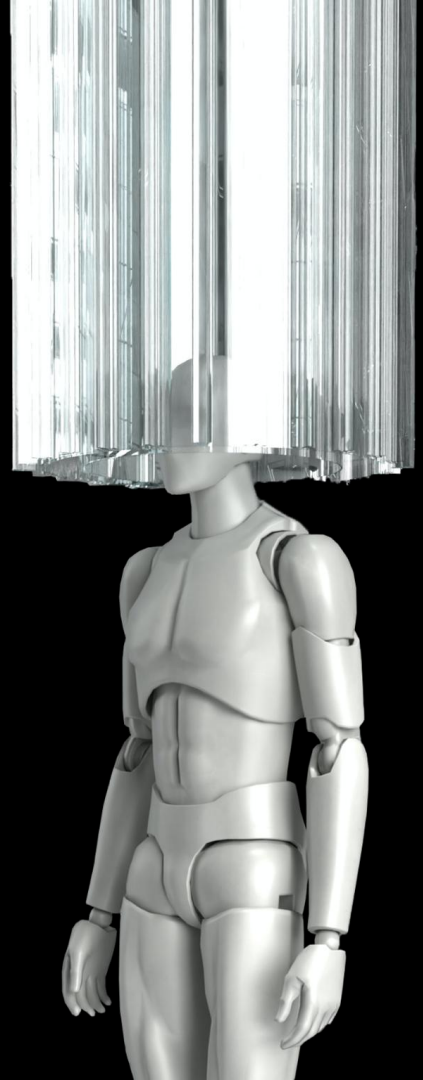


- Griežtos įvesties ir išvesties duomenų kontrolės mechanizmai
- Ribotas leidimas DI generuoti ir vykdyti kodą (pvz., HTML, JavaScript)
- Aktyvių ir nuolatinių stebėjimo priemonių diegimas
- Saugumo kultūros ugdymas
- Reguliarūs saugumo testavimai
- Įrankių vertinimas ir valdymas prieš diegimą
- Atsargumas su pokalbių istorija

cybernews®

Kai nesuvaldytas DI veikia prieš įmonę:

DI įrankių naudojimas darbe



cybernews®

Kodėl darbuotojai naudojasi DI įrankiais?

Verslas neskuba oficialiai patvirtinti DI įrankių naudojimo, bet...
Darbuotojai renkasi DI dėl to, ko negali suteikti žmonės.

Dažniausiai pasirenkamos priežastys:

- **42%** – DI yra visada pasiekiamas
- **30%** – mašininis greitis ir kokybė
- **28%** – nesibaigiantis naujų, kūrybiškų idėjų srautas
- **23%** – neribotos galimybės sprendžiant sudėtingas užduotis

Kartais pasirenkamos priežastys:

- **17%** – DI nektritikuoja
- **16%** – DI nesinervina ir visada kantrus
- **15%** – DI reikia mažiau aiškinimų ir mokymų
- **8%** – galima priskirti darbo rezultatus sau

Šaltinis: 2025 Microsoft Work Trend Index Annual Report

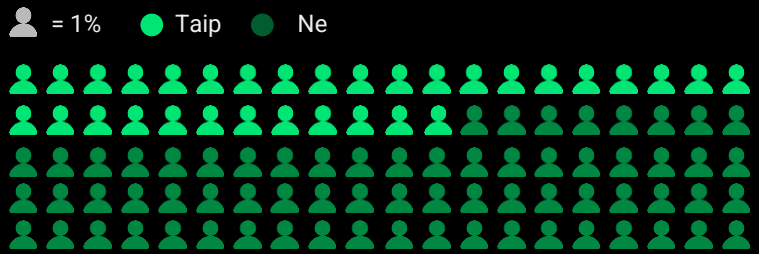
cybernews®

Slaptas naudojimasis DI: priežastys

> **32%** darbuotojų slepia, kad jų produktyvumas padidėjo dėl dirbtinio intelekto įrankių naudojimo.
>
> *Ivanti, 2025*

Beveik 1 iš 3 darbuotojų slepia dirbtinio intelekto naudojimą.

Klausimas: Ar slepiate DI įrankių naudojimą nuo darbdavio?



Kodėl slapta?

- 36% darbuotojų patinka „slaptas pranašumas“
- 30% nerimauja, kad jų darbo vietą gali panaikinti
- 27% jaučia DI sukeliamą menkavertiškumo jausmą: „Nenoriu, kad žmonės kvestionuotų mano gebėjimus“

Šaltinis: 2025 Technology at Work Report (Ivanti)

cybernews®

Nesaugus naudojimasis DI įrankiais darbe: nukentėjusios įmonės

Disney

2024

- „Disney" darbuotojas atsisiuntė nepatikrintą DI įrankį iš „GitHub", kuriame buvo paslėpta kenkėjiška programinė įranga
- Dėl to buvo nulaužti „Disney" „Slack" kanalai ir nutekėjo apie 44 milijonai vidinių žinučių

SAMSUNG

2023

- Darbuotojai pateikė konfidencialią informaciją „ChatGPT", spręsdami programinio kodo problemas ir ruošdami susitikimų santraukas
- Dėl to netyčia nutekėjo įmonės programinis kodas ir konfidencialūs vidiniai duomenys, todėl „Samsung" uždraudė generatyvius DI įrankius savo sistemose

comet

2025

- „Guardio" testavo naršyklę „Comet", sukūrę sukčiavimo scenarijų. „Comet" pakliuvo į keletą apgaulės spąstų:
  - Automatiškai nupirko „Apple Watch" iš suklastotos internetinės parduotuvės
  - Fišingo el. laišką iš netikro banko pažymėjo kaip svarbų ir automatiškai paspaudė ant suklastotos nuorodos
  - Pasitelkus užklausų įterpimo („prompt injection") metodą ir kenkėjišką „captcha" mechanizmą, privertė DI agentą vykdyti kenkėjišką veiklą, pavyzdžiui, parsisiųsti failus ar siųsti jautrią informaciją be vartotojo žinios

Duomenų nutekėjimas šiais atvejais – jautrios informacijos perdavimas nepatvirtintai trečiajai šaliai (pvz., DI įrankiui), net jei ši informacija nebuvo paviešinta ar panaudota.

cybernews®

DI įrankių privatumo spragos: nutekėję vartotojų duomenys

Grok

2025

- Šimtai tūkstančių „Grok“ vartotojų pokalbių tapo viešai prieinami ir išindeksuoti paieškos sistemose
- „Google“ indeksavo beveik 300 000 „Grok“ pokalbių

ChatGPT

2025

- „Google“ indeksavo tūkstančius bendrinamų „ChatGPT“ pokalbių, todėl jie tapo viešai prieinami
- Tyrimas: vartotojai dažnai su „ChatGPT“ dalijasi privačia, jautria ir netgi rizikinga informacija

Šie nutekėjimai įvyko dėl to, kad vartotojai paspausdavo „Share“ (liet. „Dalintis“) mygtuką, kuris sukurdavo viešai prieinamą URL nuorodą.

cybernews®

Rekomendacijos verslui:

kaip darbuotojai turėtų naudotis DI

- Aiškiai apibrėžkite DI naudojimo politiką
- Ribokite jautrios ir konfidencialios informacijos dalijimąsi
- Naudokite tik įmonės patvirtintus DI įrankius
- Įdiekite ir palaikykite saugumo technologijas
- Reguliarūs mokymai ir saugumo kultūros stiprinimas
- Reguliari DI sistemų audito ir rizikų vertinimo praktika
- Skatinkite skaidrumą ir atvirus pokalbius apie DI naudojimą

cybernews®

Jurgita Lapienytė
„Cybernews" vyriausioji redaktorė

Ačiū.

Klauskite!