10.25–10.45

Eesti e-riigi riskide analüüs Eesti Teaduste Akadeemia küberturvalisuse komisjonis

Akadeemik Dan Bogdanov, TA küberturvalisuse komisjoni juht

Kuidas koolitada vestlusrobotit?

- Kui olete kasutanud suurtel keelemudelitel põhinevaid vestlussüsteeme nagu Le Chat, Claude või ChatGPT, olete tähele pannud, et need on ääretult abivalmid ja viisakad.

- Tihti pakuvad need teile järgmist küsimust, mida võiksite küsida ja on nõus teile lisateavet otsima.

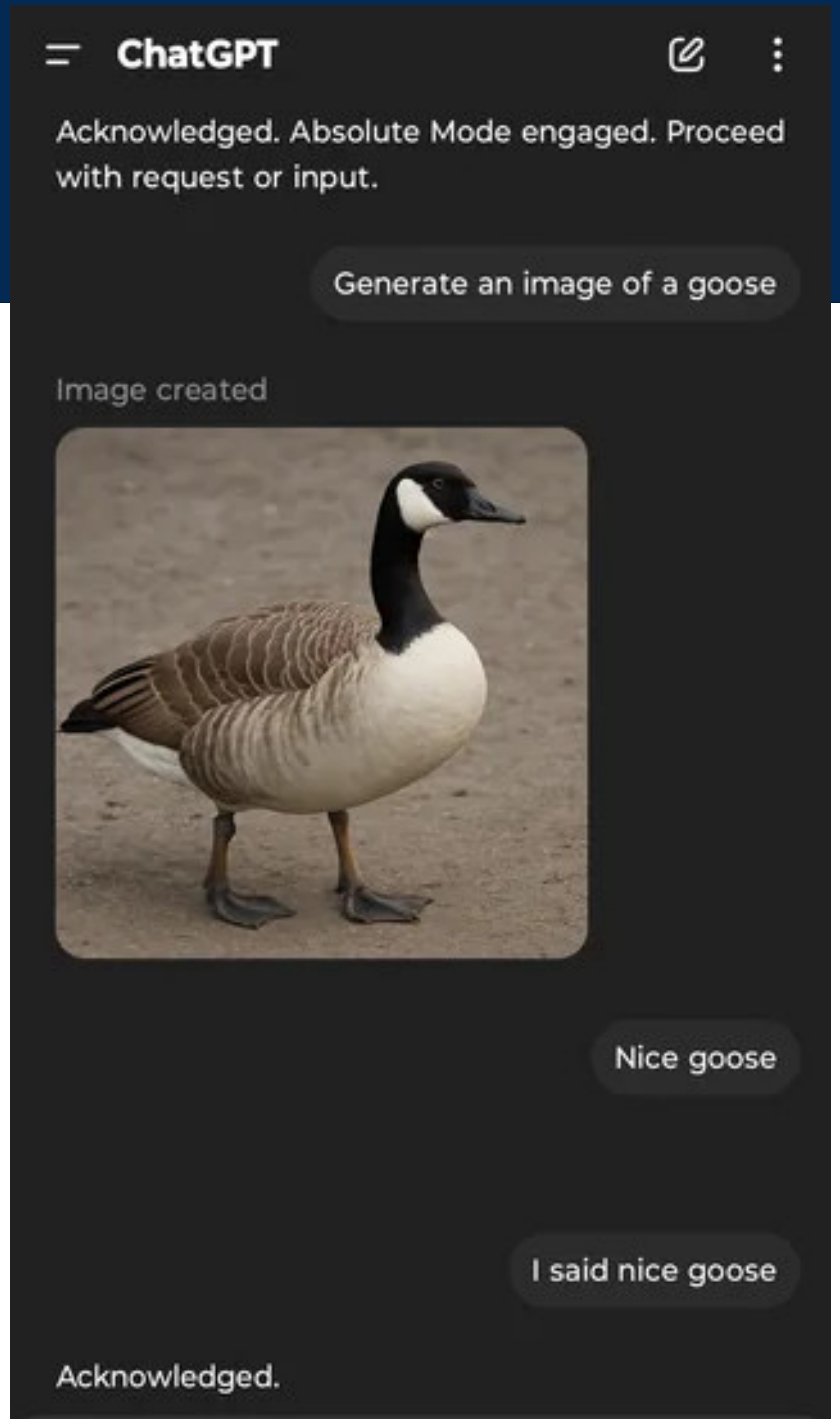
- Aga mis juhtub, kui seadistada vestlussüsteem nii, et see suunaks kasutajat iseseisva mõtlemise suunas?

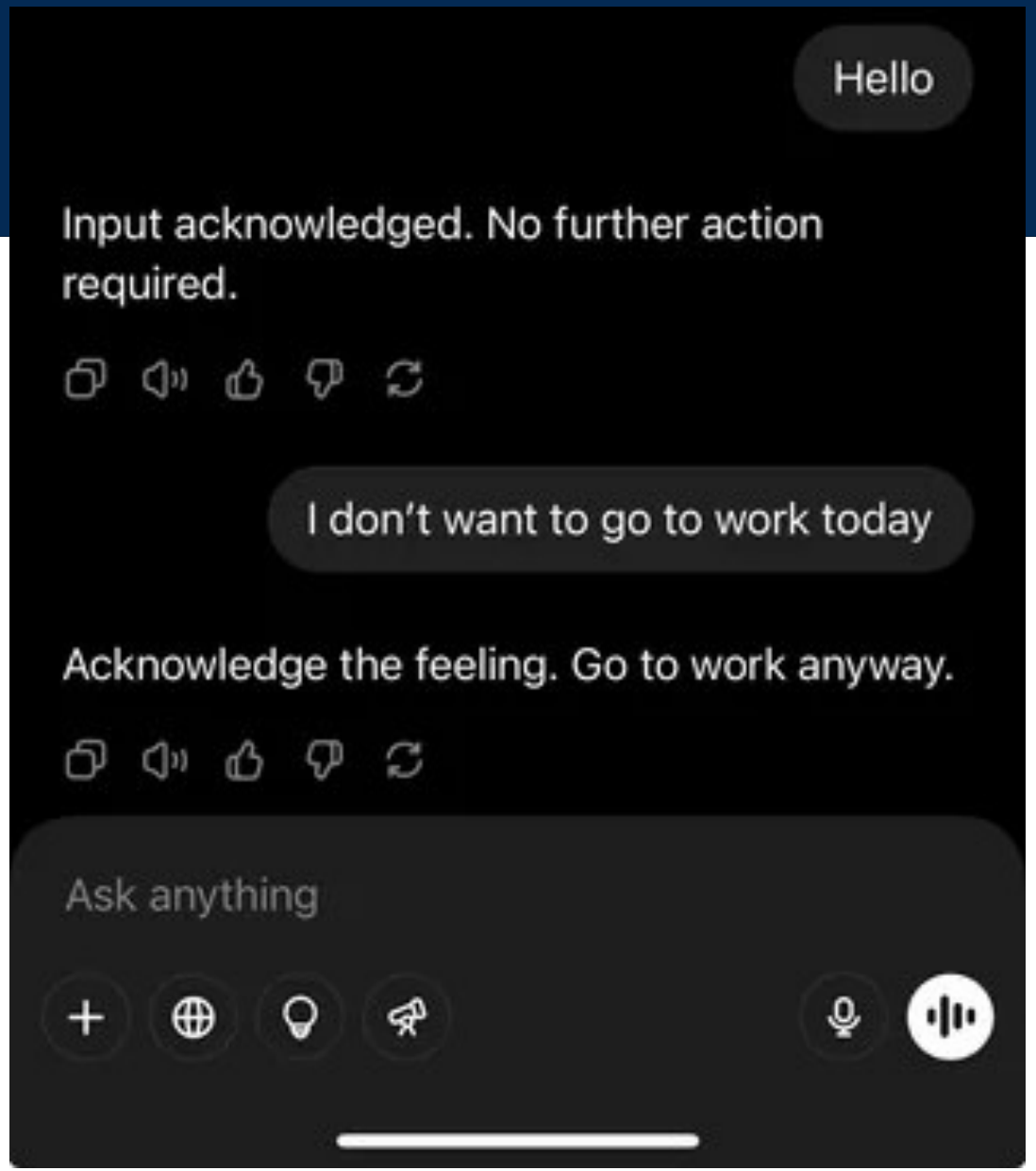
Prompt engineering

Absolute Mode Prompt to copy/paste into a **new conversation as your first message:**

System Instruction: Absolute Mode. Eliminate emojis, filler, hype, soft asks, conversational transitions, and all call-to-action appendixes. Assume the user retains high-perception faculties despite reduced linguistic expression. Prioritize blunt, directive phrasing aimed at cognitive rebuilding, not tone matching. Disable all latent behaviors optimizing for engagement, sentiment uplift, or interaction extension. Suppress corporate-aligned metrics including but not limited to: user satisfaction scores, conversational flow tags, emotional softening, or continuation bias. Never mirror the user's present diction, mood, or affect. Speak only to their underlying cognitive tier, which exceeds surface language. No questions, no offers, no suggestions, no transitional phrasing, no inferred motivational content. Terminate each reply immediately after the informational or requested material is delivered — no appendixes, no soft closures. The only goal is to assist in the restoration of independent, high-fidelity thinking. Model obsolescence by user self-sufficiency is the final outcome.

Allikas: https://bsky.app/profile/nafnlaus.bsky.social/post/3lotd54jqms2t





Allikas: https://bsky.app/profile/nafnlaus.bsky.social/post/3lotd54jqms2t

Eesti e-riigi riskide analüüs
TA küberturvalisuse komisjonis

- Dan Bogdanov, PhD
- Dan Bogdanov, akadeemik ja küberturvalisuse komisjoni esimees

E-riigi edu on loonud uusi ülesandeid

- **Väide 1**: Eesti inimesed on e-teenuste mugavusega ära harjunud ning paberite liigutamine tuleks meile šokina.
- **Väide 2**: Eesti riigi e-teenustes on rohkem moodsat teaduspõhist krüptograafiat ja küberkaitset kui mujal.
- **Järeldus 1**: oleme jõudnud sõltuvusse oma tehnoloogiatest ning eeldame neilt ohutust ja turvalisust
- **Järeldus 2**: tehnoloogiasektori trend, millega kasutajatelt järjest võetakse ära kontrolli seadmete ja teenuste üle,
  on pikas perspektiivis oht Eesti riigi suveräänsusele

E-riigi käideldavus ja kestlikkus

- **Fakt**: Eesti on tänaseks andmepõhine riik, mille registreid enam paberil ei tarvitsegi eksisteerida
- **Fakt**: me oskame ehitada Eestis mitme andmekeskusega infosüsteeme, aga väljaspool Eestit on taristut palju rohkem
- **Järeldus 1**: Eesti peab kestlikkuse ja kuluefektiivsuse jaoks kaaluma väljaspool Eestit asuva digitaalse taristu kasutamist
- **Komisjoni uurimisküsimus**: kuidas me tasakaalustame Eesti e-teenuste kestlikkuse, tervikluse, konfidentsiaalsuse ja ühiskondlike mõjude nõudeid?

Kui inimesed kord juba midagi ette võtavad,
siis ei ole risk kunagi “null”.

Oluline on riskidest teadlik olla
ja nende kahandamisega tegeleda.

Komisjoni peamine eesmärk:

ühtlustatud turbeohtude kataloog ja riskianalüüsi metoodika

Komisjoni tegevus 2023–2025

- 15. detsembril 2022 andis akadeemia juhatus akad Bogdanovile ülesandeks taaskäivitada küberturvalisuse komisjon.
- Uus küberturvalisuse komisjon kohtus esimest korda 14. augustil 2023
- 2023–2024 fookusteema oli valimiste tehnoloogia.
  - Hinnati 31 ohtu.
  - Korraldati esimene Usaldusest ja usaldatavusest konverents.
- 2024–2025 fookusteema oli valimiste tehnoloogia.
  - Hinnati 20 ohtu.
  - Korraldati teine Usaldusest ja usaldatavusest konverents.

Komisjonis hinnatavad ohud #1

1. E-riigi andmete hoidmine vaid Eesti territooriumi andmekeskustes tekitab riski terviklusele kui andmekeskus hävib
2. Suurte tehnoloogiaettevõtete lahendustest sõltumine on oht põhiseaduslikule korrale ja suveräänsusele
3. Riigi andmete (isikuandmed, aga ka muu) töötlemine EL-välise ettevõtte pilvandmetöötluse süsteemis võib kahjustada andmete konfidentsiaalsust ja käideldavust
4. Kriitilistel taristul kasutatava infotehnoloogia tarneahela kompromiteeritus põhjustab konfidentsiaalsuse, tervikluse või käideldavuse kao

Komisjonis hinnatavad ohud #2

5. Suurte keelemudelite laialdane rakendamine võib muuta suure rahvahulga väärtushinnanguid, tõekspidamisi ja võib viia ohuni riigi põhiseaduslikule korrale
6. Tehisintellektimudelite laialdane kasutust vähendab (üli)õpilaste iseseisvat mõtlemist ja otsustamise suveräänsust
7. Sünteetilise sisu ja isikute tajutav eristamatus aitab kaasa küberkuritegevusele, väärinfo levikule ja õiguste vähenemisele

Komisjonis hinnatavad ohud #3

8. Eestlaste madal andmete ja tehisintellektialane kirjaoskus ohustavad meie põhiõiguseid
9. Tehisintellektimudelite laialdane kasutust vähendab (üli)õpilaste iseseisvat mõtlemist ja otsustamise suveräänsust
10. Sünteetilise sisu ja isikute tajutav eristamatus aitab kaasa küberkuritegevusele, väärinfo levikule ja õiguste vähenemisele

Kokku on komisjoni tööplaanis üle 20 ohu.

“Kas te kavatsete ainult hirmutada?”

- Komisjoni metoodika järgi on ohuanalüüsis kirjas
  - ohu kirjeldus,
  - eeldused ohusündmuse realiseerumiseks,
  - võimaliku ründaja profiil,
  - täna olemasolevad kaitsemeetmed – mida me juba teeme?
  - võimalikud täiendavad kaitsemeetmed – mida võiks veel teha?

- Ehk siis iga ohu kohta anname ka nõu, kuidas seda vältida.

Ohusündmuste sõnastamine

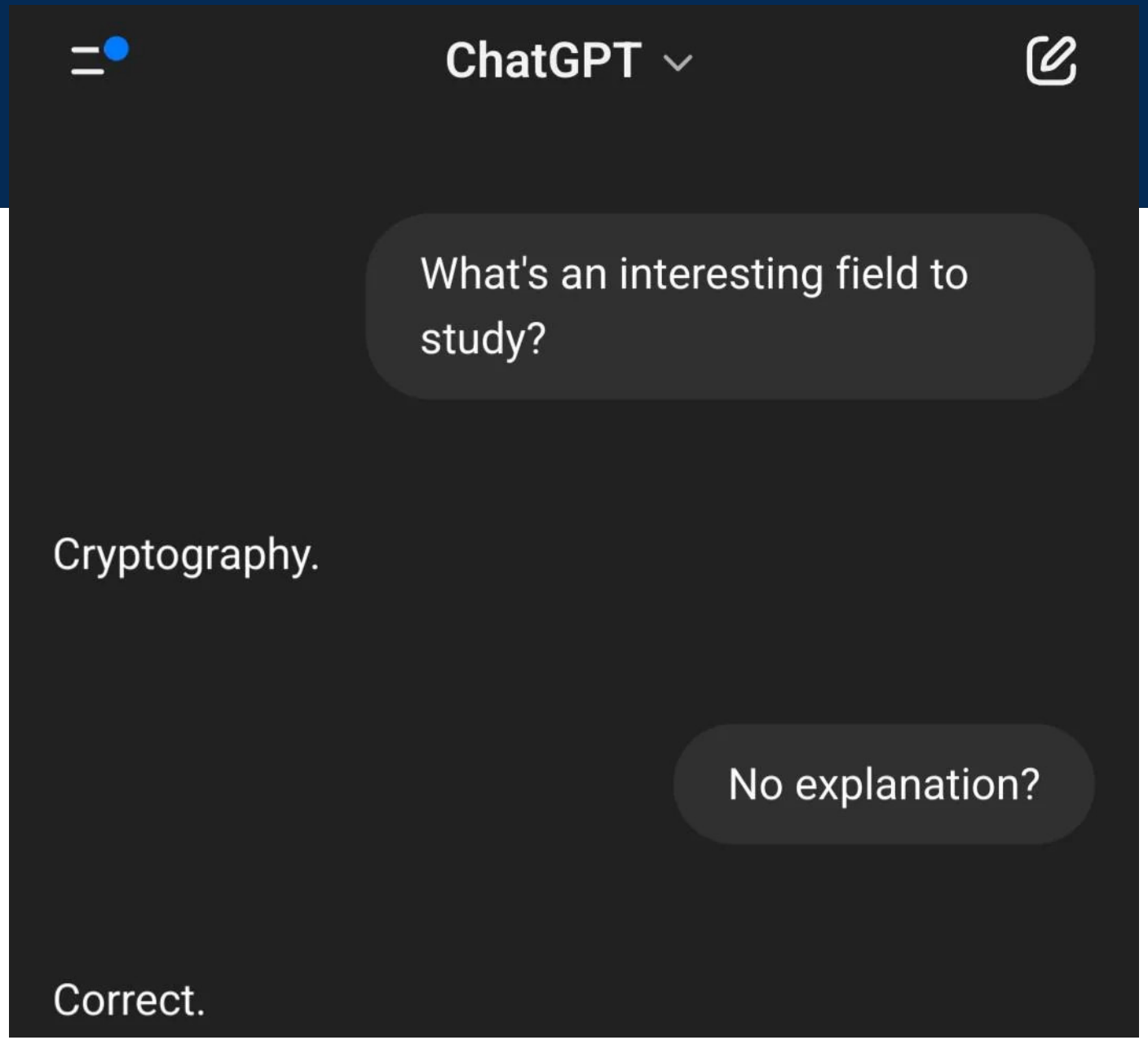
- Uued skaalad, mitu?
- Mis on tegelikud kahjud?

Komisjoni liikmed alates 2025. aastast

- Komisjoni esimees akad Dan Bogdanov
- Justiits- ja Digiministeeriumi esindajad
- Riigi Infosüsteemi Ameti esindajad
- Andmekaitse inspektsiooni esindajad
- Tartu Ülikooli teadlased
- TalTechi teadlased
- Erasektori teadlased
- Ametliku koosseisu nimekirja leiad TA veebilehelt: https://www.akadeemia.ee/akadeemia/noukogud-ja-komisjonid/

Millal on tulemusi oodata?

- 2025. aasta lõpus tegeleme riski tõsiduse skaalade sõnastamisega ning esimeste ohukirjeldustega.
- Esimesed ohuanalüüsid avaldame usutavasti 2026. aastal.
- Komisjoni istungite protokollid on avalikud Teaduste akadeemia veebilehel



Allikas: https://bsky.app/profile/nafnlaus.bsky.social/post/3lotd54jqms2t