

## Umsögn Landssamtakanna Proskahjálpar um stefnu um gervigreind, 15. mars 2021

Landssamtökin Proskahjálpar fagna því að vinna sé hafin við mótun stefnu í gervigreind.

Í þessari umsögn verður fjallað um gervigreind í víðu samhengi, og fellur þar undir vélanám (e. machine learning), málgreining (e. natural language processing), myndgreining (e. computer vision), gervitauganet (e. artificial neural networks), gagnasöfnun- og greining (e. data processing / data analytics), þjarkafræði (e. robotics) og algoritmi. Mörg þessara sviða eru þegar farin að hafa áhrif á samfélag okkar og lesa kerfin gögn, læra af þeim og bera kennsl á mynstur til að taka ákvarðanir. Gera má ráð fyrir hlutdrægni byggðri á tölfræði og af þeim sökum skapast sú hættu að ákvarðanir og ábyrgð á mati sé lögð í hendur tækninnar, þar sem þær ákvarðanir hafa áhrif á líf, líðan, heilsu, öryggi og aðra þætti í lífi lífandi einstaklinga í þeirra daglega lífi. Þessari tækni er þegar beitt á ýmsum sviðum um víða veröld, t.d. í löggæslu og dómskerfinu, heilbrigðisþjónustu og fjármálastarfsemi. Siðferðislegu álitamál sem koma upp í þessu sambandi þarf að fjalla um.<sup>1</sup>

Amnesty International vann árið 2017 að stefnu um gervigreind og mannréttindi þar sem fram kom að gervigreind væri spegill á gildi samfélagsins. Þar segir að ef við vinnum ekki að úrbótum og ef að við nýtum núverandi gildismat til þróunar, muni tæknin leiða af sér og festa í sessi þá fordóma og ójafnrétti sem eru í samfélagi nútímans.<sup>2</sup> Þá segir í Toronto yfirlýsingu Amnesty International frá árinu 2018 að stjórnvöld hafi skyldum að gegna þegar kemur að því að vernda og bera virðingu fyrir mannréttindum. Þetta eigi einnig við um fyrirtæki og einkaaðila sem hafa skyldum að gegna þegar kemur að því að bera virðingu fyrir mannréttindum ávallt og alls staðar.<sup>3</sup>

---

<sup>1</sup> Þessi skilgreining byggir á texta Mark Latonero úr skýrslunni „Governing Artificial Intelligence: Upholding Human Rights & Dignity“ frá árinu 2018, sjá bls. 8.

<sup>2</sup> Salil Shetty, „Artificial Intelligence for Good“, *Amnesty International*, 9. júní 2017. Sótt af <https://www.amnesty.org/en/latest/news/2017/06/artificial-intelligence-for-good/>

<sup>3</sup> *Amnesty International and Access Now*, „The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems“, 16. maí 2018. Sótt af [https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration\\_ENG\\_08-2018.pdf](https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf).

## Mannréttindalegar skuldbindingar íslenskra stjórnvalda

Stjórnvöld hafa skyldur til þess að efla, verja og framfylgja mannréttindalegum skuldbindingum. Þau eiga ekki að taka þátt í eða styðja við aðgerðir þegar hönnun og framkvæmd við notkun á kerfum sem nýta gervigreind, þ.m.t. í samstarfi við einkaaðila, getur ógnað mannréttindum. Gervigreind er í auknum mæli nýtt í dómskerfinu, í heilbrigðisþjónustu, í velferðarþjónustu og húsnæðismálum á alþjóðavísu. Stundum eru afleiðingarnar jákvæðar en mikil hættu er á mismunun og neikvæðum afleiðingum, sér í lagi fyrir jaðarsetta hópa fólks. Það er mikilvægt að stjórnvöld skapi sér stefnu og ramma til þess að framfylgja til að efla jákvæðar niðurstöður og draga úr skaða.<sup>4</sup>

### Mótun stefnu

Landssamtökin Þroskahjálp telja aftar brýnt að mannréttindasamtök, hagsmunafélög, stofnanir hins opinbera og þeir aðilar sem starfa að stafrænum málum sem varða gervigreind hafi vettvang til þess að starfa saman. Þau félög og einstaklingar sem vinna að mannréttindamálum munu aldrei geta varpað ljósi á eða geta séð fyrir þau margvíslegu og flóknu álitamál sem upp koma vegna gervigreindar. Að sama skapi eru þeir sem vinna að gervigreindarmálum ekki sérfræðingar í ýmsum fræðasviðum sem skipta miklu máli í þessu sambandi, laga og regluverki sem snýr að mannréttindamálum. Það er því mikilvægt að það verði til samráðsvettvangur þessara hópa til að byggja brú á milli mannréttindasamtaka og fræðasamfélagsins til þess að bera kennsl á og bregðast við þeim áhættum og óvissuþáttum sem vaxandi notkun gervigreindar skapar. Það á að vera sameiginlegt verkefni fyrrgreindra aðila að verja almenna borgara gegn skaðlegum áhrifum gervigreindar og vinna að því að tæknin sé nýtt til þess að efla mannréttindi.

Landssamtökin Þroskahjálp telja mikilvægt að gerður verði rammi um stafræn aðgengis- og réttindamál er varða mannréttindamál í breiðum skilningi, en sér í lagi þegar kemur að jaðarsettum hópum í samfélaginu og þá sérstaklega fötluðu fólki. Stjórnvöld þurfa að viðurkenna og vinna samkvæmt skuldbindingum sínum á sviði mannréttinda og fella þær skuldbindingar inn í lög, reglugerðir og stefnu um gervigreind. M.a. hvort verkferlar til að tryggja réttindi og mannhelgi fólks séu fullnægjandi og hvort lögum sé framfylgt, þegar álitamál vakna vegna notkunar gervigreindar. Rannsaka þarf áhrif gervigreindar og miðla þeim niðurstöðum sem fram koma. Hvar sem gervigreind er rannsökuð, þróuð og henni beitt þarf að

---

<sup>4</sup> „The Toronto Declaration,“ 2018.

vera markmið að bera kennsl á, sjá fyrir og minnka líkur á áhættum og skaða. Kalla þarf eftir að fyrirtæki og stjórnvöld þurfi að sæta ábyrgð brjóti þau gegn mannréttindum almennings og að gerður verður rammi til þess að taka á slíkum brotum.<sup>5</sup>

Þeir sem vinna með beinum hætti að þróun og notkun gervigreindar þurfa að hafa skýr lög og reglur til að styðjast við, en það þarf einnig staðla, samfélagsleg viðmið og efnahagslega hvata til þess að stuðla að aukinni virðingu fyrir mannréttindum. Mannréttindasamtök þurfa að sama skapi að vera virkir þátttakendur í eftirliti með gervigreind og vera reiðubúin til þess að beita þeim ramma sem ákvarðaður hefur verið í breiðri samvinnu. Samfélagið allt og þeir sem vinna að gervigreind þurfa að taka höndum saman til þess að meta áhættuþætti og þarfir viðkvæmra hópa. Á sama tíma og þeir sem vinna að mannréttindum munu aldrei bera kennsl á öll þau siðferðislegu álitamál sem upp munu rísa vegna gervigreindar býður fræðasviðið upp á þann verðmæta ramma sem þarf til að vinna að réttindum og öryggi almennings og jaðarsettra hópa, þ.e. að rýna í regluverk og eftirlit út frá mannréttindalögum og alþjóðlegum skuldbindingum.

Reglugerð um vernd einstaklinga í tengslum við vinnslu persónuupplýsinga (GDPR) byggist á mannréttindalegum skuldbindingum. Reglugerðin felur í sér vernd almennra borgara þegar kemur að gagnavernd (e. data protection) og friðhelgi sem leggur skyldur á fyrirtæki og stofnanir sem safna gögnum um íbúa Evrópusambandsins. Evrópuráðið segir frá því í skýrslu sinni frá 2018 um gervigreind og mannréttindi að það séu auknar áhyggjur af notkun algóritma og gervigreindar í samfélaginu og áhrif þess á mannréttindi.<sup>6</sup> Þá hefur Australia Human Rights Commission frá árinu 2018 unnið að verkefni sem hefur það markmið að skoða afleiðingar gervigreindar og annarrar tækni á mannréttindi, byggða á alþjóðlegum mannréttindaskuldbindingum og má nýta það sem leiðarvísi fyrir íslensk stjórnvöld.<sup>7</sup>

## Dæmi um innbyggða mismunun gervigreindar

Mikið hefur verið fjallað um hvernig gervigreind ýtir undir mismunun og fordóma. Einn algengasti þátturinn í þessari umræðu er hvernig andlitsgreining (e. facial recognition) ber ekki

---

<sup>5</sup> „The Toronto Declaration,“ 2018.

<sup>6</sup> Latonero. 2018:20-21; *Evrópuráðið*. „Algorithms and Human Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications,“ mars 2018. Sótt af <https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>

<sup>7</sup> *Australian Human Rights Commission*, „Human Rights and Technology Issues Paper, Australian Human Rights Commission,“ desember 2019. Sótt af [https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights\\_2019\\_DiscussionPaper.pdf](https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights_2019_DiscussionPaper.pdf)

kennsl á og/eða gerir ekki greinamun á fólki með dekkri húðlit og þá sérstaklega konur. Er það sagt vera vegna skorts á gögnum og ónógrar „þjálfunar“ búnaðarins. Þannig er sú kerfislega mismunun sem fólk með dekkri húðlit þarf að þola fest í sessi með hagnýtingu tækninnar. Annað dæmi er hvernig dóms- og löggæslukerfi nýta sér gervigreind sem byggist á fordómum til þess að meta líkurnar á því að sakborningur brjóti af sér aftur, svörtu fólki í óhag. Til að mynda hafa stórfyrirtæki selt lögreglunni í Bandaríkjunum aðgang að slíkum andlitsgreiningarbúnaði. Sams konar búnaður er notaður á flugvöllum, á húsnæðismarkaði og í atvinnuráðningum.<sup>8</sup> Þá eru sjálfvirk kerfi notuð í Bandaríkjunum til þess að meta líkur á vanrækslu og ofbeldi gegn börnum. Slík kerfi hafa áhrif á ákvarðanatöku fagðila, sem eru líklegri til að efast um dómgreind sína. Kerfin byggjast á upplýsingum um fjölskyldur sem nýta sér opinbera aðstoð, sem hefur meiri áhrif á félagslega viðkvæma hópa, sem t.d. búa við fátækt eða glíma við fötlun og/eða veikindi. Þessi mismununaráhrif geta leitt af sér skaða á öðrum sviðum mannréttinda, svo sem varðandi aðgang að menntun, húsnæði, á sviði fjölskyldumála og atvinnu. Jafnvel þó að það sé ekki ætlunin að viðhalda mismunun er alltaf sú hættu að afleiðingin verði þannig, ef eini mælikvarðinn á árangur er skilvirkni og hagnaður, jafnvel á kostnað samfélagslegrar ábyrgðar og virðingr fyrir mannréttindum.<sup>9</sup>

Þá hefur gervigreind haft mikil áhrif á pólitíska og samfélagslega umræðu á internetinu. Svokallaðir bottar, sem eiga að líkja eftir mannlegri hegðun, eru notaðir til margra gagnlegra hluta eins og t.d. í leitarvélum og sjálfvirkri svarþjónustu fyrirtækja og opinberra aðila, þ.e. vélmenni sem svara spurningum fólks. En þeir hafa einnig verið nýttir til þess að dreifa falsfréttum og skapa glundroða og óvissu. Vegna þess hve líkir þeir eru raunverulegu fólki er erfitt fyrir almenning að greina hvort um ræðir. Það er ljóst að þetta bitnar frekar á hópum sem skortir færni og þjálfun til þess að beita gagnrýnni hugsun, t.a.m. vegna fötlunar og/eða félagslegrar eða efnahagslegrar mismununar.<sup>10</sup>

Annað dæmi er notkun tauganetstækni þegar kynhneigð fólks var greind með því að safna og greina myndir af stefnumótasíðum sem sýndi að tölvur höfðu meiri getu til þess að greina kynhneigð en manneskjur. Var rannsóknin framkvæmd án vitneskju og samþykkis einstaklingana sem í hlut áttu og sýnir það hvernig vanvirðing fyrir friðhelgi fólks eykur hættur

---

<sup>8</sup> Latonero. 2018:10; Kade Crockford, „How is Face Recognition Surveillance Technology Racist?“ *ACLU*, 16. júní 2020. Sótt af <https://www.aclu.org/news/privacy-technology/how-is-face-recognition-surveillance-technology-racist/>

<sup>9</sup> Latonero. 2018:10-11.

<sup>10</sup> Latonero. 2018:12.

á að eftirlit með borgurum með notkun á gervigreind sé beitt. Þetta skapar mikla hættu fyrir þennan jaðarsetta hóp fólks, sér í lagi í ríkjum þar sem réttindi hinsegin fólks eru fótum troðin.<sup>11</sup>

ACLU, samtök um borgaraleg réttindi í Bandaríkjunum, létu rannsaka andlitsgreiningarbúnað stórfyrirtækisins Amazon árið 2018. Þar voru skoðaðar myndir af 535 þingmönnum í bandaríska þinginu og þær bornar saman við 25 þúsund myndir af einstaklingum sem höfðu verið handteknir. Rannsóknin sýndi að kerfin bjuggu ekki aðeins til 28 rangar samsvaranir heldur voru þau einnig hlutdræg þegar kom að húðlit fólks. Amazon hafði selt þennan búnað til lögreglunnar. ACLU hefur lýst yfir áhyggjum af eftirliti hins opinbera með notkun andlitsgreingar, sem er umfangsmikið, ógagnsætt og engin lög og reglugerðir gilda um.<sup>12</sup>

Þá var það mat hollensks dómsstóls að eftirlitskerfi með bótaþegum hafi verið brot á mannréttindum þar sem fátækt og langveikt fólk var beitt óhóflegu eftirliti og brotið var einkalífsréttindum og friðhelgi þess þegar tækni af þessu tagi var nýtt til þess að reyna að uppræta bótasvik. Slíkt eftirlit þekkist víða um heim og hefur hvorki fengið nægilega umfjöllun né hefur rammi verið mótaður um notkun gervigreindar í þessu samhengi.<sup>13</sup>

Þetta eru allt dæmi, úr alþjóðlegu samhengi, um hvernig gervigreind er notuð og henni misbeitt. Þessi notkun og misbeiting hefur grafalvarlegar afleiðingar fyrir jaðarsetta hópa þar sem staðsetning, notkun samfélagsmiðla, upplýsingar um heilsu, stjórnmalaskoðanir og annars konar gögn eru notuð gegn fólki. Alþjóðleg mannréttindalög geta hjálpað þeim sem þróa gervigreind að greina, bera kennsl á og bregðast við hættu sem stafar af tækni þeirra. Þessi hætta snýr meðal annars að söfnun gagna um einstaklinga og hópa, sem ógnar friðhelgi og öryggi einstaklinga.<sup>14</sup>

---

<sup>11</sup> Latonero. 2018:13.

<sup>12</sup> Latonero. 2018:13.

<sup>13</sup> Jon Henley og Robert Booth, „Welfare surveillance system violates human rights, Dutch court rules,“ *The Guardian*. 5. febrúar 2020. Sótt af <https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules>.

<sup>14</sup> Latonero. 2018:14.

## Áhættuþættir fyrir fatlað fólk

- Atvinnumál:
  - Fatlað fólk er þegar jaðarsett á vinnumarkaði. Stórfyrirtæki erlendis eru þegar farin að nýta sér gervigreind til þess að sigta út umsóknir um störf. Fatlað fólk er sérstaklega berskjaldað í þessu samhengi. Landssamtökin Þroskahjálp hafa haft samband við stórar ráðningarskrifstofur til að afla upplýsinga um notkun slíkra kerfa en slíkra upplýsinga þarf að afla með markvissum og fyrirfram skilgreindum hætti af hálfu hins opinbera. Mannréttindasamtök ein geta ekki borið ábyrgð á slíku eftirliti heldur þarf virkt samstarf og samráð þeirra, hins opinbera og aðila sem þróa og nýta gervigreind.
  - Fatlað fólk er afar berskjaldað fyrir sjálfvirknivæðingu starfa og eru þau störf sem líkleg eru til að hverfa á næstu árum og áratugum oft unnin af einstaklingum með þroskahömlun og skyldar fatlanir, t.d. afgreiðslustörf í verslunum, áfylling vara o.fl.
- Myndgreining:
  - Ef hægt er að bera kennsl á fötlun einstaklinga út frá myndefni (t.a.m. próffilmmyndir á samfélagsmiðlum) vaknar sú spurning hvort ekki sé ástæða til að ætla að fatlað fólk, og sérstaklega fólk með þroskahömlun og skyldar fatlanir, sé sérlega berskjaldað fyrir svikum á netinu þar sem fólk gæti vegna fötlunar sinnar skort færni til þess að bera kennsl á svikapósta og aðrar vefveiðar (e. phishing).
- Stafrænt aðgengi:
  - Skortur á aðgengi að tæknilausnum er mikið áhyggjuefni þrátt fyrir skýrar lagaheimildir og skuldbindingar stjórnvalda til þess að tryggja stafrænt aðgengi fólks. Þetta er áhyggjuefni sem eykst með meiri notkun tækninnar, m.a. hjá hinu opinbera. Þetta á við um notkun almennra borgara t.d. á Stafrænu Íslandi og mínum síðum hjá hinum ýmsu opinberu aðilum. Þetta býr til gjá milli ófatlaðra og fatlaðra íbúa. Þetta sýnir jafnframt að lög og reglugerðir eru ekki nóg. Oft þarf þrýsting samtaka og aðgerðarsinna til þess að ná fram réttindum og breytingum.
- Gagnaöflun:
  - Gagnaöflun getur verið gagnleg þegar kemur að því að sýna fram á og draga úr mismunun en sumir hópar eru sérstaklega berskjaldaðir fyrir gagnaöflun og miðlun gagna um þá vegna fordóma í samfélaginu. Verja þarf fatlað fólk sérstaklega í þessu smahengi. Til dæmis þegar kemur að dreifingu gagna um fötlun og veikindi í úrskurðum dómstóla sem eru aðgengilegir almenningi og geta haft neikvæð áhrif á atvinnumöguleika, að fá húsnæði og marga aðra þætti.

## Tækifæri

Ljóst er að það eru ekki einungis hættur sem stafa af notkun gervigreindar heldur eru einnig mikil tækifæri til þess að auka lífsgæði og réttindi fatlaðs fólks. Allir hlutaðeigandi aðilar (tæknigeirinn, hagsmunasamtök, fræðasamfélagið og stjórnvöld) þurfa að þróa og nýta gervigreind og skylda tækni með þeim hætti að þær stuðli að auknum tækifærum til virkrar þátttöku í samfélaginu. Tryggja þarf að hún stuðli í verki að því að fólk fái notið mikilsverðra mannréttinda og að fólk fái jafnframt notið óskertra einkalífsréttinda og friðhelgi. Þetta getur til dæmis gerst með nýjum aðferðum við að efla greiningar og meðferðir á sjúkdómum og auka aðgengi að heilbrigðisþjónustu.

Líta má svo á að Ísland hafi sérstöðu í heiminum þegar kemur bæði stöðu mannréttinda og tækniþekkingu. Það eru því mikil sóknartækifæri í að verða leiðtogar á alþjóðavísu í góðu siðferði og vernd mannréttinda við þróun og nýtingu gervigreindar, og að sýna fram á að hægt sé að byggja brýr milli fræðaheimsins, hagsmunaaðila, mannréttindasamtaka og tæknigeirans. Stjórnvöld þurfa að berjast fyrir því að jákvæð efnahagsleg, félagsleg og önnur réttindi séu eflað með tækni og vísindum. Íslensk stjórnvöld eiga einnig að hafa mikinn metnað og frumkvæði í að beita sér fyrir mannréttindum á þessu sviði á alþjóðavettvangi.

## Heimildir

*Amnesty International and Access Now.* „The Toronto Declaration: Protecting the right to equality and non-discrimination in machine learning systems“. 16. maí 2018. Sótt af [https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration\\_ENG\\_08-2018.pdf](https://www.accessnow.org/cms/assets/uploads/2018/08/The-Toronto-Declaration_ENG_08-2018.pdf)

*Australian Human Rights Commission.* „Human Rights and Technology Issues Paper, Australian Human Rights Commission.“ Desember 2019. Sótt af [https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights\\_2019\\_DiscussionPaper.pdf](https://tech.humanrights.gov.au/sites/default/files/2019-12/TechRights_2019_DiscussionPaper.pdf)

*Evrópuráðið.* „Algorithms and Human Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications.“ Mars 2018. Sótt af <https://rm.coe.int/algorithms-and-human-rights-en-rev/16807956b5>

Jon Henley og Robert Booth. „Welfare surveillance system violates human rights, Dutch court rules.“ *The Guardian.* 5. febrúar 2020. Sótt af <https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules>

Kade Crockford. „How is Face Recognition Surveillance Technology Racist?“ *ACLU.* 16. júní 2020. Sótt af <https://www.aclu.org/news/privacy-technology/how-is-face-recognition-surveillance-technology-racist/>

Mark Latonero. „Governing Artificial Intelligence: Upholding Human Rights & Dignity.“ 18. október 2018. Sótt af <https://datasociety.net/library/governing-artificial-intelligence/>

Salil Shetty. „Artificial Intelligence for Good.“ *Amnesty International.* 9. júní 2017. Sótt af <https://www.amnesty.org/en/latest/news/2017/06/artificial-intelligence-for-good/>