

# Réttlát umskipti á gervigreindaröld

---

*33. þing Landssambands íslenskra verzlunarmanna*

*19. október 2023*

Victor Karl Magnússon, sérfræðingur hjá VR

# Aðkoma launafólks að tæknibreytingum

---



- Fjórða iðnbyltingin er hafin og mun hafa víðtæk áhrif á vinnumarkaði framtíðar.
- Gervigreind gegnir þar lykilhlutverki og mun leiða til sjálfvirknivæðingar á ýmsum sviðum.
- Mikilvægt fyrir verkalyðsfélög að gæta hagsmuni launafólks á tímum samfélagsbreytinga.
- En með hvaða hætti?

# Mögulegar áskoranir

---

- Hvernig eigum við að þjálfa starfsfólk til þess að takast á við tæknibreytingar og nýjan veruleika?
- Hvernig grípum við þann hóp sem missir vinnuna í kjölfar sjálfvirknivæðingar?
- Hvernig eigum við að móta regluverk í kringum gervigreind? Hvað má og má ekki nota gervigreind í?
- Hvernig tryggjum við að ávinningur tækniframfara skili sín launafólks?

# Stefna VR/LÍV – úr kröfugerð 2023

---

The logo consists of the letters 'VR' in a white, bold, sans-serif font, centered within a dark blue square.

”VR og LÍV vilja að vinnuvika starfsfólks stýttist enn frekar án launaskerðingar. Fyrirtæki nýti tækifæri sem skapast í kjölfar hagræðingar sem sjálfvirknivæðing og tækniframfarir stuðla að til þess að stytta vinnutíma starfsfólks.“

## 1) Hvað er gervigreind?

- Sígild gervigreind og vélarnám.
- Generatív gervigreind og ChatGPT.

## 2) Hvaða áhrif mun gervigreind og aðrar tæknibreytingar hafa á samfélagið?

- Framtíðarstörf og sjálfvirknivæðing.

## 3) Hvernig getum við nýtt þessa tækni af skynsemi?

- Gervigreind og gagnrýnin hugsun.
- Réttur til gagnsæis og útskýringar.
- Tæknibreytingar og stytting vinnuvikunnar.

# Uppruni gervigreindar



Alan Turing (1912-1954)

- Alan Turing birti grein í heimspekitímaritinu *Mind* árið 1950 þar sem hann velti því fyrir sér hvort, og þá hvernig, tölvur gætu „hugsað“.
- Árið 1956 var haldin ráðstefna við Dartmouth háskóla þar sem tölvunarfræðingar og stærfræðingar veltu sér upp úr hugmyndum Turing og lögðu til að þær yrðu rannsakaðar sem sérstök fræðigrein, „gervigreind“.
- Gervigreind hefur það markmið að þróa tölvur sem geta leyst flókin verkefni sem virðast krefjast vitneskju, dómgreindar eða innsæis af einhverju tagi.
- Sú rannsóknaráætlun sem var sett af stað árið 1956 varði í nokkra áratugi en afrakstur hennar var hin svokallaða **sígilda gervigreind (e. classical AI, good-old-fashioned AI)**.
- Sígild gervigreind er frábrugðin þeirri gervigreind sem er algengust í dag, sem styðst við svokallað **vélarnám (e. machine learning)**.

# Sígild gervigreind og Deep Blue

- Sígild gervigreind byggir á þeirri hugmynd að hægt sé að þjálfra tölvur til alls kyns verka með því að **mata þær með staðreyndum og almennum reglum** sem þær geta beitt til þess að draga ályktanir.
- Sígilt gervigreindartól sem spilar skák þarf þannig að mata með staðreyndum um mannganginn en líka reglum sem hjálpar tölvunni að velja leiki sem eru vænlegir til sigurs.
- Deep Blue tölvan frá IBM er eitt þekktast afrek sígildrar gervigreindar en hún sigraði heimsmeistarinn Gary Kasparov í þekktu einvígi árið 1997.
- Sígild gervigreindartól eins og Deep Blue eru með **harðkóðuð viðbrögð** við öllum möguleikum. Heill her af skákmönnum og tölvunarfræðingum kenndu Deep Blue hvaða leiki ætti að velja í hvaða kringumstæðum og stóluðu síðan á að reiknigeta tölvunnar myndi leiða hana til sigurs.



Gary Kasparov teflir gegn Deep Blue, 1997

# Vélarnám og AlphaGo

- Sígild gervigreind hentaði ágætlega til þess að spila skák en tækni rakst á veggum annars staðar. Það er ómögulegt að harðkóða reglur sem segja til um hvernig tölva eigi að bregðast við öllum hugsanlegum aðstæðum, sérstaklega þegar leysa á verkefni sem reyna á „mannlega skynsemi“.
- Ný nálgun í gervigreind varð því til upp úr 1980 sem er allsráðandi í dag – **vélarnám (e. machine learning)**.
- Hugmyndin með vélarnámi er að leyfa tölvunni að **læra hluti á eigin spýtur** í stað þess að mata hana með endalausum reglum (sbr. **AlphaGo**).
- Þessi tækni veltur á því að hafa aðgang að gífurlega stórum **gagnasöfnum** sem tölvan getur skoðað til þess að bera kennsl á tölfræðileg mynstur.



Lee Sodol spilar gegn AlphaGo tölvunni frá DeepMind, 2016



# Generatív gervigreind

---

- Generatív gervigreind (e. *generative AI*) getur framleitt **nýtt efni** eftir pöntun, t.d. texta, myndir og myndbönd.
- Chat GPT (texti) og Dalle (myndir) eru dæmi um generatíva gervigreind sem fyrirtækið OpenAI hefur þróað í samstarfi við Microsoft.
- Tæknin byggir á **vélarnámi**, þ.e. á tölfræðilegum reikniritum sem taka stór gagnasöfn og framleiða síðan „ný“ gögn í safnið.

# DALL-E



Þessi mynd var búin til á nokkrum sekúndum með því að slá eftirfarandi beiðni inn í prufu-útgáfu af Dall-E:

*Oil painting of union workers marching on May 1<sup>st</sup>, use blue and white colours.*

# Chat GPT

VI Can you find a few peer-reviewed studies on the effects remote work has on worker performance and summarize them for me?

Sure, here are summaries of a few peer-reviewed studies on the effects of remote work on worker performance:

1. Study: "Does Working from Home Work? Evidence from a Chinese Experiment" (2015)

This study, published in the Quarterly Journal of Economics, examined the impact of remote work on worker performance in a Chinese travel agency. The researchers found that remote workers were 13% more productive than their office-based counterparts, primarily due to fewer distractions and fewer breaks.

2. Study: "Telecommuting and the Role of Supervisory Power" (2018)

This study, published in the Journal of Business and Psychology, examined the impact of telecommuting (remote work) on worker performance and found that it was positively related to perceived autonomy and flexibility, as well as supervisor support. However, the authors noted that the effect of telecommuting on performance was moderated by supervisory power, suggesting that managers play a critical role in ensuring remote workers are productive.

3. Study: "Work-Life Balance and Telework: A Cross-Cultural Study of Four Countries" (2016)

This study, published in the Journal of Business and Psychology, examined the impact of telework (remote work) on work-life balance and job performance in four countries (the United States, India, Germany, and China). The researchers found that telework was positively related to work-life balance, which in turn was positively related to job performance. However, the effect of telework on performance was stronger in individualistic cultures (such as the United States) compared to collectivistic cultures (such as China).

Overall, these studies suggest that remote work can have a positive impact on worker performance, especially when it is supported by managers and supervisors and when workers have the autonomy and flexibility they need to be productive. However, the effects of remote work may vary across different cultures and industries, and additional research is needed to fully understand the relationship between remote work and performance.

Chat GPT getur skrifað ljóð, forrit, minnisblöð, o.sfrv.

Líkanið hefur aðgang að upplýsingum á internetinu og getur nýtt sér þær til að skrifa sérhæfðan texta og leysa verkefni, t.d. nær Chat GPT samræmdu prófi sem veitir lögmannsréttindi í Bandaríkjunum (e. *bar exam*).

Mállíkön eins og Chat GPT eru líka sögð breyta stöðu leitarvéla.

Í stað þess að slá inn leitartexta á Google eða Google Scholar til þess að finna rannsóknir um áhrif fjarvinnu á framleiðni, þá skilar Chat GPT nokkrum rannsóknum ásamt samantektum á örfáum sekúndum í dæminu til hliðar.

- 1) Hvað er gervigreind?
- 2) **Hvaða áhrif mun gervigreind og aðrar tæknibreytingar hafa á samfélagið?**
  - Framtíðarstörf og sjálfvirknivæðing.
- 3) Hvernig getum við nýtt þessa tækni af skynsemi?

# Goldman Sachs, 2023

---

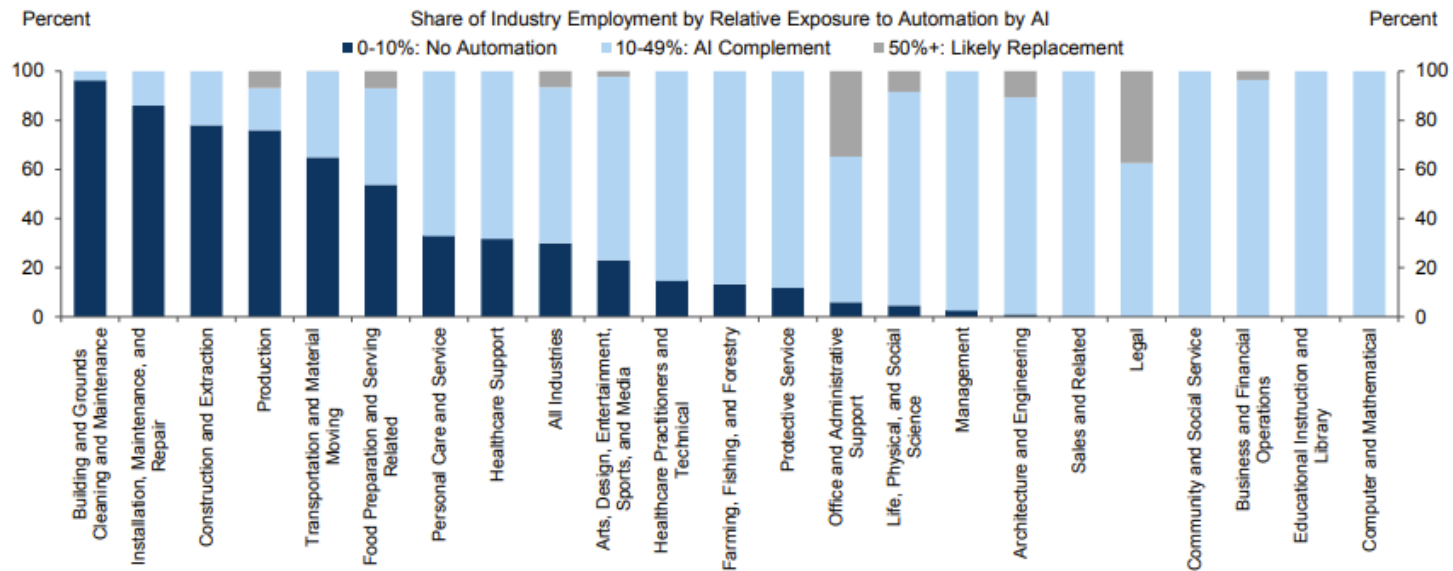
- Hagfræðingar bandaríska fjárfestingabankans Goldman Sachs birtu rannsóknarskýrslu í lok mars 2023 um áhrif gervigreindar á vinnumarkaðinn. Spá til 10 ára:
  - 67% starfa í Bandaríkjunum og Evrópu gætu orðið fyrir **einhvers konar** sjálfvirknivæðingu af sökum gervigreindar.
  - 300 milljónir starfa á heimsvísu gætu horfið en ný verða til.
  - 7% hækkun GDP á heimsvísu.
  - Skrifstofustörf líklegri til þess að verða fyrir áhrifum en önnur.
  - Iðnvæddustu lönd heims líklegri til þess að verða fyrir áhrifum en önnur.
  - **Generatív gervigreind** mikilvæg.

# Vinna á tímum gervigreindar

- Generatív gervigreind hefur mesta þýðingu fyrir **skrifstofufólk** og **sérfræðinga** – hentar best fyrir afmörkuð verkefni þar sem árangur er mælanlegur.
- Hagfræðingar Goldman Sachs spá því að **46%** verkefna almenns skrifstofufólks gætu verið leyst af gervigreind og **44%** verkefna sem tengjast lögfræðilegri starfsemi.
- Áhrifin verða miklu minni á störf sem krefjast líkamlegrar vinnu, sem er áhugavert frávik frá tækniframförum fyrri tíma.
  - Það er auðveldara að kenna tölvu að verða stórmeistari í skák en að hella upp á kaffi (sbr. Smith 2019).

# Áhrif eftir starfsgreinum

**Exhibit 8: Replacement in Legal and Administrative Fields, Little Effect in Manual and Outdoor Jobs, and Productivity-Enhancement Everywhere Else**



Source: Goldman Sachs Global Investment Research

# Störf sem verkefnabúnt

- Störf felast sjaldnast í því að leysa eitt tiltekið verkefni aftur og aftur. Það væri nákvæmara að lýsa störfum sem **búnt verkefna**.
  - Bóksali þarf t.d. að afgreiða á kassa, veita ráðgjöf varðandi bækur, panta bækur inn í búðina, huga að útstillingum, vera í samskiptum við höfunda og útgefendur, skipuleggja viðburði o.fl.
- Gervigreind mun sjaldnast leysa störf algjörlega af hólmi en frekar **ákveðna þætti** í verkefnabúntum þeirra (Brynjolfsson 2022).
- Sumt starfsfólk verður afkastameira og skilvirkara, annað gæti misst vinnuna (e. *augmentation vs. automatation*).



# Ný störf vega á móti atvinnumissi

---

- Tækniframfarir leysa ákveðin verkefni af hólmi en skapa líka ný störf – 60% fólks á vinnumarkaði eru í starfi sem var ekki til árið 1940 (Brynjolfsson 2022, bls. 280).
- Þessi áhrif munu milda áhrif sjálfvirknivæðingar en verkalýðsfélögin þurfa að vera vakandi fyrir því hversu **hratt** þessar breytingar ganga í gegn og hvort að það sé **sami hópur** sem missir störf og fái ný.

- 1) Hvað er gervigreind?
- 2) Hvaða áhrif mun gervigreind og aðrar tæknibreytingar hafa á samfélagið?
- 3) **Hvernig getum við nýtt þessa tækni af skynsemi?**
  - Gervigreind og gagnrýnin hugsun.
  - Réttur til gagnsæis og útskýringar.
  - Tæknibreytingar og stytting vinnuvikunnar.

# Regluverk og kjarasamningar



# Réttur til útskýringar

- Vélarnám er oft ógegnsætt – stundum veit engin hvers vegna gervigreindin tekur eina ákvörðun frekar en aðra.
- Í sumum aðstæðum skiptir **öllu máli** að vita hvers vegna ein ákvörðun var tekin, til dæmis þegar kemur að ákvörðunum stjórnvalda og dómstála, ráðningar í störf, veitingu lána, o.sfrv.
- Frægt dæmi um dulda fordóma í ráðningartóli sem Amazon notaði árið 2014.
- Vaxandi kröfur um **rétt fólks til útskýringar** á ákvörðunum sem snerta líf þeirra.
- Verkalýðsfélög ættu að íhuga hvenær tryggja ætti rétt launafólks í þessu samhengi.

# Gagnrýnin hugsun

---

- Gervigreind hefur náð merkilegum árangri á ýmsum sviðum en hefur engu að síður oft rangt fyrir sér.
- Starfsfólk sem hyggst nýta sér gervigreind ætti því ekki að stóla á tæknina í blindi heldur tileinka sér **gagnrýna hugsun** við meðferð hennar.
- Páll Skúlason skilgreindi gagnrýna hugsun sem þá „hugsun sem fellst ekki á neina skoðun eða fullyrðingu nema hún hafi fyrst rannsakað hvað í henni felst og fundið fullnægjandi rök fyrir henni“.
- Skapar vandamál fyrir vélarnám sem er oft ógegnsætt – við vitum t.d. ekki hvers vegna AlphaGo velur einn leik frekar en annan.

# Lögfræðingur í vanda

The New York Times

## *Here's What Happens When Your Lawyer Uses ChatGPT*

A lawyer representing a man who sued an airline relied on artificial intelligence to help prepare a court filing. It did not go well.

Give this article



1.1K



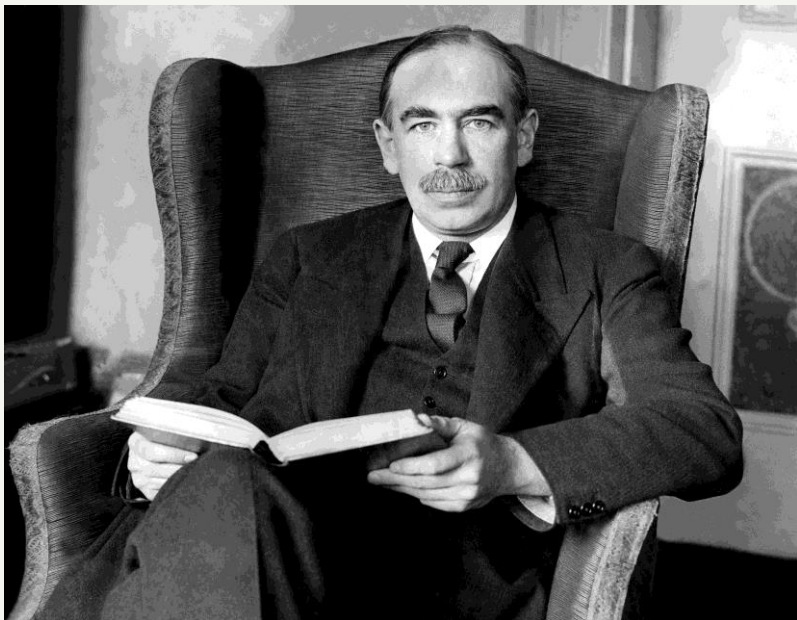
- Það eru mörg dæmi um það að gervigreind geri mistök eða skáldi hreinlega upp svör við tilteknum spurningum.
- Hættulegt að treysta tækninni í blindni, eins og nýlegt dæmi um hrakfarir bandaríks lögfræðings sýnir.
- Lögfræðingurinn notaði ChatGPT til þess að skrifa dómsskjöl, þar sem vísað var til dóma um sambærileg mál. Seinna meir kom í ljós að gervigreindin **skáldaði** upp þessi dómsmál.
- Sýnir að áfram verður þörf á starfsfólki með sérhæfða þekkingu sem getur lagt mat á gæði þess efnis sem gervigreindin framleiðir.

# Skipting verðmæta

- Gervigreindarbyltingin mun skapa mikil verðmæti – hagfræðingar Goldman Sachs spá því að GDP hækki um 7% á heimsvísu á næstu 10 árum í tengslum við tækniframfarar á sviði gervigreindar.
- Hvernig skiptum við þeim verðmætum? Hvernig getum við styrkt stöðu launafólks á vinnumarkaði framtíðar? Ein leið er að setja tækniframfarir í samhengi við kjaramál eins og **styttingu vinnuvikunnar**, sbr. stefnu VR og LÍV í kröfugerð frá 2023:

*„VR og LÍV vilja að vinnuvika starfsfólks styttist enn frekar án launaskerðingar. Fyrirtæki nýti tækifæri sem skapast í kjölfar hagræðingar sem sjálfvirknivæðing og tækniframfarir stuðla að til þess að stytta vinnutíma starfsfólks.“*

# Hvernig verður vinnuvikan eftir 100 ár?



John Maynard Keynes (1883-1946)

- Breski hagfræðingurinn John Maynard Keynes birti grein árið 1930 þar sem hann spáði því að vinnuvikan yrði 15 stundir árið 2000.
- Rök Keynes voru einföld: framleiðni hlyti að stórukast á næstu áratugum í kjölfar tækniframfara og sjálfvirknivæðingar.
- Þannig þyrfti alltaf æ minni tíma til þess að framleiða nóg til þess að viðhalda lífsgæðum og hagvexti.
- Keynes hafði rétt fyrir sér m.t.t. framleiðni og tækniframfara en ofmat styttingu vinnuvikunnar.
- Það er hins vegar spurning hvað næstu 100 ár bera í skauti sér.



- Áhrif gervigreindar verða mest hjá skrifstofufólki og sérfræðingum en það er erfitt að búa til sannfærandi tölulegar spár vegna skorts á viðeigandi gögnum um hæfnipætti.
- Mörg störf gætu glatast en flest munu einfaldlega **breytast** þegar gervigreind tekur við ákveðnum verkefnum en ekki öðrum (e. *automation vs. augmentation*).
- Þau sem geta **unnið með** gervigreindinni til þess að auka framleiðni sína verða í sterkri samningastöðu til framtíðar.
- Nauðsynlegt er að setja tækniframfarir í samhengi við kjaramál eins og styttingu vinnuvikunnar – verðmætin sem gervigreind skapar þarf líka að skila sín til launafólks en ekki bara til atvinnurekenda.

# Heimildir og ítarefni

---

Briggs, Joseph and Davesh Kodani, (2023). The Potentially Large Effects of Artificial Intelligence on Economic Growth. *Goldman Sachs Economics Research*.

Brynjolfsson, E. (2022). The Turing trap: The promise & peril of human-like Artificial Intelligence. *Daedalus*, 151(2), 272-287.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, 358(6370), 1530-1534.

Frank, Morgan R., et al. (2019). Toward understanding the impact of artificial intelligence on labor. *Proceedings of the National Academy of Sciences* 116.14: 6531-6539. URL=  
<https://www.pnas.org/doi/full/10.1073/pnas.1900949116>

Forsætisráðuneyti Íslands (2019). „Ísland og fjórða iðnbýltingin.“ Skýrsla framtíðarnefndar. URL=  
<https://www.stjornarradid.is/library/04-Raduneytin/ForsAetisraduneytid/Framtidarnefnd/Fjorda-idnbyltingin-skyrsla.pdf>

Handel, M. J. (2016). The O\* NET content model: strengths and limitations. *Journal for Labour Market Research*, 49(2), 157-176.

Smith, B. C. (2019). *The Promise of Artificial Intelligence: Reckoning and Judgment*. MIT Press.