

ERINDISBRÉF

Starfshópur um smíði verklagsreglna um notkun gervigreindar í starfsemi Reykjavíkurborgar

Ábyrgðarmaður: Óskar J. Sandholt

Inngangur:

Notkun gervigreindar (e. Artificial Intelligence) er orðin almenn og aðgengileg. Nauðsynlegt er að Reykjavíkurborg birti tilmæli er varðar notkun gervigreindarhugbúnaðar í starfsemi borgarinnar, bæði til að hvetja fólk til hagnýtingar á tækni sem getur auðveldað þeim vinnu sína, en á sama tíma tryggja að fólk sé meðvitað um takmarkanir gervigreindarhugbúnaðar og hvaða hættur þeim geti fylgt.

Spunagreind (e. Generative AI) er sú gerð gervigreindar sem hefur fengið mesta athygli að undanförunu enda hafa orðið umfangsmiklar framfarir á stuttum tíma. Eins hefur aðgengi að spunagreind aukist til muna. Með spunagreind er einfalt að búa til nýtt efni með einföldum fyrirmælum (e. prompt). Þannig er einfalt að fá drög að texta, tónlist, myndir o.s.frv. Spunagreind byggir á þeim gögnum sem hún hefur aðgang að og er þjálfuð á gögnum sem hún hefur séð áður. Dæmi um spunagreind er til dæmis hugbúnaður sem getur skrifað nýja texta í stíl viðurkenndra höfundar eða mynstur sem líkja eftir ákveðnu listaverki.

Gervigreind gerir ekki greinamun á réttu og röngu og getur ekki nýtt rökhugsun til að átta sig ef niðurstöður sem hún gefur stangast á við lög eða ýta undir misrétti. Líkönin greina heldur ekki heimildir, svo ómögulegt er fyrir þau að vita hvort þau byggji svör sín á réttum eða röngum upplýsingum. Líkönin eru hins vegar fljót að læra og geta því auðveldlega búið til sín eigin rök eða búið til samhengi á milli hluta sem eiga sér ekki stoð í raunveruleikanum.

Mikilvægt er að í starfsemi borgarinnar sé gervigreind notuð með ábyrgum hætti. Í því flest m.a. að tryggja að notkun, hönnun og þróun á gervigreindarlíkönnum sé upplýst og ábyrg, hugað sé að gæðum gagnasafna og farið eftir siðferðisstöðlum. Eftir því sem gervigreindarlíkön verða öflugri, klárari og algengari er nauðsynlegt að tryggja að þau séu notuð og þróuð á ábyrgan hátt.

Hlutverk:

Hlutverk hópsins er að móta verklagsreglur um notkun gervigreindar í starfsemi Reykjavíkurborgar.

Helstu verkefni/viðfangsefni:

- Setja ramma fyrir notkun gervigreindar í starfsemi Reykjavíkur.
- Setja fram leiðbeiningar fyrir starfsfólk um notkun gervigreindar þegar kemur að daglegum störfum.

Starfshópurinn skipa:

Karen María Jónsdóttir (formaður), skrifstofustjóri skrifstofu sviðsstjóra, þjónustu- og nýsköpunarsvið

Aldís Geirdal Sverrisdóttir, teymisstjóri lögfræðipjónustu, þjónustu- og nýsköpunarsvið

Búi Bjarmar Aðalsteinsson, stafrænn leiðtogi, þjónustu- og nýsköpunarsvið

Inga Rós Gunnarsdóttir, skrifstofustjóri gagnapjónustu, þjónustu- og nýsköpunarsvið

Próstur Sigurðsson, skrifstofustjóri Stafrænnar Reykjavíkur, þjónustu- og nýsköpunarsvið

Starfsmaður/verkefnisstjóri starfshópsins:

Eva Pandora Baldursdóttir, verkefnastjóri stjórnsýslumála, þjónustu- og nýsköpunarsvið

Til ráðgjafar og samstarfs:

Starfshópurinn vinni hagsmunaaðilagreiningu sem síðan verði lögð til grundvallar því samráði og samstarfi sem þarf að hafa vegna vinnu hópsins. Starfshópurinn leiti einnig ráðgjafar innan og utan borgarkerfis hjá aðilum með sérþekkingu á viðfangsefninu eftir því sem við á.

Starfstími:

Starfshópurinn skili ábyrgðarmanni kostnaðarmetnum tillögum fyrir lok desember 2023.

Reykjavík, 1. nóvember 2023



Óskar J. Sandholt