

# TEHISARU KASUTAMINE HARIDUSES JA TEADUSES. EETILISED VÄLJAKUTSED

**MARGIT SUTROP**

Tartu Ülikooli tehisaru pädevuspäev

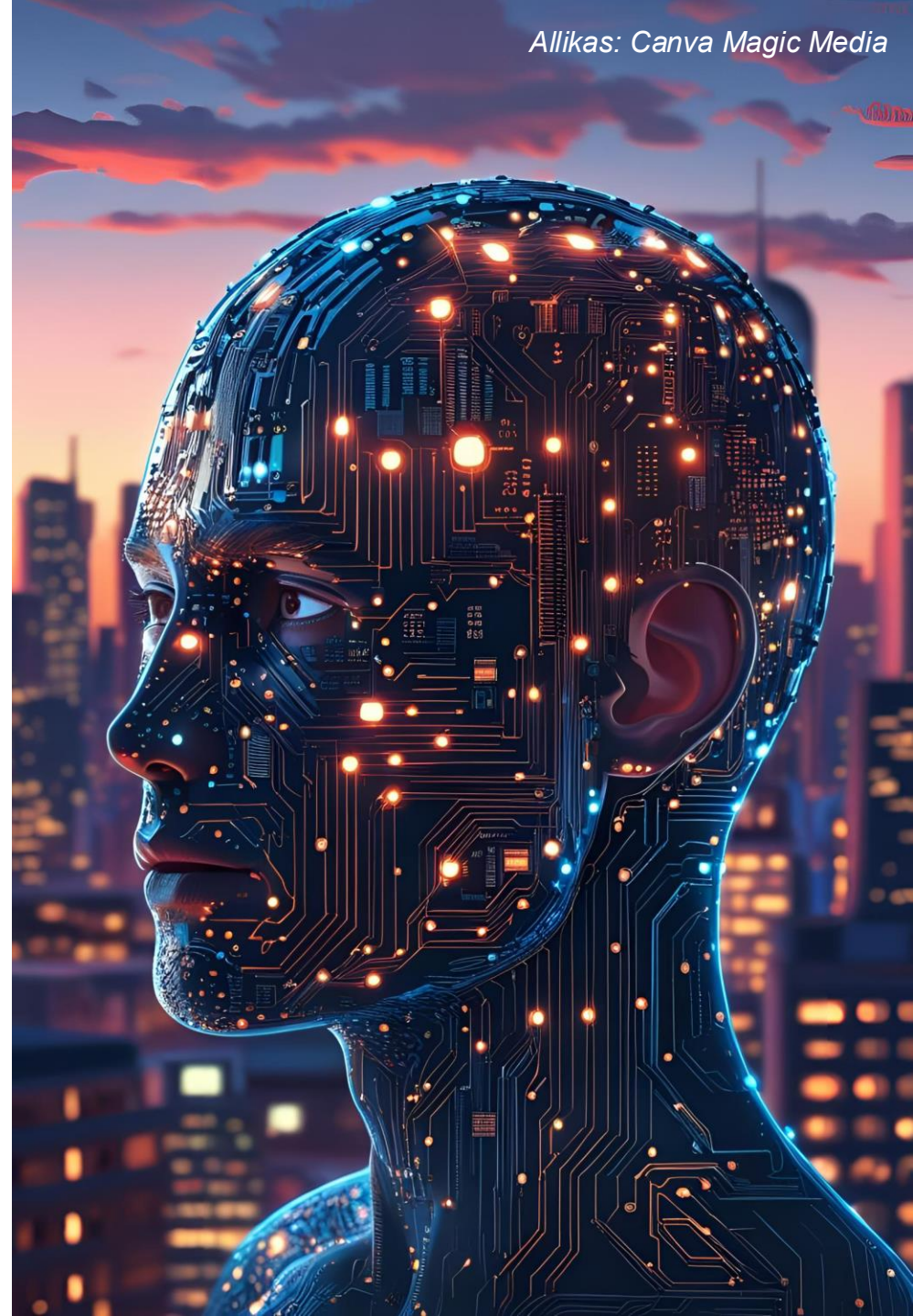
22. september 2025

# TEHISINTELLEKTI EETIKA VÕRGUSTIK

- 28.11.2024 asutasime tehisintellekti eetika võrgustiku
- Võrgustiku eesmärk on koordineeritult ühendada Eestis tehisintellekti ja eetika valdkonnaga seotud tegevused.
- Eesmärk on kasvatada tehisaru eetikaalast pädevust Eestis; õppida üksteiselt, jagada teadmisi, hoida end kursis üksteise tegemistega ning tutvuda käimasolevate projektidega
- Tegevus: Korraldame koostööseminare, õppekohtumisi ning vastavaid teadus- ja koostööprojekte nii Eesti kui ka välismaa ekspertidega.
- Võrgustik toetab Eesti tehisintellekti tippkeskuse (EXAI) tööd, aidates kaasa tehisaru arendamisele ja vastutustundlikule rakendamisele.
- Liitumiseks kirjuta: Kaisa-Maria.Hagel@ut.ee

# KAVA

- Tehisaru definitsioon ja arengutasemed
- Tehisaru eetika
- Tehisaru valitsemise üleilmne raamistik
- Tehisaru hariduses ja teaduses
- Võimalused ja ohud



# MILLEST ME RÄÄGIME, KUI ME RÄÄGIME TEHISARUST?

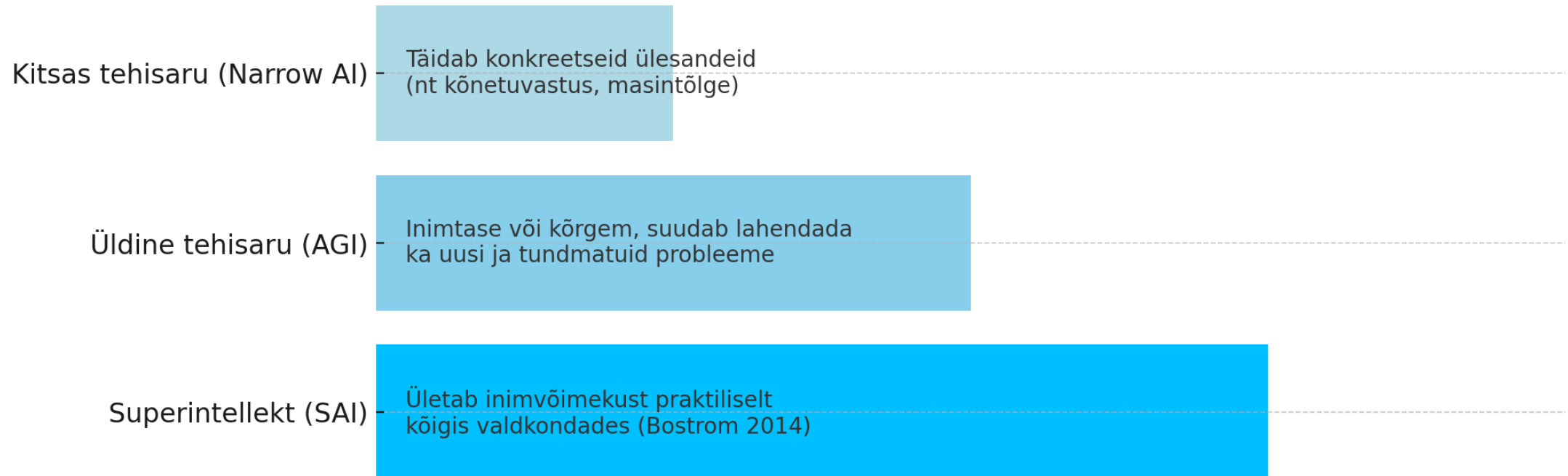
**Tehisaru süsteemid** on inimeste loodud tarkvara (ja vahel ka riistvara), mis:

- **tajuvad oma keskkonda** andmete kaudu,
- **tõlgendavad ja analüüsivad andmeid**,
- **teevad otsuseid**, et saavutada mingi eesmärk,
- ning **õpivad ja kohandavad oma käitumist** varasemate kogemuste põhjal.
- Al võib põhineda sümbolipõhistel **reeglitel** või õppida **arvulistest mustritest** (nt masinõpe).

*Allikas: Euroopa Komisjon, 2019a*

# TEHISARU TÜÜBID EHK ARENGU-TASEMED

## Tehisaru arengutasemed



→ Kasvav intelligentsus ja võimekus

# TEHISARU ÜLEILMNE VALITSEMISMAASTIK

- **OECD tehisintellekti põhimõtted** (2019) – esimene valitsustevahelisõne standard. Tehisintellekti globaalses partnerluses (GPAI) osaleb 44 liikmesriiki  
<https://www.oecd.org/en/topics/ai-principles.html>
- **UNESCO AI eetika soovitused** (2021). Ratifitseerinud 194 riiki  
<https://unesco.ee/dokumendid/unesco-soovitused-eesti-keeles/>
- **UNESCO: AI and education: protecting the rights of learners**  
DOI: <https://doi.org/10.54675/ROQH4287>
- **EL-i tehisintellekti määrus** (2024): riskipõhised tasemed, kõrge riskiga süsteemide kohustused  
<https://www.europarl.europa.eu/topics/et/article/20230601STO93804/euroopa-tehisintellekti-maarus-eli-plaan-reguleerida-tehisintellekti>

# TEHISARU EETIKA

- **Tehisaru eetika** on praktilise eetika haru, mis uurib tehisintellekti (AI) arendamise, kasutamise ja mõju eetilisi küsimusi.
- **Eetika põhineb väärtustel. Väärtused** on asjad, mida me hindame ja mida me ei taha kaotada.
- **Eetiline käitumine** tähendab tegutsemist kooskõlas väärtustega.
- Kuna **väärtused võivad pörkuda**, siis on sageli vaja ka väärtuste kaalumist ja põhjendamist.
- Väärtuste kaitsmiseks sõnastatakse **juriidilised normid ja printsiibid**.

# UNESCO arvates tehisaru mõjud

- **Haridus:** digitaalses ühiskonnas elamine nõuab uut moodi haridust, eetilist eneseanalüüsi, kriitilist mõtlemist, vastutustundlikku disaini ja uusi oskusi, võttes samas arvesse mõju tööturule, töölasele konkurentsivõimele ja kodanikuosalusele.
- **Teadus** kõige laiemas tähenduses hõlmab kõiki akadeemilisi valdkondi loodus- ja meditsiiniteadustest sotsiaal- ja humanitaarteadusteni. Tehisintellekti tehnoloogiad pakuvad uusi uurimisvõimalusi ja käsitleviise, avaldavad mõju meie teadusliku mõistmise ja seletuse kontseptsioonidele ning loovad otsuste tegemiseks uusi aluseid.

<https://www.unesco.org/en/articles/what-you-need-know-about-unescos-new-ai-competency-frameworks-students-and-teachers?hub=32618>

# MIDA TEHISARU TEHA SUUDAB?

- 7 aastat tagasi ennustasid teadlased, et tehisaru ületab inimesi **tõlkimises** aastaks 2024, gümnaasiumi **essee**de kirjutamises aastaks 2026, **veoki juhtimises** aastaks 2027 ja **kirurgina** töötamises aastaks 2053.
- Teadlased arvasid, et on 50% tõenäosus, et AI ületab inimesed kõikides ülesannetes 45 aasta jooksul ning automatiseerib kõik inimtööd 120 aastaga. (Grace, K. jt 2018)
- Tänapäeval kasutatakse tehisintellekti laialdaselt digimajanduses, e-kaubanduses, voogedastusteenustes ja vestlusrobotites.
- ChatGPT, mille OpenAI käivitas 2022. aasta novembris, on olnud omamoodi šokk ja ületanud ka visionääride ootusi. Tehisaru muljetavaldav **kontekstitaju** ja **dialoogivõime** aitavad inimesi paljude ülesannete puhul.

# Tehisaru hariduses

- ❖ Õppimise personaliseerimine – tehisaru saab kohandada õppeprotsessi vastavalt õppija vajadustele, tempole ja tasemele
- ❖ Õppejõu töö toetamine – rutiinsete ülesannete tegemisel aja säästmine (vastuste kontroll, tagasiside loomine, õpiväljundite sõnastamine, õppekavade koostamine, materjalide otsimine)
- ❖ Õppijate kaasamine – interaktiivsed vestluspartnerid simulatsioonid, virtuaalsed laborid
- ❖ Õppejõu rolli muutumine- materjali edastamise asemel juhendaja ja suunaja, kriitilise mõtlemise arendaja

# Salman Khan (2025)

„To take some advantage of this technology requires some creativity and bravery too. /.../ To get the very best of this technology, we need to reconsider what is now possible. We also need to ponder how to mitigate the risks, our fears, and our hesitations. This requires rethinking everything from **the role of teachers**, to how kids **use multimedia** ... to how to help graduates find jobs once they enter the **workforce**.“

*Khan, S. (2025). Brave New Worlds. How AI Will Revolutionize Education (and Why That's a Good Thing). Penguin 2025*

# TEHISARU ABI TEADUSLIKE TEKSTIDE KIRJUTAMISEL

- Aitab ületada kirjutamisärevust,
- parandab kirjutamise tõhusust,
- teksti sidusust,
- argumentatsiooni ülesehitust,
- aitab lühendada, muuta koherentsemaks,
- aitab struktureerida,
- töötada läbi andmemassiive,
- otsida kirjandust, näiteid.

# TEHISARU KASUTAMISE OHUD

- ChatGPT-I on endiselt tõsiseid väljakutseid korrektsel viitamisel,
- andmete täpsuse ja genereeritud sisu osas, sh ebatäpne või aegunud teave,
- sooliste või rassiliste eelarvamuste levitamine
- oht originaalsele mõtlemisele: plagiaat (sh tahtmatu)
- autorsuse ähmastamine.

*Head ülevaateartiklid: Cheng 2025; Yuandyan Liu 2025*

# TEHISARU TEADUSES: VÕIMALUSED

- Uurimisprotsessi kiirendamine – andmeanalüüs, hüpoteeside genereerimine, tulemuste analüüs
- Kirjutamise tugi
- Eksperimentide ja simulatsioonide optimeerimine
- Kiirem ja parem juurdepääs teadustulemustele

# TEHISARU TEADUSES: OHUD

- ❖ Hallutsineerivad vastused, läbipaistmatus
- ❖ Akadeemiline ausus
- ❖ Viidete ja faktide kontrollimise vajadus
- ❖ Ebaõiglase konkurentsieelise tekkimine, kui üks kasutab, teine mitte (kas kõik saavad kasu või süveneb lõhe?)
- ❖ Andmekaitse ja privaatsus
- ❖ Vastutus: kes vastutab, kui AI teeb vigu?

# TEHISARU PEAB VASTAMA NEILE VÄÄRTUSTELE

- Inimkeskne AI
- Autonomia ja inimväärikuse austamine.
- Ausus
- Privaatsus, andmekaitse
- Õiglus
- Läbipaistvus

# KOKKUVÕTE

- Tehisaru on/peab olema vahend, mitte eesmärk
- Tehisaru kasutamine toob kaasa vajaduse mõtestada ümber õpetamist.
- Võimalused ja ohud tuleb läbi mõelda.
- Ohud tuleb muuta võimalusteks.
- Väärtused, mis on ohus, on ausus, iseseisev mõtlemine, loovus.

# KIRJANDUS

- Cheng, A., Calhoun, A. & Reedy, G. (2025). Artificial intelligence-assisted academic writing: recommendations for ethical use. *Adv Simul* **10** (22). <https://doi.org/10.1186/s41077-025-00350-6>
- EU Commission (2019a). *A definition of AI: main capabilities and disciplines*. Available online at <https://www.aepd.es/media/docs/ai-definition.pdf>
- EU Commission (2019b). *Ethics guidelines for trustworthy AI*. Available online at <https://ec.europa.eu/futurium/en/ai-alliance-consultation>
- Grace, K. et al. (2018). "When will AI exceed human performance? Evidence from AI experts", <https://arxiv.org/abs/1705.08807>
- Iphofen, R. and Kritikos, M. (2019). "Regulating artificial intelligence and robotics: ethics by design in a digital society". *Contemporary Social Science* **2041**, 1–15. doi: 10.1080/21582041.2018.1563803.
- Khan, S. (2025). *Brave New Worlds. How AI Will Revolutionize Education (and Why That's a Good Thing)*. Penguin 2025
- Sutrop, M. (2019). Should We Trust Artificial Intelligence? *Trames*, **23** (4), 499–522. <https://doi.org/10.3176/tr.2019.4.07>
- Sutrop, M. (2020). Challenges of Aligning Artificial Intelligence with Human Values. *Acta Baltica Historiae et Philosophiae Scientiarum*, **8** (2), 54–72. DOI: 10.11590/abhps.2020.2.04.
- Yuandyan Liu, et al.,. 2025. ChatGPT application in academic writing: a review of potential, limitations, and ethical challenges", *Arquivos Brasileiros de Oftamologia*, **88** (3): e 2024-0269 DOI: 10.5935/0004-2749.2024-0269.