



AÐGERÐARÁÆTLUN UM GERVIGREIND, 2024-2026:

Athugasemdir og tillögur að breytingum og viðbótum við aðgerðaráætlun stjórnvalda í gervigreind

Gervigreindarsetur Háskólans í Reykjavík
Vitvélstofnun Íslands ses
STEM Ísland

FORMÁLI: EÐLI GERVIGREINDAR ER Í STÖÐUGRI ÞRÓUN

Gervigreind er vísindalegt rannsóknarsvið með langa sögu að baki. Gervigreind er meira en afurðir rannsókna og þróunar á sviðinu og er því hvorki ákveðin tækni né ákveðnar hagnýtar aðferðir, heldur mengi vísindalegrar þekkingar um eðli og virkni hugsunar og greindar.

Við upphaf gervigreindarsviðsins, sem var stofnað árið 1956, takmarkaðist starfsemi þess að mestu við rannsóknarstofnanir. Gervigreind snýst um að sjálfvirkinnæða hugsun – verðugt vísindalegt verkefni sem hefur þá óvenjulegu hliðarafurð að hugmyndir hennar eru hagnýtanlegar á mörgum sviðum þjóðfélagsins, þó ekki öllum.

Rannsóknir á tækninni sem oftast er nú vísað til sem „gervigreind,” svo sem ChatGPT, Midjourney, o.fl., má rekja um 90 ár aftur í tímann. Tæknin byggir á hugmyndum um gervitauganet á fjórða áratug síðustu aldar, sem voru innblásnar af því hvernig náttúruleg tauganet höndla upplýsingar. Fyrsta frumgerð að gervitauganeti var búin til árið 1952. Á 7. og 8. áratugnum voru slík kerfi smíðuð sem forrit í tölvum þess tíma. Það tók um 60 til 90 ár fyrir þessar hugmyndir að verða hagnýtanlegar.

Nútíma gervigreindartækni, sem að miklu leyti er byggð á gervitauganetum, er mörgum annmörkum háð, eins og er sífellt að koma betur í ljós. Þótt sumar þeirra takmarkana, sem flestar má rekja til eðlis og uppbyggingar gervitauganeta, verði mögulega rutt úr vegi á næstu árum, eru aðrar sem liggja djúpt, til dæmis útskýringar á hegðun þeirra, og krefjast því nýrra

nálgana. Líkt og á öðrum vísindasviðum eru nýjar grundvallarnálganir háðar akademískum rannsóknum sem geta litið til lengri tíma en hagnýtar rannsóknir.

Helstu takmarkanir nútíma gervigreindar tengjast eftirfarandi þáttum:

1. Kröfum um gífurlegt gagnamagn. Gervigreindartæknin sem bandarískir tæknirisar eru að próa gera ráð fyrir tífalt til tíupúsundfalt meira gagnamagni en Ísland býr yfir, hvert svo sem sviðið er. Ísland er smágagnaland, ekki land gífurgagna. Þetta á reyndar líka við um flest smærri lönd Evrópu.
2. Kröfum um mikla reiknigetú. Reiknigetan til að útbúa grunntækni í nútíma gervigreind kostar ekki aðeins fjármagn, sem fer beint úr landi, heldur krefst hún sífellt meiri orku, sem vinnur gegn loftslagsmarkmiðum.
3. Duldam þróunar- og viðhaldskostnaði. Þegar tækni byggðri á gervitauganetum er skipt út telur kostnaðurinn oft nánast jafn mikið og upprunalegi kostnaðurinn við tæknina sem hún leysir af hólmi.
4. Óáreiðanlegri and siðferðislega varasamri hegðun. Þar sem gervigreind byggð á gervitauganetum er „þjálfuð“ á gögnum af Internetinu, sem öll eru framleidd af notendum þess, erfir hún og jafnvel magnar upp allar þær rangærslur, fordóma, hatursorðræðu og misskilning sem þar er að finna. Þar að auki er „þekkingarframleiðsla“ þessarar kerfa byggð á aðferðum sem ekki eru „jarðtengdar,” þ.e.a.s. byggja ekki á traustri þekkingu um hvernig heimurinn virkar. Þar er komin uppspretta óáreiðanlegrar og ófyrirséðrar hegðunar slíkra kerfa.

Hagnýting nútíma gervigreindar á sviðum heilsuverndar og læknisþjónustu, umferðarstjórnunar, björgunaraðgerða, lögreglustarfa, lögfræði, stórtækra áætlanagerða og annarra mikilvægra verkferla er því í mörgum tilfellum óábyrg og í versta falli siðlaus.

Á þeim tæplega 70 árum sem hafa liðið frá stofnun gervigreindarsviðsins hafa, til viðbótar gervitauganetum, fjöldi annara afurða, hugmynda, frumgerða, og tækni litið dagsins ljós. Ýmis ný gervigreindartækni mun því koma útúr rannsóknarsetrum á næstu árum og áratugum. Ljóst er að sum þessarar tækni mun leysa nútíma gervigreindartækni af hólmi, sérstaklega sú sem býður betri nálgun en eitt eða fleiri af ofantöldum vandamálum.

Í gervigreindarstefnum Norðurlandanna, og fjölda annarra landa í Evrópu, er mikið er rætt um trausta, örugga og siðferðislega hagnýtingu tækninnar. Helstu ástæður slíkra áhyggja er ekki að rekja til sjálfvirknivæðingarinnar sjálfrar, enda mæla lög flestra landa þegar hvað megi og megi ekki með tækni, heldur til eðli og hegðunar nútíma gervigreindar og ofantalinna takmarkana og óæskilegra hliðarverkana.

ÞÖRF Á LENGRI SJÓNDEILDARHRING

Þegar kemur að gervigreind þykir okkur afar varhugavert að setja aðgerðaráætlun (eða áætlanir) sem ná aðeins til tveggja ára, þar sem hún mun hafa áhrif og marka djúp spor í þjóðfélög allra

landa næstu 50 árin. Við teljum að trúverðug gervigreindaráætlun ætti að marka stefnu fyrir næstu 10 árin hið minnsta. Að slíkri áætlun þurfa mun fleiri að koma en starfsmenn ráðuneyta og örfáir aðilar frá atvinnuvegunum – þar þarf sérþekkingu vísindamanna sem hafa þekkingu og færni til að útlista næstu skref á því sviði, geta greint neikvæðar afleiðingar tiltekinnar tækni, og geta ráðlagt um hvar fjármagni er best varið til að ná settum markmiðum um sjálfvirknivæðingu.

ÍSLAND ÞARF SMÁGAGNAGERVIGREIND

Fyrir um tveimur áratugum stóð Afríka frammi fyrir vandamáli: Til að fylgja vestrænum þjóðum í samskiptatækni í að nútímavæða símakerfi heimsálfunnar þyrfti að grafa gífurlegt magn skurða, með tilheyrandi kostnaði, og leggja í þá símalínur, verkefni sem myndi taka a.m.k. 20 ár. Í staðin var ákveðið að hoppa yfir þann áfanga og fara beint í farsímvæðingu.

Að sama skapi leggjum við til að í stað þess að reyna að keppa við Bandaríkin í þróun og hagnýtingu nútíma gervigreindar stökkvi Ísland beint í næstu tegund gervigreindar, sem mun leysa flest ofantalin vandamál og því vera hagnýtanleg í verkefnum sem krefjast öruggar, og fyrirsjáanlegrar hegðunar sjálfvirkni, sem eru nauðsylegar forsendur siðferðislegar hagnýtingar ekki bara gervigreindartækni, heldur allrar tækni.

Sú gervigreindartækni sem myndi nýtast Íslandi best væri gervigreind sem getur unnið með smágögn. Slík gervigreind er ekki til „úti í búð,” en hún er nú þegar til inni á nokkrum gervigreindarsetrum, meðal annars í HR og Vitvélastofnun Íslands. Með því að byrja að undirbúa hagnýtingu slíkrar gervigreindar strax getur Ísland tekið hástökk framúr öðrum löndum og orðið leiðandi allra landa sem eru í sömu stöðu og við, þ.e. að eiga ekki mikið – eða engin – gífurgögn.

Sé rétt á málum haldið mun þannig gervigreind líka geta skilið orsök og afleiðingu, sem er lykillinn að gervigreind sem getur útskýrt hvað hún er að gera. Þegar slík gervigreind er orðin nógu þroskuð mun hún taka yfir alla sjálfvirknivæðingu á spítölum, flugumferðarstjórnun, lögfræðistarfsemi, löggæslu, netöryggi, og svo mætti lengi telja.

GERVIGREIND: HVAÐ ÞARF ÍSLAND AÐ GERA?

Mikilvægt er að styrkja tengsl hagnýtra gervigreindarransókna í akademíu og atvinnuvegi. Þessi tengsl eru lífæð hagnýtingar rannsókna og þróunar – „færiband nýsköpunar” sem færir hugmyndir, fólk, verkefni og lausnir af frumstigi yfir á hagnýtanlegt form. Ríkisstuðningur er hér mikilvægur vegna þess að sjóndeildarhringur akademíunnar og atvinnuvegarins skarast ekki, heldur eru aðskilin að a.m.k. 6 árum. Nýsköpunarfæribandið hefur aldrei verið jafnmikilvægt eins og fyrir gervigreind, því gervigreind hefur möguleika á margföldunaráhrifum. Ríkið getur stutt nýsköpunarfæribandið með margvíslegum hætti, en við viljum nefna hér nokkur atriði sem

er tiltölulega auðvelt að útfæra strax, þarf ekki endilega að byrja með offorsi, og hefur mikla möguleika á að vaxa á komandi árum og áratugum:

1. Styðja við starfsnám (e. Internship) í gervigreind fyrir masters (MSc) og doktorsnemendur (PhD) í háskólum landsins.
2. Styðja beint starfsemi eins og Vitvélastofnun Íslands, sjálfseignarstofnun sem hefur yfir áratugar reynslu í hagnýtingu gervigreindar með opnum hugbúnaði, með það fyrir augum að tryggja stöðugleika til samstarfs við smáfyrirtæki í rannsóknum og þróun gervigreindar sem lofar hagnýtingu á tímabili 5 til 9 ár. Þannig má brúa bilið milli háskólarannsókna, sem horfa 10 ár eða lengra fram í tímann, og atvinnuvegarins, sem horfir sjaldan lengra en 4 ár fram í tímann.
3. Auka framlög í bæði rannsóknarsjóði og tækniþróunarsjóði með sérstaka áherslu á smágagnagervigreind.
4. Styðja sérstaklega við doktorsnám í gervigreind og gagnavísindum til þess að fjölga í þeim hópi sem mun þróa næstu kynslóð tækninnar.
5. Fjárfesta í aukinni færni og betri menntun í gervigreindar- og gagnatengdum greinum á háskólastigi og skyldunámi á lægri menntastigum.

Aukin staðbundin færni í gervigreind er ein af forsendum þess að hægt sé að innleiða gervigreind markvisst í íslenskt samfélag. Með hvatakerfi árangurstengdrar fjármögnunar háskólastigsins verður stutt við fjölgun nemenda í gervigreindar- og gagnatengdum greinum og aukna þekkingu á notkun gervigreindar almennt. Sérstök áhersla þarf að leggja stuðning við núverandi M.Sc. nám í gervigreind við Háskólann í Reykjavík, sem er meðal þeirra fremstu á norðurlöndunum, gerð nýrra námskeiða og stuðningi við rannsóknir og þróun, ásamt auknum tengslum við innlend og erlend fyrirtæki. Jafnframt þarf áfram stuðning við nám í netöryggi og gervigreind, í samstarfi Háskóla Íslands, Miðstöð stafrænnar nýsköpunar (MSN; European Digital Innovation Hub Iceland / EDIH-IS), Vitvélastofnun Íslands og fleiri. MSN er fulltrúi Íslands í stóru samstarfsneti yfir 225 slíkra miðstöðva um alla Evrópu, sem að hluta til er fjármagnað með DIGITAL Europe áætlun framkvæmdastjórnar Evrópusambandsins. Til að styðja þetta mikilvæga framtak frekar er nauðsynlegt að 50% mótframlag komi frá ríkisfjármögnun, eins og tíðkast í öðrum Evrópuþjóðum. Þessi stuðningur hefur ekki enn verið veittur. Gera þarf átak í að kynna og nýta opin gögn við kennslu í námsáföngum á öllum skólastigum og samstarf aukið milli háskóla og tæknifyrirtækja landsins í því skyni að nemendur fái hagnýta þjálfun í notkun gagna- og gervigreindar.

Til að auka sérfræðiþekkingu má fjárfesta í að styrkja doktorsnema og nýdoktara beint í gegnum markvissa fjárveitingu, fremur en einungis úr samkeppnissjóðum. Þetta mun tryggja að fleiri nemendur geti lokið háskólagráðu á hæsta stigi í gervigreind og gagnatengdum greinum, sem stuðlar að aukinni nýsköpun og rannsóknargetu. Jafnframt á að leggja áherslu á að nýta betur þáþekkingu sem próffessorar og sérfræðingar við Háskólann í Reykjavík og Háskóla Íslands búa yfir, með því að auka rannsóknatækifæri þeirra, styrkja erlend og innlent samstarf og rannsóknahópa.

Líkt og nágranalöndin getur beinn fjármagnsstuðningur við stofnanir eins og Vitvélstofnun Íslands, sjálfseignarstofnun sem sérhæfir sig í tækniyfirfærslu á sviði gervigreindar og samstarfi fyrirtækja og háskóla, og STEM Ísland, sjálfseignarstofnun sem vinnur að því að efla STEM, menntun á grunnskólastigi, þ.m.t. í gervigreind í gegnum samstarf við NVIDIA, tryggt stöðugleika á sviði gervigreindarpróunar og hagnýtingar hennar. Aðeins með óslitinni keðju nýsköpunar, frá rannsóknarsetrum til atvinnuvegarins, mun öflugasta tæknin rata stystu leiðina að sjálfvirknivæðingu menntunar og atvinnuvegarins.

Ísland hefur jafnframt tækifæri til þess að forðast að verða eftirbátur og leiksoppur erlendra stórfyrirtækja, eins og vill gerast þegar ný tækni verður allt um lykjanði og afleiðingar hennar koma of seint í ljós, sbr. áhrif samfélagsmiðla á samskipti og uppeldi barna. Þvert á móti er nú tækifæri til þess að verða leiðandi í einstaklingsmiðaðri og tryggri sjálfvirknivæðingu sem nýtir mannauðin og þekkinguna héraðs til fulls og eykur þar með sjálfstæði, sjálfbærni og sjálfræði Íslands í þessum efnunum.

Með stefnumótandi fjárfestingum og framtíðarsýn þeirri sem lýst er í þessu bréfi getur Ísland byggt upp færni og þekkingu á notkun gervigreindar í þjóðfélaginu á breiðum grundvelli, og stuðlað að betra samfélagi til lengri tíma.

21. nóvember, 2024

F.h. Gervigreindarseturs Háskólans í Reykjavík

Stefán Ólafsson, PhD
Jón Guðnason, PhD
Kristinn R. Þórisson, PhD
Torfi Þórhallsson, DPhil

F.h. Vitvélstofnunar Íslands, ses

Kristinn R. Þórisson, PhD
Bridget Burger, MEd

F.h. STEM Ísland

Bridget Burger, MEd