Reykjavík, 18. maí 2026
MSS25010120

MINNISBLAÐ

Viðtakandi: Borgarstjóri.
Sendandi: Borgarlögmaður.

Samspil gervigreindarnotkunar og persónuverndarlöggjafar í störfum sveitarfélaga

Í framhaldi af vinnufundi yfirstjórnar Reykjavíkurborgar í janúar 2025, þar sem meðal annars var haldið erindi um gervigreind, óskaði þáverandi borgarstjóri eftir því að borgarlögmaður tæki saman minnisblað um helstu atriði persónuverndarlöggjafar sem hafa þarf í huga og gera þarf ráðstafanir vegna í tengslum við áform Reykjavíkurborgar um að byrja að taka gervigreind í gagnið.

I.
Hugtakið gervigreind

Meðal þess sem tölvutæknin hefur fært fram á síðustu árum og reynt hefur á í vaxandi mæli við vinnslu persónuupplýsinga er gervigreind. Ekki er til ein almenn lagaleg skilgreining á hugtakinu gervigreind í íslenskum rétti en það hefur aftur á móti verið skilgreint í stefnu Íslands um gervigreind frá apríl 2021.[1] Þá hafa íslenskir lögfræðingar í nokkrum fræðigreindum sett fram sínar skilgreiningar á hugtakinu.[2]

Með stefnu Íslands um gervigreind hefur hugtakið gervigreind verið skilgreint á tvenns konar máta. Í fyrsta lagi má sjá eftirfarandi skilgreiningu: „Gervigreind er fræðasvið innan tölvunarfræði en tengist þó einnig fjölmörgum öðrum sviðum. Hún er víðtæk og fjölmörg reiknirit falla undir sviðið sem skiptist svo í frekari undirsvið eins og þekkingarkerfi, vitvélar og vélnám.“ Í öðru lagi er hugtakið skilgreint með einföldum hætti og í víðara samhengi sem leið okkar til að fá vélar, í víðum skilningi, til að vinna mannanna verk. Sé mannleg greind skilgreind sem hæfileikinn til að öðlast og nota þekkingu og hæfni þá er gervigreind sá hæfileiki að búa til tölvukerfi sem hefur hæfileika til að öðlast og nota þekkingu og hæfni. Þetta á sérstaklega við um verkefni sem aðeins maðurinn gat gert áður en til gervigreindar kom. Samkvæmt stefnu Íslands er gervigreind þannig annars vegar fræðasvið innan tölvunarfræði og hins vegar leið eða hæfileiki mannfólksins til að skapa vélar til að framkvæma mannlega vinnu sem krefst þekkingar og hæfni, einkum í tengslum við verkefni sem einungis voru á færi manna að vinna áður en mennirnir sköpuðu slíkar vélar.

---

[1] Um er að ræða samantekt nefndar sem falið var að rita gervigreindarstefnu Íslands og skila tillögum að skýrri framtíðarsýn um hvernig íslenskt samfélag geti unnið með gervigreind öllum til hagsbóta, sjá hér: https://www.stjornarradid.is/library/01--Frettatengt---myndir-og-skrar/FOR/Fylgiskjol-i-frett/Stefna%20%c3%8dslands%20um%20gervigreind

[2] Lára Herborg Ólafsdóttir og Sindri M. Stephensen: „Gervigreind og höfundaréttur“. Tímarit lögfræðinga, 2. tbl. 2019, bls. 159-170. Í greininni setja þau fram eftirfarandi skilgreiningu á hugtakinu gervigreind: „Þau vísindi að láta tölvur eða forrit framkvæma hluti sem krefjast myndu vitsmunalegrar þekkingar, eða greindar, ef einstaklingur eða vitsmunavera framkvæmdi slíka hluti, svo sem að taka ákvarðanir byggðar á skynsemisrökum, að alhæfa um eitthvað eða læra af fyrri reynslu.“

II.
Lagalegt umhverfi gervigreindar

Í gildi eru hér á landi lög nr. 90/2018 um persónuvernd og vinnslu persónuupplýsinga en með þeim var innleidd reglugerð Evrópusambandsins um sama efni sem ætlað er að styrkja stjórnarskrárvarin réttindi borgaranna til friðhelgi einkalífs, meðal annars með hliðsjón af þeirri tæknibyltingu sem nú á sér stað. Í lögunum er ekki vikið sérstaklega að gervigreind og hvaða kröfur séu gerðar til vinnslu persónuupplýsinga í gegnum hana. Einungis er tekið fram í lögskýringargögnum að eftirlits- og ráðgjafahlutverk Persónuverndar muni aukast verulega meðal annars vegna nýrra og krefjandi áskorana tengdum vinnslu persónuupplýsinga á borð við gervigreind, sjálflærandi vélar o.fl.

Reglugerð Evrópusambandsins um gervigreind nr. 2024/1689 tók gildi 1. ágúst 2024 og verður hún innleidd í áföngum fram til 2. ágúst 2027. Reglugerðinni er ætlað að samræma reglur um gervigreind innan Evrópusambandsins og greiða leiðina bæði fyrir tæknilega þróun og nýsköpun á sviði gervigreindar. Enn á eftir að taka reglugerðina upp í EES-samninginn áður en hún verður hluti af íslenskum lögum. Íslensk stjórnvöld, stofnanir og fyrirtæki þurfa því að svo stöddu fyrst og fremst að vera meðvituð um ákvæði reglugerðarinnar ef fyrirhuguð er notkun á gervigreind og greina þá hvort hún sé að öllu leyti samrýmanleg og skoða gæði þeirra gagna sem eru notuð af slíkri tækni, meðal annars með hliðsjón af persónuvernd. Reglugerðin beitir áhættumiðaðri nálgun á mismunandi gervigreindarkerfi, þar sem kerfi með óásættanlega áhættu eru bönnuð og kerfi með mikla áhættu eiga að vera undir ströngu eftirliti. Gervigreindarkerfi sem hafa í för með sér takmarkaða áhættu verða að uppfylla tilteknar gagnsæiskröfur en engar kröfur eru gerðar til kerfa sem ekki falla undir neinn af þessum flokkum þar sem til staðar er þá lágmarksáhætta.

III.
Samspil gervigreindar og persónuverndarlöggjafar

Samspil gervigreindar og persónuverndarlöggjafar einkennist þannig af því að gervigreindarkerfi verða að fylgja almennum reglum um vinnslu persónuupplýsinga samkvæmt lögum nr. 90/2018 auk þess sem gervigreindarreglugerð Evrópusambandsins mun leggja viðbótarskyldur á slík kerfi þegar hún verður innleidd að fullu. Meginreglur persónuverndar, svo sem meðalhóf og lágmörkun gagna, eiga við um alla þróun og notkun gervigreindar sem styðst við persónugreinanleg gögn. Þetta þýðir að um leið og persónuupplýsingar koma við sögu, t.d. í þjálfunargögnum, innslætti notenda í löggum eða í niðurstöðum sem tengjast einstaklingi, gilda meginreglur og skyldur samkvæmt persónuverndarlöggjöfinni. Á sama tíma setur gervigreindar-reglugerðin tæknilegar og skipulagslegar kröfur um áreiðanleika og áhættustýringu á ákveðna flokka gervigreindarkerfa. Í framkvæmd þarf því sama verkefni oft að standast tvö regluverk sem skarast að hluta, þ.e. annars vegar persónuverndarlög sem leggja megináherslu á lögmæti, gagnsæi og réttindi einstaklinga og hins vegar gervigreindarreglugerðina sem kveður á um öryggi, stjórnun áhættu, gæði gagna og rekjanleika kerfa.

Meginreglur persónuverndarlaga eru meðal annars sanngirnisreglan, tilgangsreglan, meðalhófsreglan, áreiðanleikareglan, varðveislureglan og öryggisreglan, sbr. a- til f-liði 1. mgr. 5. gr. persónuverndarlaga. Þá er í lögunum einnig lögð áhersla á að vinnsla persónuupplýsinga þurfi að vera nauðsynleg, þ.e. að önnur leið sé almennt ekki fær, sbr. 1.-6. tölulið 1. mgr. 9. gr. laganna. Sá aðili sem vinnur með persónuupplýsingar í gegnum gervigreindarforrit þarf því að ganga úr skugga um að meginreglum persónuverndarlaga sé fylgt á öllum stigum enda ber að hafa í huga að gervigreind vinnur persónuupplýsingar ólíkt á ólíkum stigum.

Sanngirnisreglan

Sanngirnisreglan felur í sér skyldu til að upplýsa hinn skráða einstakling um að verið sé að vinna með persónuupplýsingar hans. Eins og tekið er fram í vel flestum þeim gervigreindarforritum sem í notkun eru nú geta niðurstöður þeirra verið rangar eða mismunandi ef þjálfunargögnin gefa bjagaða mynd af raunveruleikanum eða ef þau tengjast ekki viðkomandi sviði. Með öðrum orðum geta gervigreindarforrit gert mistök. Slík hlutdrægni gagna getur auðveldlega leitt til mismununar og ósanngirni gagnvart hagsmunum hins skráða einstaklings.[3]

Sanngirnisreglan krefst þess að öll vinnsla fari fram með virðingu fyrir hagsmunum hins skráða einstaklings og að gögnin séu notuð í samræmi við það sem hann má með réttu búast við. Reglan krefst einnig þess að ábyrgðaraðili[4] grípi til ráðstafana til að koma í veg fyrir handahófskennda mismunun á einstaklingum. Notkun viðeigandi stærðfræðilegra eða tölfræðilegra aðferða eru mögulegar ráðstafanir en það eitt og sér er þó almennt ekki talið nægjanlegt til að tryggja fylgni við sanngirnisregluna. Gervigreind þarf að vera þjálfuð á viðeigandi, réttum og fjölbreyttum gögnum svo hún geti lært að bera kennsl á mikilvægar upplýsingar og mynstur í gagnasafninu. Það þýðir að hún má ekki leggja áherslu á upplýsingar sem tengjast kynþætti eða þjóðernisuppruna, stjórnmálaskoðunum, trúar- eða lífsskoðunum, aðild að stéttarfélagi, erfðafræðilegri stöðu, heilsufari eða kynhneigð ef getur leitt til handahófskenndrar mismununar. Ef því verður haldið fram að notkun gervigreindar leiði til ósanngjarnra eða mismunandi niðurstaðna getur Persónuvernd rannsakað hvort sanngirnisreglunni hafi í raun verið fylgt við vinnslu persónuupplýsinga.

Þegar vinnsla persónuupplýsinga fer fram með notkun gervigreindar er mikilvægt að gagnsæi sé einnig tryggt. Ákveðin áskorun felst í því þar sem erfitt getur verið að skilja gervigreindarkerfi sem eru háþróuð kerfi auk þess sem svokallaður svartur kassi (e. black box) þeirra gerir það afar erfitt að útskýra hvernig upplýsingar eru vegnar og metnar í einstaka tilvikum. Með hugtakinu „svarti kassinn“ er átt við að algengt sé að notendur gervigreindar og þeir sem verða fyrir áhrifum hennar hafi ekki nægjanlegar upplýsingar um hvernig gervigreindin vinnur úr fyrirliggjandi gögnum. Skýrist þessi óvissa einkum af eðli gervigreindar og hvernig hún er uppbyggð til að nýta vélnám og algrím.[5] í vinnslu sinni á gögnum til að komast að niðurstöðu með sjálfvirkum hætti. Þrátt fyrir hátt flækjustig gildir meginreglan um gagnsæi og því er það ófrávíkjanlegt skilyrði að ábyrgðaraðilar uppfylli grundvallarreglur persónuverndar um fræðsluskyldu og tryggi rétt hins skráða til að fá staðfestingu frá ábyrgðaraðila um það hvort hann vinni með persónuupplýsingar hans, sbr. 17. gr. persónuverndarlaga.

Tilgangsreglan

Tilgangsreglan felur í sér að safna verður persónuupplýsingum í skýrum og málefnalegum tilgangi. Persónuupplýsingum má til að mynda ekki safna vegna þess eins að þær kunni að koma að notum síðar meir. Þá er almennt óheimilt að nota persónuupplýsingar í öðrum tilgangi en upphaflega stóð til. Með áskilnaði um skýran tilgang er meðal annars tekið fram í lögskýringargögnum að átt sé við að hann skuli vera nægjanlega vel afmarkaður og skilgreindur til að vinnsla sé auðskilin og gagnsæ auk þess sem koma skuli í veg fyrir víðtækan tilgang til að fella megi undir hann nánast hvað

---

[3] Með tilvísun til hins skráða er átt við persónugreindan eða persónugreinanlegan einstakling, sbr. 2. tölulið 1. mgr. 3. gr. persónuverndarlaga. Einstaklingur telst persónugreinanlegur ef unnt er að persónugreina hann, beint eða óbeint, svo sem með tilvísun í auðkenni eins og nafn, kennitölu, staðsetningargögn, netauðkenni eða einn eða fleiri þætti sem einkenna hann í líkamlegu, lífeðlisfræðilegu, erfðafræðilegu, andlegu, efnalegu, menningarlegu eða félagslegu tilliti,

[4] Ábyrgðaraðili er einstaklingur, lögaðili, stjórnvald eða annar aðili sem ákveður einn eða í samvinnu við aðra tilgang og aðferðir við vinnslu persónuupplýsinga, sbr. 6. tölulið 1. mgr. 3. gr. persónuverndarlaga.

[5] Forskrift eða lýsing sem segir hvernig leysa megi tiltekið reiknivandamál.

sem er. Slíkt er enda ósamrýmanlegt sjónarmiðum um persónuvernd. Því viðkvæmari persónuupplýsingar sem unnið er með og því meiri afleiðingar sem notkun þeirra getur haft í för með sér þeim mun mikilvægara er að tilgangurinn sé skýrt afmarkaður.[6]

Meðalhófsreglan
Meðalhófsreglan felur í sér að persónuupplýsingar þurfa að vera nægilegar, viðeigandi og takmarkaðar við nauðsyn vinnslunnar, sbr. 3. tölulið 1. mgr. 8. gr. persónuverndarlaga. Fyrrnefnd skilyrði hafa verið nefnd, „lágmörkun gagna". Reglan skarast við framangreindar tilgangsreglur en mat á skilyrðum meðalhófsreglunnar byggist á tilgangi vinnslunnar. Til þess að skilyrði meðalhófsreglunnar séu uppfyllt þarf ábyrgðaraðili að tryggja að persónuupplýsingar séu ekki varðveittar lengur en nauðsynlegt er. Það krefst þess að settur sé tímafrestur varðandi endurskoðun og/eða eyðileggingu gagnanna. Af framangreindu leiðir því að meðalhófsreglunni er ætlað að tryggja að vinnsla persónuupplýsinga gangi ekki lengra en þörf krefur til þess að markmiði vinnslunnar sé náð.

Áreiðanleikareglan
Sú regla felur í sér að persónuupplýsingar þurfa að vera áreiðanlegar og uppfærðar eftir þörfum, sbr. 4. tölulið 1. mgr. 8. gr. persónuverndarlaga. Séu persónuupplýsingar óáreiðanlegar ber að leiðrétta þær eða að öðrum kosti eyða þeim þá þegar. Við mat á því hve fljótt ábyrgðaraðila ber að leiðrétta eða eyða persónuupplýsingum þarf meðal annars að horfa til efnis og vinnslu persónuupplýsinganna.

Áreiðanleiki innan gervigreindar vísar að meginstefnu til þess hversu oft gervigreind giskar á rétt svar samanborið við prófunargögn. Í mörgum tilvikum má álykta að svör eða úttök gervigreindarinnar séu persónugögn og því skiptir máli að greina á milli ólíkra skilgreininga hugtaksins áreiðanleika á grundvelli persónuverndarlaga og í skilningi gervigreindar. Hafa þarf í huga að áreiðanleikaregla persónuverndarréttar nær til allra persónugagna sem unnin eru, þar með talið persónugagna sem verða til við notkun gervigreindar, hvort heldur sem er inntak, þjálfunargögn eða niðurstöður. Reglan felur þó ekki í sér kröfu um fullkomna eða 100% nákvæmni. Þess í stað ber ábyrgðaraðila að gera viðeigandi og sanngjarnar ráðstafanir til að tryggja að persónugögn séu ekki röng eða villandi í ljósi tilgangs vinnslunnar. Sérstaklega skiptir þetta máli þegar niðurstöður gervigreindar eru lagðar til grundvallar ákvörðunum um einstaklinga enda geta ónákvæm gögn leitt til brota gegn meginreglunni um lögmæti og sanngirni.

Varðveislureglan
Varðveislureglan felur í sér að persónuupplýsingar þurfa að vera varðveittar á sérstöku formi, sbr. 5. tölulið 1. mgr. 8. gr. persónuverndarlaga. Áskilnaðurinn um form felur í sér að það má ekki vera hægt að bera kennsl á skráðan einstakling lengur en þörf krefur. Í því samhengi getur skipt máli hvort að persónuupplýsingar séu á persónugreinanlegu formi en telja má að með því að tryggja að formið sé ópersónugreinanlegt aukist svigrúmið bæði hvað varðar geymslu og geymslutíma. Við mat á þeim tímaramma er meðal annars horft til tilgangsreglunnar sem og aðstæðna hverju sinni.

Öryggisreglan
Persónuupplýsingar þurfa að vera unnar með þeim hætti að öryggi þeirra sé tryggt, sbr. 6. tölulið 1. mgr. 8. gr. persónuverndarlaga. Umrædd meginregla endurspeglast víða í persónuverndarlögunum en vinnsla persónuupplýsinga á að vera með þeim hætti að

[6] Alþingistíðindi, 2017-2018, A-deild, bls. 5176.

trúnaður sé hafður í fyrirrúmi og að viðeigandi öryggi sé tryggt. Það felur í sér áskilnað þess efnis að ábyrgðaraðili komi í veg fyrir ólögmætan aðgang sem og nýtingu persónuupplýsinga.

Ábyrgðarskyldan
Ábyrgðarskyldan felur í sér að ábyrgðaraðili beri ábyrgð á því að vinnsla persónuupplýsinga uppfylli meginreglur 1. mgr. 8. gr., sbr. 2. mgr. sömu greinar persónuverndarlaga. Ábyrgðaraðilar bera því ábyrgð á ákvörðunum gervigreindar, þar sem einstaklingar eiga rétt til mannlegrar íhlutunar við sjálfvirka ákvörðunartöku í öllum tilvikum nema ef til staðar er lagaheimild sem kveður á um annað.

IV.
Samantekt

Af því sem hér hefur verið rakið er ljóst að þegar Reykjavíkurborg ákveður að notast við gervigreind sem tilgang og aðferð við vinnslu persónuupplýsinga haggast ekki sú grundvallarskylda hennar sem ábyrgðaraðila að virða meginreglur persónuverndarlöggjafarinnar. Það má heldur ekki gleymast að Reykjavíkurborg er ábyrgðaraðili þegar gervigreind er notuð í stjórnsýslu óháð því hvort kerfið sé þróað af þriðja aðila, keypt sem þjónusta eða sé notað í tilraunaskyni. Vinnsla persónuupplýsinga með gervigreind þarf því ávallt að byggja á skýrum lagaheimildum í sérlögum, stjórnsýslulögum eða öðrum lögmætum grundvelli. Það ber því að tryggja að einstaklingar/íbúar séu upplýstir um að gervigreind sé notuð, viti í hvaða tilgangi hún er notuð og hvaða áhrif notkunin getur haft á réttarstöðu þeirra. Þetta á sérstaklega við þegar gervigreind er notuð til mats, forgangsröðunar eða greiningar, þegar sjálfvirk vinnsla styður við ákvarðanir starfsmanna eða þegar íbúar eiga í samskiptum við sjálfvirk kerfi, t.d. spjallmenni.

Notkun gervigreindar í ákvarðanatöku sveitarfélags þarf því að samrýmast ákvæðum persónuverndarlaga og meginreglum stjórnsýsluréttar. Í því felst meðal annars að óheimilt er að byggja stjórnvaldsákvörðun eingöngu á sjálfvirkri vinnslu. Tryggja þarf raunhæfa mannlega íhlutun. Þá þarf skráður einstaklingur að geta fengið skýringar á forsendum ákvörðunar og notið þar af leiðandi réttlátrar málsmeðferðar. Telja verður að þetta hafi sérstaka þýðingu í málum er varða félagsþjónustu, húsnæðismál, skóla- og velferðarþjónustu og öðrum þeim málum þar sem ákvörðun hefur veruleg áhrif á réttarstöðu einstaklings.

Það skal tekið fram hér að Reykjavíkurborg hefur sett sér tilmæli um notkun gervigreindar sem byggja á tilteknum viðmiðum og hafa það að markmiði að tryggja að notkunin sé í takt við verkefnaáherslur, öryggiskröfur og siðferðileg gildi hjá sveitarfélaginu. Í tilmælunum er tekið fram að til þess að tryggja ábyrga og örugga notkun gervigreindar í opinberri þjónustu skuli hafa í huga að persónuvernd og öryggi gagna verða að vera í forgrunni. Þá eru framangreindar meginreglur persónuverndarlöggjafarinnar áréttaðar sem og þeir áhættuflokkar notkunar gervigreindar sem kveðið er á um í gervigreindarreglugerðinni sem notandi þarf að þekkja.

F.h. borgarlögmanns,

Þórður Guðmundsson
lögmaður

Reykjavík