



AI NT vertintojo darbe:

ką galima, ko negalima ir kas atsako už klaidą?

Dr. Neringa Gaubienė

AI reguliavimas ir valdysena



Apie mane



**Dr. Neringa
Gaubienė**

Ekspertė ir verslo konsultantė

AI reguliavimo, skaitmeninių technologijų teisės ir kibernetinio saugumo srityse

Tarptautinė patirtis

Europos Komisijos AI biuro ekspertų grupės narė

Akademinei veikla

Profesoriaus asistentė VU, dėsto kursą „Teisė ir skaitmeninės technologijos“

Konferencijų pranešėja

Reguliariai dalyvauja konferencijose Lietuvoje ir užsienyje

**AI reguliavimas ir valdysena – Nr. 1 Lietuvoje | Teisinių paslaugų rinkoje – Influencerė
Nr. 1 | IT & Tech įtakingiausios moterys Lietuvoje – Nr. 2**

Kas **nenaudojate AI savo darbinėje
veikloje?**

AI nebėra eksperimentas, bet jo patikimumas vis dar kelia klausimų

54 %

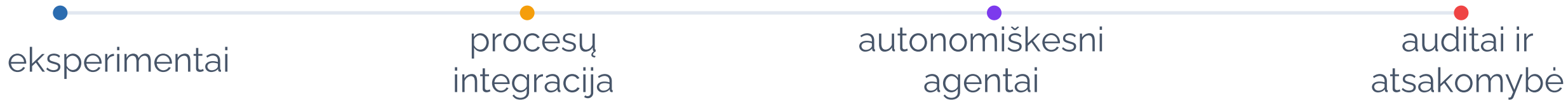
naudoja AI darbinėje
veikloje (2025)

69 %

AI Lietuvoje naudoja 7 iš
10 šalies gyventojų
(2025)

70 %

respondentų laiko, kad
AI vis dar yra
nepatikimas (2025)



Klausimas jau ne „ar naudoti AI“, o „kaip naudoti taip, kad būtų patikima“

Perkeliame diskusiją nuo technologijos žavesio prie saugaus profesinio naudojimo

1

Suprasti

ką AI gali daryti NT
vertinime ir kur jo ribos

2

Atskirti

leistinus pagalbinius
veiksmus nuo neleistino
galutinio sprendimo
„delegavimo“

3

Patikrinti

duomenis, prielaidas ir
skaičiavimus

4

Pagrįsti

kodėl priimtas sprendimas
yra teisingas

Turite gebėti atsakyti į klausimus: kur AI padeda, kur privaloma sustoti, kaip tikrinti sugeneruotą turinį ir kaip dokumentuoti procesus.

Ką tai reiškia NT vertintojui?



AI gali padėti iki sprendimo. Sprendimo negalima palikti modeliui.

Ar skaitėte DI naudojimo instrukciją?

Kaip tinkamai naudotis DI?

1. Aiškiai suformuluok užduotį



2. Pateik daugiau konteksto

Žiemos sezono idėjos:

- Florencija, Roma,
- kalnuose, pavaziuoti per savėitę.



3. Patikrink ir redaguok rezultatus



4. Užduok papildomų klausimų

Pasakyk daugiau apie Dolomitams!

Žinoma! Dolomitai – tai įspūdingi kalnų masyvai....



DI sugeneruotas paveikslėlis

Svarbu: AI prognozuoja labiausiai tikėtiną atsakymą, tačiau AI nemąsto

AI negali garantuoti absoliutaus tikslumo ar absoliutaus saugumo.

Kas vyksta modelyje?

Modelis iš įvesto teksto prognozuoja, koks žodis ar sakinyš tikėtinai turėtų sekti toliau.

Ką tai reiškia praktikoje?

Įtinkamas tekstas gali būti klaidingas. Sklandumas nėra patikimumo įrodymas.



Duomenų apsauga neturėtų būti stabdis, bet dažnai yra

Saugumas prasideda nuo klausimų, kuriuos užduodame prieš diegiant ar kuriant DI sistemas

„Varnelės“ uždėjimas
nėra saugumas



*„Duomenys nebus naudojami
modelio mokymui“.*

Shadow AI



*„Shadow AI - tai darbuotojų
savarankiškai naudojami AI
įrankiai be organizacijos žinios,
leidimo ar kontrolės.*

100 % patikimas
nuasmeninimas
nėra pasiekiamas



Kai kuriamas saugumo
įspūdis, bet ne rizikos
valdymas.

Duomenų taisyklė: ką saugu kelti į AI?

Kaip kalbame apie saugų AI naudojimą, pirmas klausimas yra ne kaip teisingai suformuoti užklausą, o kokius duomenis naudosiu.



ŽALIA

vieši duomenys, nuasmeninti pavyzdžiai, bendri klausimai

galima naudoti su patikra



GELTONA

kliento pateikti dokumentai, adresai, nuotraukos, sutartys

tik patvirtintu įrankiu ir su aiškiu pagrindu



RAUDONA

asmens kodai, banko info, konfidencialūs sandoriai, neanoniminės bylos

nekelti į viešus įrankius

Prieš keldami informaciją į AI, įvertinkite duomenų kategoriją ir įrankio tipą — viešas, įmonės valdomas, sutartinis / uždaras.



"Jeigu gauni prekę nemokamai, tikėtina, kad tu pats esi prekę".

Volstrito taisyklė AI įrankių kontekste

Klasikinės AI sistemų rizikos

1

Netikslumas ir haliucinacijos

AI gali pateikti įtikinamą, bet klaidingą atsakymą.

2

Šališkumas ir diskriminacija

Rezultatai gali atkartoti duomenyse esančius neatitikimus ar stereotipus.

3

“Blac box”

Gali būti sunku paaiškinti, kodėl AI sistema pateikė konkretų rezultatą.

4

Duomenų ir privatumo pažeidimai

Naudojant AI sistemą gali būti atskleisti jautrūs ar konfidencialūs duomenys.

5

Saugumo ir manipuliavimo rizikos

Sistema gali būti apeinama, manipuluojama arba naudojama nenumatytais tikslais.

6

Atsakomybės išskaidymas

Kyla neaiškumų, kas priima galutinį sprendimą ir atsako už klaidas.

Ką galima daryti su AI?

Tai, ką galima patikrinti, paaiškinti ir dokumentuoti.

- ✓ parengti klausimų sąrašą objektui
- ✓ sutrumpinti ar perrašyti aiškinamąjį tekstą
- ✓ sustruktūrinti duomenis į lentelę
- ✓ rasti neatitikimus dokumentuose
- ✓ parengti ataskaitos juodraštį
- ✓ paaiškinti metodą ne specialistui
- ✓ sugeneruoti rizikų sąrašą
- ✓ patikrinti kalbą, stilių, aiškumą

Kontrolinis klausimas: ar galiu patikimai patikrinti AI atsakymą ir paaiškinti, kodėl jį priėmiau arba atmečiau?

Ko negalima daryti su AI?

Ribojimai prasideda ten, kur prarandama kontrolė ar konfidencialumas.

✗ **Negalima deleguoti galutinės išvados**

✗ **Negalima įkelti nevaldytų duomenų**

asmens, kliento, banko, sandorio ar komercinių paslapčių į viešus įrankius

✗ **Negalima aklai pasikliauti AI sugeneruotu turiniu**

AI pateikta informacija nėra patikimumo įrodymas

✗ **Negalima slėpti reikšmingo AI naudojimo**

kai AI sugeneruotas turinys daro poveikį klientui ar plačiajai visuomenei

Kas atsako už klaidą?

AI naudojimas nepanaikina vertintojo ir organizacijos pareigų, įskaitant ir pareigą atsakyti už sukeltą žalą.

Vertintojas

profesinis sprendimas, metodas,
patikra

Įmonė

vidinė tvarka, įrankiai, prieigos,
mokymai

**KLAIDA
VERTINIME**

AI tiekėjas

sutartinės garantijos, sauga,
paslaugos ribos

Klientas / duomenys

neteisingi pateikti duomenys,
sutarties sąlygos

AI Aktas: kas aktualu NT vertintojui?

AI raštingumas

nuo 2025-02-02: mokyti darbuotojus pagal naudojimo kontekstą

Skaidrumas

nuo 2026-08-02: tam tikrais atvejais aiškiai informuoti apie AI sąveiką ar turinį

Draudžiamos praktikos

manipuliavimas, socialinis vertinimas ir kitos nepriimtinos rizikos praktikos

Aukštos rizikos atvejai

NT vertinimas pats savaime nėra tipinis didelės rizikos atvejis, bet gali sietis su kreditu ar viešomis paslaugomis

Pravartu: turėkite AI naudojimo politiką ir mokymų įrodymą.

Skaidrumas

Ne kiekvienas AI naudojimas turi būti aprašytas, bet reikšmingas naudojimas turi būti atskleistas.

Kada reikia atskleisti AI panaudojimą?

- AI turinys ar išvada pateikiami klientui
- AI turėjo reikšmingą įtaką ataskaitos tekstui ar analizei
- AI naudojamas sąveikai su klientu
- AI sugeneruotas turinys publikuojamas viešai

Ką fiksuoti?

- įrankio pavadinimą ir versiją
- naudojimo tikslą
- duomenų kategoriją
- kas peržiūrėjo ir patvirtino
- ką pakeitė žmogus

Gera formuluotė

„AI įrankis naudotas kaip pagalbinė priemonė teksto struktūrai / santraukai parengti. Galutinė analizė, prielaidos, skaičiavimai ir vertės išvada patikrinti ir patvirtinti vertintojo.“

AI raštingumas – ne mokymas „kaip paspausti mygtuką“

6 žingsnių taisyklė:

1

Kas yra
AI?

2

Kaip
veikia AI?

3

Kokios
yra
rizikos?

4

Kokios
galimybės
?

5

Etiškas AI
kūrimas/naudoji
mas

6

AI teisinis
reguliavimas

Intensyvus ir ilgalaikis AI naudojimas slopina kritinį mąstymą, atmintį ir kūrybiškumą

Updated: JUN 23, 2025 5:16 PM EET

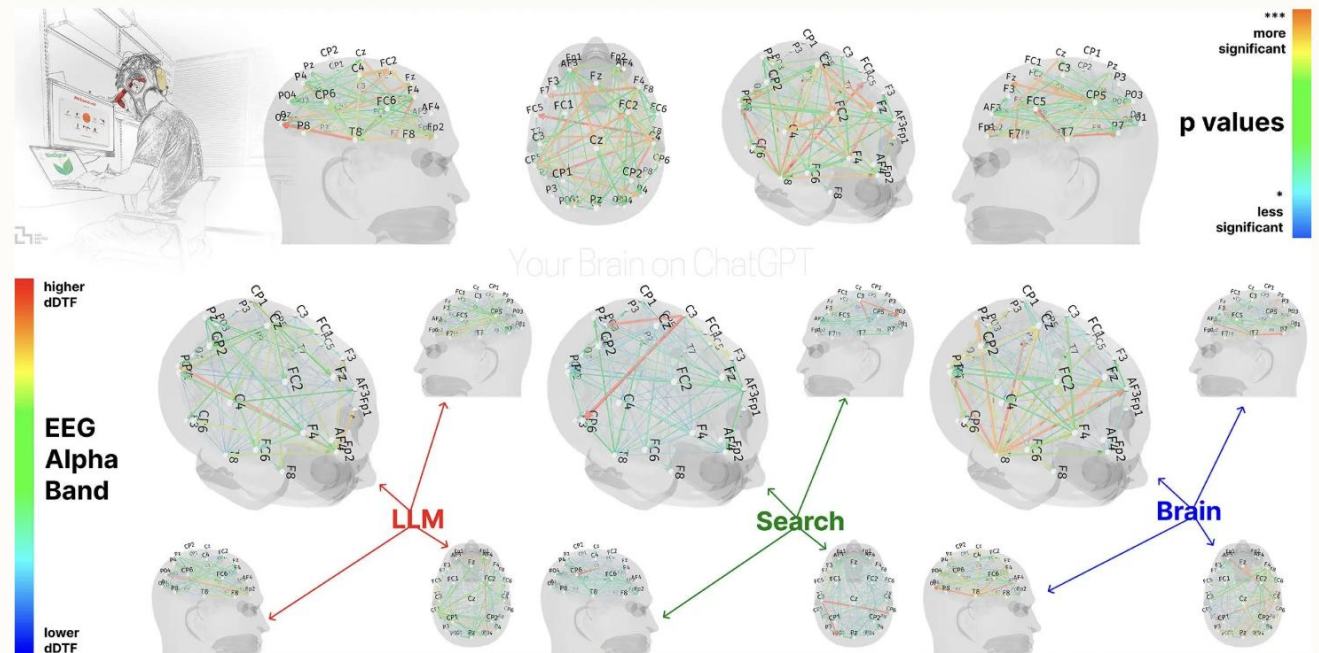
ChatGPT May Be Eroding Critical Thinking Skills, According to a New MIT Study

TECH

ARTIFICIAL INTELLIGENCE



by **Andrew R. Chow**
CORRESPONDENT



Ką verta prisiminti po šio pranešimo:

- 1 AI yra asistentas, ne vertintojas. Sprendimus priimate jūs.
- 2 Niekada nekelkite konfidencialių duomenų į viešą įrankį.
- 3 Kiekvieną AI rezultatą tikrinkite.
- 4 Įtikinamas AI sugeneruotas atsakymas dar nereiškia, kad jis teisingas.
- 5 Atskleiskite reikšmingą AI naudojimą.

Svarbu! Atsakomybė už galutinį vertinimą lieka vertintojui.

Neringa Gaubienė, PhD

Guiding **responsible** AI knowledge, driving success



Ačiū.

Bendraukime

Tel. +370 614 24694

El. p. neringa.gaubiene@tf.vu.lt

