

---

# Generatiivinen tekoäly ja viettelevät visualisaatiot

---

*Anna-Mari Wallenberg*

---

ProCom – Viestinnän ammattilaiset ry  
ISBN 978-952-65488-5-2

Data- ja algoritmivetoisissa yhteiskunnissa hyödynnetään massiivisiin aineistoihin perustuvia tilastoja, simulaatioita ja ennusteita jo lähes kaikilla yhteiskunnan osa-alueilla. Datavisualisaatioiden eli kaavioiden, kuvituksien ja diagrammien viestinnällinen rooli on koko ajan keskeisempi muun muassa mediassa, tieteessä ja poliittisessa päätöksenteossa. Niiden avulla havainnollistetaan ja viestitään datapohjaisia ilmiöitä.

Generatiivinen tekoäly avaa mahdollisuuksia uuden sukupolven entistä tehokkaammille, ilmaisuvoimaisimmille ja uskottavimmille visualisaatioille. Se myös automatisoi monia visualisoinnin työläitä välivaiheita. Samalla tekoäly monimutkaistaa visualisaatioiden tiedollisia ominaisuuksia tavalla, joka nostaa esiin kysymyksiä niiden viestinnällisen käytön reunaehdoista.

---

## Johdanto

Visualisaatiot, eli erilaiset graafit, diagrammit ja kaaviot, ovat keskeisiä viestimisen välineitä monilla yhteiskuntien osa-alueilla, kuten mediassa, poliittisessa päätöksenteossa ja hallinnon organisoimisessa. Visualisaatioita hyödynnetään laajalti myös esimerkiksi koulutuksessa, tieteessä ja mainosalalla (Couldry & Hepp 2017).

Visualisaatioiden merkitys korostuu erityisesti algoritmisoituneissa ja datavetoisissa yhteiskunnissa, joissa päätöksenteko usein tapahtuu saatavilla olevan datan pohjalta. Ennusteet, simulaatiot ja tilastot muodostavat viitekehyksen, jonka pohjalta keskustelemme ilmaston muutoksesta, tekoälyn vaikutuksista työpaikkoihin tai hallituksen talouspoliittisista linjauksista.

Visualisaatiot ovat yksi harvoista keinosta avata monimutkaisia datapohjaisia analyyseja tai ennusteita kognitioystävällisesti (Padilla ym. 2021). Niiden tarkastelu yksittäisten numeroiden tai riippuvuuksien tasolla on ihmiselle työlästä ja usein myös mahdotonta. Datapohjaisten tilastollisten ilmiöiden tulkitseminen edellyttää lisäksi usein huomattavaa asiantuntemusta laskennallisista analyysi- ja ennustemenetelmistä. Visualisaatiot tarjoavatkin kognitioystävällisen tavan viestiä datapohjaisista analyyseista tai ennusteista myös kohderyhmille, joilla itsellään ei ole riittävää tietoteknistä osaamista.

Generatiivinen tekoäly avaa oven uuden sukupolven visualisaatio-otekniikoille<sup>1</sup>. Se mahdollistaa aiempaa kognitioystävällisemmät kuvitukset, uudistaa visuaalista kieltä sekä laajentaa visualisaatioiden ilmaisuvoimaa. Generatiivisen tekoälyn avulla voidaan lisäksi automatisoida monia visualisoinnin välivaiheita ja tehostaa työläitä tuotantoprosesseja.

Samalla tekoäly kuitenkin muuttaa visualisaatioiden tiedollista ja viestinnällistä luonnetta. Generatiiviset algoritmit eivät vain leikkaa, kopioi ja liitä, vaan aidosti *generoivat* eli *tuottavat* sisältöjä. Generatiivinen tekoäly muokkaakin perinteisiä suunnittelu- ja tuotantoprosesseja tavalla, joka pakottaa arvioimaan niiden vaikutuksia esimerkiksi viestinnälle osittain uudesta näkökulmasta.

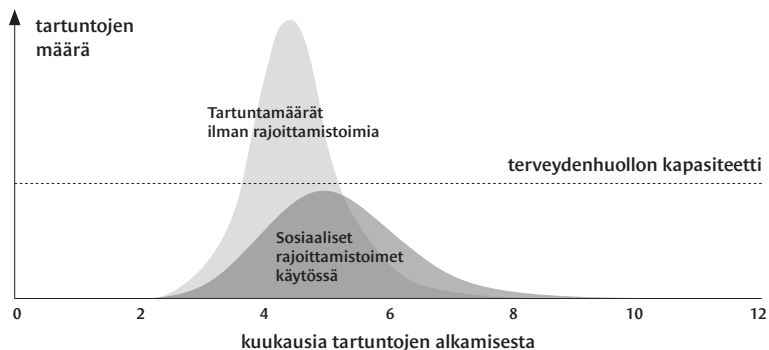
---

1 Kaupallisia ja avoimen lähdekoodin sovelluksia on jo useita.

Mitä monimutkaisemmaksi ja itsenäisemmiksi algoritmiset visuaalisaatioteknologiat kehittyvät, sitä vaikeampi käyttäjän on hahmottaa, kuinka visualisointiprosessit tapahtuvat tai millaiset tekijät niitä ohjaavat. Erityisesti tiede- ja viranomaisviestinnän kaltaisilla aloilla, joissa tiedollinen läpinäkyvyys on kriittistä, tästä seuraa haasteita. Artikkelissa tarkastellaan näitä haasteita ja käytetään Midjourney-nimistä kuvia generoivaa algoritmia havainnollistamaan niiden aiheuttamaa problematiikkaa kuvan 4 avulla.

## Datavisualisaatiot

Covid-19-pandemian eräs anti oli, että se tutustutti laajan yleisön, toimittajat ja päätöksentekijät datapohjaisiin epidemiologisiin ennusteisiin ja niiden visualisaatioihin. Tartuntataulukoista, koronakuolleisuuskaavioista ja pandemian etenemistä kuvaavista kasvukäyristä muodostui pitkälti käsitteistö, jolla viruksesta viestittiin. Moni muistaakin esimerkiksi terveydenhuollon kestäkykyä havainnollistaneet THL:n pylväs- ja viivadiagrammit, joiden pohjalta miltei koko koronakeskustelu käytiin keväällä 2020:



**Kuva 1.** Reproduktio THL:n alkuperäisestä kaaviosta. Alkuperäinen kaavio: thl.fi.

---

THL:n yksinkertainen graafi ilmaisi muutaman, selkeän visuaalisen muuttujan, kuten geometrysten muotojen (käyrät), värin (kuva 1: harmaan eri asteet) ja valikoitujen yksityiskohtien (esim. tummuusasteiden kontrastit) avulla kaksi kehityskulkua tavalla, josta oli helppo päätellä niiden vaikutukset terveydenhuollon kestokykyyn.

Kognitiontutkimuksen näkökulmasta graafi hyödyntää tehokkaasti ihmisaivoissa näköinformaation käsittelyyn osallistuvia lajityypillisiä, osittain automaattisia havainto- ja tiedonkäsittelymekanismeja (Ecker ym. 2022). Aivot käsittelevät visuaalista informaatiota monikanavaisesti, mikä mahdollistaa nopean karkean kokonaiskuvan muodostamisen ja helpottaa näkökentän kohteiden muistinvaraista tunnistamista ja vertailua (Ullman 1995).

THL:n graafi on esimerkki myös siitä, kuinka visualisaatiot toimivat kognition ulkoisena tukena. Ne vähentävät muisti- ja prosessointikuormitusta ja mahdollistavat tiedonkäsittelyresurssien suuntaamisen ongelmien ratkaisemiseen. Hyvät ja selkeät visualisaatiot parantavat lisäksi päättelyiden ja päätöksenteon laatua erityisesti monitulkintaisten, epäselvien ja abstraktien ongelmien kohdalla (Tak ym. 2015).

Covid-19-pandemian aikana visualisaatiot olivatkin erityisen hyödyllisiä tilastollista päättelyä vaativien ennusteiden viestimisessä. Tilastollinen päättely, jota vaaditaan esimerkiksi kuolleisuusriskin arvioinnissa, on ihmisaivoille vaikeaa. Erityisen haastavaa riskien ja muiden todennäköisyyksiin liittyvien ilmiöiden arviointi on silloin, kun todennäköisyydet kuvataan numeerisesti ns. suhteellisina todennäköisyyksinä (esim. 0,2 tai 2 %). Ihmisen on huomattavasti helpompaa hahmottaa todennäköisyydet, jos ne kuvataan ns. luonnollisina frekvensseinä (esim. kaksi sadasta).

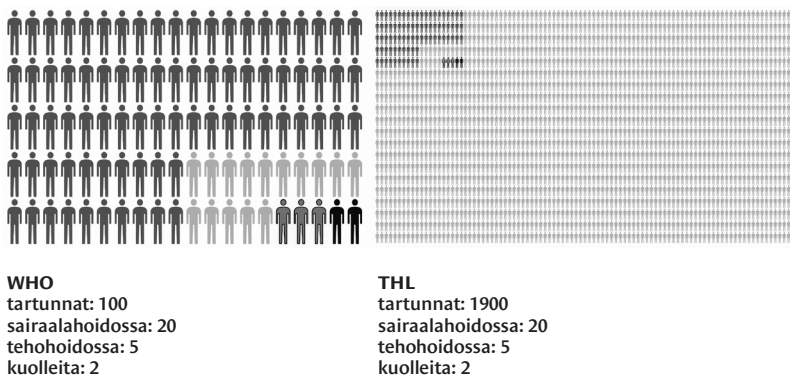
Helsingin Sanomien graafi (kuva 2) osoittaa, kuinka esitystapa helpottaa tilastollisten tekijöiden hahmottamista. Se liittyy THL:n ja WHO:n kuolleisuusarvioiden eroon, joka herätti hämmennystä keväällä 2020.

WHO:n arvioissa kuolleisuus oli n. 2 % tartuntojen määrästä. THL:n laskelmien mukaan se olisi ollut n. 0,1 %, mikä sai monet epäilemään THL:n arvioiden luotettavuutta.

Monelta jäi kuitenkin huomaamatta, että eriävät käsitykset johtuivat pääasiallisesti siitä, että WHO käytti epidemian alkuvaiheessa Kiinassa kerättyä aineistoa ja huomioi arvioissaan vain siinä vahvistetut

tartunnat<sup>2</sup>. WHO arvioi siis kuolleisuuden ja kuolleisuusriskin laske-  
malla sen suuruuden varmistettujen tartuntojen määrän perusteella  
(kaksi sadasta tartunnan saaneesta kuolee). THL:n arviossa puolestaan  
lähtökohta oli, että todellisten tartuntojen määrä olisi epidemiologisten  
laskelmien pohjalta moninkertainen (n. 1900 tartuntaa) verrattuna  
vahvistettuihin tartuntoihin (100). THL:n arvion mukaan kuolleisuus  
olisikin ollut huomattavasti maltillisempaa (0,05–0,1) kuin WHO:n  
arviossa (0,2).

Monen oli vaikeaa hahmottaa arvioiden erojen syytä, koska se  
ilmaistiin yleensä vain numeerisesti (esim. 0,05 vs 0,2). Jos ero ilmaistaan  
visualisaatiolla, jossa luvut kuvataan luonnollisina frekvensseinä ja  
konkretisoidaan ne ihmistä muistuttavalla symbolilla (kaksi sadasta  
tai kaksi 1900:stä ihmisestä), niiden hahmottaminen helpottuu huomattavasti (kuva 2):



*Kuva 2.*

WHO:n ja THL:n kuolleisuusarvioiden erot. Visualisaatiossa hyödynnetty  
samaa ajatusta kuin Helsingin Sanomien 19.3.2020 julkaistussa graafissa.  
Grafiikka Anna-Mari Wallenberg ja Jukka Niittymaa.

<sup>2</sup> Aineisto oli peräisin epidemian alkuvaiheesta Kiinasta.

---

Visualisaatio jäsentää esitettävän informaation tavalla, joka ohjaa kognitiivista järjestelmää tunnistamaan, mikä on oleellista vertailun kannalta. Lisäksi se tukee itse päättely- ja vertailuprosesseja konkreettisen esitystavan avulla.

## **Viettelevät visualisaatiot, stilisoidut faktat**

Kuolleisuusrvioiden erojen visualisaatiot kertovat, kuinka kognitiivisesti toteutetut havainnollistukset tukevat ihmisen tiedonkäsittely- ja ongelmanratkaisujärjestelmien toimintaa. Ne muotoilevat ja tulkitsevat ongelmat tavoilla, joista kognitiivisen järjestelmän on helppo tulkita, millaisten päättelyiden tai arvioiden avulla ongelma voi ratketa.

Kääntöpuoli on, että tämä tekee visualisaatioista myös kognitiivisesti seduktiivisia eli vietteleviä. Viettelevyys tarkoittaa, että visualisaatiot esijäsentävät informaation niin, että se johdattelee tiedonkäsittelyjärjestelmän kohti toivottuja päätöksiä tai tulkintoja hyödyntämällä tiedonkäsittelymekanismien automaattisia toimintaperiaatteita. Visualisaatiot eivät siis vain tulkitse, suuntaa tarkkaavaisuutta tai tue muistia, vaan ne ohjaavat kohti tietynlaisia päätelmiä, arvioita tai päätöksiä.

Visualisaatioiden seduktiivisuudesta kertovat esimerkiksi tutkimukset, joissa havaittiin, että erilaisten visualisaatioiden avulla voidaan tuottaa toisistaan poikkeavia tulkintoja (samasta) Covid-19:n kuolleisuusriskistä jopa silloinkin, kun ihmiset tietävät, että kyse on samasta kuolleisuusriskistä (Padilla ym. 2021). Tätä visualisaatioiden ominaisuutta vahvistaa niihin liittyvä mielikuva tiedollisesta luotettavuudesta ja paikkansapitävyydestä (Kennedy & Hill 2017).

Todellisuudessa visualisaatiot ovat aina yksinkertaistettuja, abstrahoituja ja jossain määrin epärealistisia kuvauksia kohteistaan. Ne ”stilisoivat” eli tyylittelevät kohteitaan poimimalla, korostamalla ja kuvittamalla valikoituja piirteitä<sup>3</sup>. Datavisualisaatioihin liittyy aina myös taiteellisia ja pragmaattisia elementtejä, sillä tiedollisten näkökulmien ohella niiden suunnittelussa otetaan huomioon esteettiset ja käytän-

---

3 Stilisoinnin käsite ks. Hirschman 2016.

---

nölliset tekijät (Ambrosio 2015). Visualisaatiot ovatkin tietoisten ja tiedostamattomien suunnitteluratkaisujen, viestinnällisten valintojen ja kompromissien lopputulos, jossa tasapainotellaan realistisuuden, esteettisyyden ja kognitiivisen käytettävyyden välillä.

## **Datavisualisaatiot ja tiedollisen ympäristön myrkyttyminen**

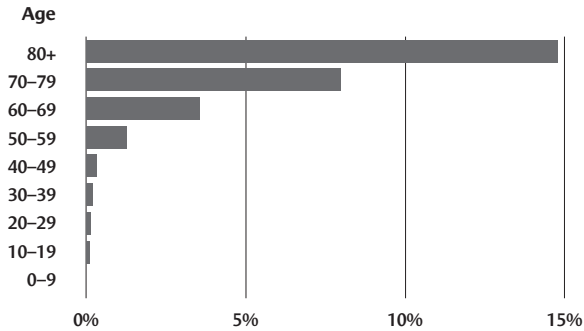
Datavisualisaatiot tarjoavat tehokkaan ja usein myös ainoan tavan viestiä monimutkaisista datapohjaisista sisällöistä ymmärrettävästi. Datavisualisaatioiden tiedetään kuitenkin ruokkivan, kuten McCosker ja Wilson (2014) toteavat, ”fantasiaa tietämisestä”. Pelkästään visualisaation lisääminen tutkimusraporttiin tai esitysmateriaaliin luo vaikutelman totuudenkaltaisuudesta ja paikkansapitävyydestä (Kennedy & Hill 2017; Padilla ym. 2021).

Visualisaatiot ovat tehokkaampia virheellisen tiedon levittäjiä ja asenteiden muokkaajia kuin perinteiset tekstipohjaiset mis- tai disinformaation kanavat (Vaccari & Chadwick 2020 Padilla ym. 2021). Tutkimusnäytön pohjalta datavisualisaatioiden merkitys onkin ollut yllättävän suuri esimerkiksi Covid-19-pandemian aikaisessa ”informaatioepidemiassa” (Iwendi ym. 2022)<sup>4</sup>. Niitä käytettiin mm. tarkoituksellisesti rokote- ja koronaskeptisten asenteiden ruokkimiseen sosiaalisessa mediassa (disinformaatio) sekä tahattomasti (ns. misinformaatio) pandemian etenemiseen liittyvien virheellisten käsitysten levittämiseen (Shelton 2020; Reinholtz ym. 2021).

BBC:n ”kokonaiskuolleisuuskaavio” on eräs esimerkki jälkimmäisestä, eli tahattomasta, mutta harhaanjohtavasta datavisualisaatiosta (kuva 3):

---

4 United Nations Department of Global Communications (2020). <https://www.un.org/en/un-coronavirus-communications-team/un-tackling-%E2%80%98infodemic%E2%80%99-misinformation-and-cybercrime-covid-19> (accessed 29 September 2020).



**Kuva 3.**  
Reproduktio BBC:n kokonaiskuolleisuuskaaviosta<sup>5</sup>.

Kaaviossa yli 80-vuotiaiden kuolleisuutta kuvaava pylväs dominoi tulkintaa<sup>6</sup>. Se vihjaa, että miltei kaikki ikäkohorttiin kuuluvat, koronatautinnan saaneet kuolisivat. Keskeinen tieto, oikeassa alakulmassa mainittu prosenttiosuus, jää monelta huomaamatta. Todellisuudessa kaavio kuvaa yli 80-vuotiaiden osuutta (vajaat 15 %) tartuntaan kuolleista<sup>7</sup>, ei koronan aiheuttamia kuolemia 80-vuotiaiden ikäkohortissa.

Vaikka BBC vastuullisena journalistisena toimijana ensin korjasi kaaviota, taustoitti sitä mm. Twitterissä ja lopulta poisti sen, vahinko ehti tapahtua. Visualisaatiota levitettiin laajalti. Se lietsoi ikäihmisten ja heidän läheistensä pelkoja, sillä arvostetun mediatalon uutisointiin uskottiin. BBC:n kaavion tai sen korjausyritysten vaikutuksista ei ole systemaattista tutkimustietoa, mutta tutkijat epäilevät, että sen kaltaiset

---

5 Tässä käytetty lähteenä:  
<https://www.tableau.com/blog/covid-19-resources-data-viz-best-practices>

6 Erinomainen analyysi BBC:n kaavion ongelmista:  
<https://www.tableau.com/blog/covid-19-resources-data-viz-best-practices>

7 Kaavio oli harhaanjohtava myös siksi, että se ei edes kertonut globaalista kokonaiskuolleisuudesta. Se kuvasi Kiinan tautien ehkäisy- ja valvontakeskuksen epidemian alkuvaiheissa raportoimia 44 000 koronaan liittyntä kuolemantapausta, joista vajaat 15 % olivat yli 80-vuotiaita.



---

tapaukset ovat polttoainetta lähinnä salaliittoteorioille ja epäluuloille (Iwendi ym. 2022).

Esimerkiksi Yhdysvalloista tiedetään pandemian ajalta vastaavia tapauksia, joissa viranomaisten tarkoituksena oli ennaltaehkäistä tartuntojen holtitonta kasvua varoittamalla kansalaisia yksinkertaistetuilla ja osittain liioitelluilla koronakuolleisuusvisualisaatioilla. Kun liioittelu selvisi, se johti lähinnä epäilyihin viranomaisten vilpittömyydestä, eikä niinkään pyrkimyksiin ennaltaehkäistä pandemiasta aiheutuvia haittoja toimimalla viranomaisten toivomalla tavalla (Shelton 2020; Iwendi ym. 2022).

Onkin hyvä tiedostaa, että harhaanjohtava visuaalinen viestintä voi synnyttää ilmapiiirin, joka ruokkii epäluottamusta viranomaisiin, tutkijoihin ja muihin perinteisesti luotettaviksi arvioituihin tiedonlähteisiin. Tutkijoiden mukaan mis- ja disinformaatio eivät aiheuta vain vääriä ja perusteettomia käsityksiä, vaan ne vahingoittavat koko tiedollista ympäristöä estämällä myös paikkansa pitäviä uskomuksia (Fallis 2021). Ne eivät siis johda ”vain” välitetyn sisällön paikkansapitävyden kyseenalaistamiseen, vaan tieto niille altistumisesta on itsessään psykologisesti vahingollista (Fallis 2021; Ecker ym. 2020. Kuten Yhdysvaltojen tilanne jossain määrin osoittaa, seurauksena on tiedollisesti myrkyttynyt ilmapiiiri, joka rapauttaa yhteistyön, kollektiivisen päätöksenteon ja yhteisöllisyyden kognitiivisia ja sosiaalisia mekanismeja. Tämä puolestaan voi kääntyä viranomaisia vastaan, heikentää yhteiskunnallista vakautta ja vaarantaa laajalti mm. demokratian keskeisiä toimintaedellytyksiä (Vaccari & Chadwick 2020; Balliet & Van Lange 2013).

## **Generatiivinen tekoäly tiedeviestinnässä**

Generatiivisen tekoälyn avulla on mahdollista luoda entistä tehokkaampia, uskottavampia ja vakuuttavampia visualisaatioita. Tekoäly avaa oven uuden sukupolven havainnollistamiselle, jossa voidaan yhdistää liikkuvaa kuvaa, ääntä ja tekstiä kolmiulotteisiin, kokemusmaailman kielelle käännettyihin simulaatioihin. Tämä synnyttää myös huolen generatiivisen tekoälyn hyödyntämisestä entistä aidommalta vaikuttavien syväväennösten ja muiden dis- ja misinformaatioiden muotojen laajamittaiseen tuottamiseen.

Datavisualisaatioiden kohdalla huolet kohdistuvat myös tapaan,

---

jolla tekoäly mutkistaa visualisaatioiden ja niiden kohteiden välistä suhdetta. Esimerkiksi tieteessä datavisualisointi on perinteisesti ollut erityinen asiantuntijatehtävä, jossa visualisoijilla on ollut tiedollinen pääsy myös datan käsittelyyn, tulkitsemiseen ja analysoinnin eri vaiheisiin. Se on mahdollistanut, että tarvittaessa on voitu artikuloida ja kommunikoida graafista suunnittelua, visuaalista kieltä tai sisältöjen poimimista ohjanneet tekijät. On tiedetty, mitä on visualisoitu ja miksi.

Generatiivinen tekoäly ja ylipäänsä monimutkaiset algoritmiset visualisointiohjelmistot muokkaavat tätä tiedollista yhteyttä. Algoritmit saattavat automatisoida visualisoinnin välivaiheita tavoilla, jotka eivät ole läpinäkyviä visualisoijalle. Tämä vaikeuttaa visualisaatioiden tiedollisten reunaehtojen tai niiden epävarmuuksien viestimistä.

Kun datavisualisaatioita käytetään esimerkiksi yhteiskunnallisen päätöksenteon tukena, on kuitenkin kriittistä, että nimenomaan epävarmuudet, keskeiset tiedolliset ominaisuudet ja tutkimuksen hypoteettinen luonne viestitään selkeästi. Esimerkiksi ilmastonmuutosmallit, kuten NOAA-simulaatiot, ovat datapohjaisia ennustemalleja, joiden käyttö päätöksenteon tukena edellyttää niiden hypoteettisen luonteen artikulointia.

NOAA-mallit harjoitetaan valtavilla data-aineistoilla, jotka koostuvat mm. satelliittidatasta, ja vuosikymmenten aikana eri puolilta maapalloa kerätyistä, sääilmiöihin liittyvistä valtavista mittausaineistoista. Malleissa käytetään kymmeniä monimutkaisia algoritmeja, joiden vaatimaan laskentaan tarvitaan tehokkaimpia supertietokoneita.

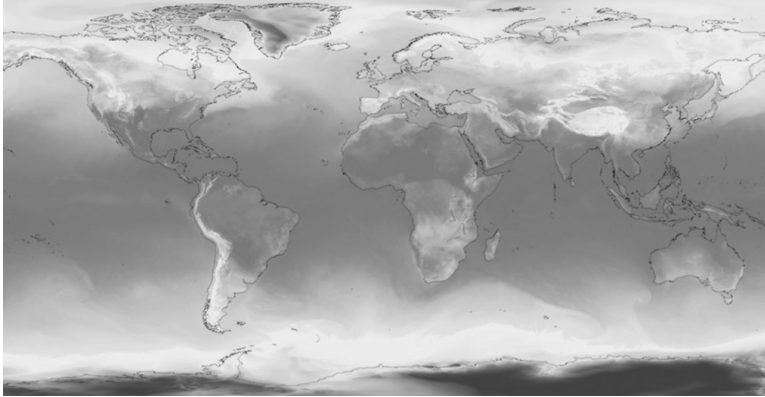
NOAA-malleissa planeetta kuvataan miljoonina kuutioina, joiden välillä energia siirtyy termodynamiikan lakien perusteella. Simulaatiot laskevat ennusteita historiallisen aineiston perusteella (esimerkiksi ennuste vuosien 1950–1965 pohjalta vuosille 1975–1985), ja testaavat niitä historiallisen datan avulla (esim. vertaamalla ennustetta mittausdataan vuosilta 1975–1985). Testitulosten pohjalta laaditaan varsinaiset simulaatiot esimerkiksi maanpinnan lämpötilan (mahdollisesta) kehityksestä mallin sisällä. Lämpötilan nousu viittaa siihen, että mallin laskeman ennusteen lämpötilojen keskiarvot ylittävät ketun 20. vuosisadan keskiarvon.<sup>8</sup>

---

8 NOAA-simulaatioissa planeetta kuvataan miljoonina kuutioina, joiden välillä energia siirtyy termodynamiikan lakien perusteella. Mallissa lämpötilan nousu viittaa mallin laskeman ennusteen lämpötilojen keskiarvoihin tietyllä aikavälillä (eli tietyn aikavälin lämpötilan ennustetaan olevan korkeampi kuin historiallisesta datasta laskettu 20. vuosisadan keskiarvo).

---

Perinteisesti NOAA-ennusteita on visualisoitu valikoimalla erilaisia viiva- ja pylväsdiagrammeja ja erityisesti laajalle yleisölle tarkoitetuilla kaksiulotteisilla kartoilla ([www.climate.gov](http://www.climate.gov)). Karttojen avulla konkretisoidaan ennusteet tavalla, joka helpottaa lämpötilojen muutoksen vaikutuksien hahmottamista:



**Kuva 4:**  
Esimerkki kaksiulotteisesta, karttapohjaisesta NOAA-visualisaatiosta.  
Lähde: NOAA-visualisaatiot.

NOAA-viitekehyksessä visualisaatioiden tiedollisen käytön rajoitteista pyritään viestimään selkeästi avaamalla niiden suhde kuvattuihin simulaatioihin, tarjoamalla pääsy laskennassa käytettyihin datajoukkoihin sekä kertomalla simulaatioiden teoreettisesta ja laskennallisesta taustasta<sup>9</sup>. Lisäksi itse visualisointitekniikoita avataan käyttäjille esimerkiksi tarjoamalla mahdollisuus kokeilla työkalua, jonka avulla havainnollistetaan visualisaatioiden riippuvuutta datasta ([www.climate.gov](http://www.climate.gov)).

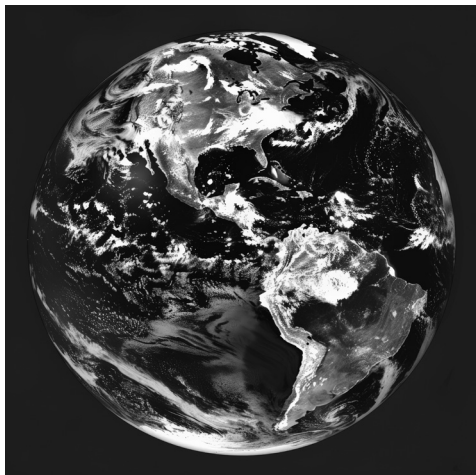
Kun generatiivisen tekoälyn avulla visualisoidaan NOAA-simulaatioiden tapaisia malleja, visualisaation ja simulaation välinen yhteys

---

<sup>9</sup> Climate.gov-sivustolla kerrotaan yleistajuisesti ja tyhjentävästi ilmastomallien laskennasta, mallien visualisoinnista ja tarjotaan pääsy sekä dataan että myös erilaisiin visualisointityökaluihin.

---

katkeaa tavalla, joka heikentää niiden tiedollista luonnetta, synnyttää uusia epävarmuuksia sekä vaikeuttaa ilmastomallien tiedollisten rajoitteiden viestimistä ja ymmärtämistä. Alla oleva Midjourneylla promptattu visualisaatio on tästä esimerkki:



*Kuva 5.*

Midjourneylla tehty NOAA-tyylinen visualisaatio.

Kuva: Jukka Niittymaa ja Midjourney (v6). Luotu 17.4.2024 promptilla: black and white NOAA-style visualization of global warming from a satellite picture --s 50 --style raw.

Midjourneyn tuottama visualisaatio on hyvin uskollinen NOAA-visualisaatioiden visuaaliselle kielelle siinä määrin, että sitä on vaikea erottaa aidosta NOAA-visualisaatiosta. Olennaista kuitenkin on, että Midjourneylla tehdyllä visualisaatiolla ei ole yhteyttä alkuperäiseen simulaatiodataan, eikä Midjourneyn käyttäjä pysty yksilöimään, missä määrin visualisaatio visualisoi itse NOAA-malleja.

Vielä ongelmallisempaa on, ettei Midjourneyn käyttäjä pysty myöskään yksilöimään, mitä kehote- eli promptipohjainen visualisaatio tosiasiaa visualisoi. Erityisesti kaupallisten sovellusten kohdalla järjestelmät ovat käyttäjilleen liikesalaisuussyistä pitkälti mustia laatikoita. Käyttäjälle ei kerrota, miten generointityökalu toimii, millaisella aineis-

---

tolla ja miten se on harjoitettu, tai kuinka ne ratkaisevat esimerkiksi ns. hallusinointiin liittyviä haasteita (Dai ym. 2023).

Monimutkaiset kuvangenerointityökalut, ja ylipäänsä algoritmit, lisäävätkin visualisointiprosessiin läpinäkyvyyttä ainakin kahdella tasolla. Ensiksi ne vaikeuttavat sen yksilöimistä, miksi ja miten algoritminen järjestelmä tuottaa lopputuloksen, jonka se tuottaa (läpinäkyvyys selitettävyytenä, ks. Mittelstadt ym. 2019). Toiseksi algoritmiset työkalut katkaisevat yhteyden simulointiprosessiin ja vaikeuttavat viestimistä siitä, mitkä visualisaatioiden tiedolliset reunaehdot ja epävarmuudet ovat (ns. kommunikaatiivinen läpinäkyvyyttä).

## Lopuksi

Mitä enemmän generointityökaluja tulevaisuudessa hyödynnetään monimutkaisten, datapohjaisten ilmiöiden visualisoinnissa, sitä tärkeämpää on, että huomioidaan algoritmien vaikutukset esimerkiksi läpinäkyvyyteen. Datavisualisaatiot, ja ylipäänsä visualisaatiot, ovat lähtökohtaisesti tyyliä, yksinkertaistettuja ja osittain vääristeleviä kuvauksia niiden kohteista. Ne ovat aina suunnitteluratkaisujen, käytettyjen tekniikoiden ja viestinnällisten valintojen kompromissi, jossa tasapainotellaan informatiivisuuden, esteettisyyden ja käytettävyyden välillä.

Eryityisesti tiedollisen luotettavuuden suhteen sensitiivisillä aloilla, kuten tiedeviestinnässä, on keskeistä, että tasapainottelua ohjanneet valinnat voidaan tarvittaessa artikuloida ja avata. On kyettävä kertomaan, mitä on visualisoitu ja miksi. Generatiivinen tekoäly kuitenkin automatisoi visualisaatioiden laatimista tavalla, joka voi olla läpinäkyvyyden mätön. Eryityisesti alueilla, joissa visualisaatiot ovat ainoa tapa viestiä ymmärrettävällä tavalla ilmiöistä, generointityökalujen vaikutus on hyvä tiedostaa.

On hyvä muistaa, että visualisaatiot ovat kognitiivisesti tehokas viestinnän muoto. Visualisaatioiden avulla voidaan toisaalta tukea kognitiivisen järjestelmän toimintaa. Toisaalta visualisaatioilla voidaan myös ohjata ja johdatella. Niihin liittyy mielikuvia tiedollisesta varmuudesta, ja ne viestivät varmuudesta.

---

Generatiivinen tekoäly luo rajattomasti mahdollisuuksia aiempaa uskottavampien, puhuttelevimpien ja ilmaisuvoimaisimpien visualisaatioiden laatimiselle. Kun tekoälyä käytetään visualisoinnissa, on keskeistä, että huomioidaan käytettyjen työkalujen vaikutukset ja tunnistetaan viestinnän välineiden ja niiden tiedollisten ominaisuuksien merkitys. On hyvä myös huomata, että vaikka teknologisilla valinnoilla on suuri vaikutus, teknologiaa itseään ei silti tarvitse demonisoida. Teknologian käyttö on myös jatkossa alisteista hyväksyttävän viestinnän normeille ja päämäärille. Vastuu niiden muodostamisesta ja noudattamisesta on edelleen kuitenkin yksin viestijöillä eli meillä ihmisillä.

---

## Näistä voit aloittaa:

- 1) MIT:in tutkijoiden interaktiivinen sivusto, jolla kuvataan, miten Covid-19-pandemian aikana koronaskeptikot vaikuttivat sosiaaliseen mediaan datavisualisaatioiden avulla: <https://vis.csail.mit.edu/covid-story/>
  - 2) Sivu, jolla käsitellään datavisualisaatioiden ongelmia: <https://medium.com/@thomas.ellyatt/bad-data-visualisation-real-life-examples-out-there-in-the-wild-eb5032329aeb>
  - 3) Tiedon julkistamisen neuvottelukunnan tiedeviestinnän suositukset: [https://www.tjnk.fi/sites/tjnk.fi/files/tiedeviestinnan\\_suosituks\\_2018.pdf](https://www.tjnk.fi/sites/tjnk.fi/files/tiedeviestinnan_suosituks_2018.pdf)
- 



**Anna-Mari Wallenberg** on kognitiotieteen yliopistonlehtori ja dosentti. Hän tutkii älykkäiden järjestelmien tiedonkäsittelyominaisuuksia, niiden välistä vuorovaikutusta ja algoritmien yhteiskunnallisia vaikutuksia. Viime vuosina Wallenberg on työskennellyt erityisesti tekoälyn etiikan parissa sekä pohtinut mm. tekoälyasetukseen liittyviä kysymyksiä.

---

## Kirjallisuus

- Ambrosio, Chiara (2015). Objectivity and representative practices across artistic and scientific visualization. Teoksessa: A. Carusi; A. S. Hoel; T. Webmoor, & S. Woolgar (toim.). *Visualization in the age of computerization*. (s. 118–144). London: Routledge.
- Balliet, Daniel & Van Lange, Paul (2013). Trust, conflict, and cooperation: A meta-analysis. *Psychological Bulletin*, 139:5, 1090–1112.
- Couldry, Nick & Hepp, Andreas. (2017). *The Mediated Construction of Reality*. Cambridge: Polity Press.
- Dai, Zihan; Ziwei, Ji; Dan Su & Pascale, Fung (2023). Plausible may not be faithful: Probing object hallucination in vision-language pre-training. Teoksessa: *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2023, Dubrovnik, 2023*, 2128–2140.
- Fallis, Don (2021). The epistemic threat of deepfakes. *Philosophy of Technology*, 34: 623–643.
- Ecker, Ullrich; Lewandowsky, Stephan; Cook, John; Fazio, Liza; Brashier, Nadia & Amazeen, Michelle (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews*. <https://doi.org/10.1038/s44159-021-00006-y>
- Hirschman, Daniel (2016). “Stylized facts in the social sciences.” *Sociological Science* 3:604–626, DOI 10.15195/v3.a26
- Kennedy, Helen & Hill, Rosemary (2017). The pleasure and pain of visualising data in times of data power. *Television and New Media*, 18:8, 769–782.
- McCosker, Anthony & Wilken, Rowan (2014). Rethinking ‘big data’ as visual knowledge: the sublime and the diagrammatic in data visualisation, *Visual Studies*, 29:2, 155–164.
- Mittelstadt, Brent; Russell, Chris & Wachter, Sandra (2019). *Explaining explanations in AI*. <https://arxiv.org/abs/1811.01439>
- NOAA-visualisaatiot. [https://www.nnvl.noaa.gov/view/globaldata.html#GFST?timespan=3-hour&date=2024-04-18\\_at\\_06&lat=6.25&lon=0&zoom=2](https://www.nnvl.noaa.gov/view/globaldata.html#GFST?timespan=3-hour&date=2024-04-18_at_06&lat=6.25&lon=0&zoom=2)
- Padilla, Lace; Hosseinpour, Helia & Fygenson, Racquel (2021). Impact of COVID-19 forecast visualizations on pandemic risk perceptions. *Sci Rep 12, 2014* (2022). <https://doi.org/10.1038/s41598-022-05353-1>
- Reinholtz, Nicholas; Maglio, Sam & Spiller, Stephen (2021). Stocks, flows, and risk response to pandemic data. *Journal of Experimental Psychology: Applied*, 27:4, 657–668. <https://doi.org/10.1037/xap0000395>
- Shelton, Taylor. (2020). A post-truth pandemic? *Big Data & Society*, 7:2. <https://doi.org/10.1177/2053951720965612>

---

Tak, Suzanne; Toet, Alexander & van Erp, Jan (2015). Public understanding of visual representations of uncertainty in temperature forecasts. *Journal of Cognitive Engineering and Decision Making*. 2015;9(3):241–262. doi:10.1177/1555343415591275

Ullman, Shimon (1995). Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex. *Cerebral Cortex*, 5:1, 1–11.

Vaccari, Cristian & Chadwick, Andrew. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6:1, 1–13. <https://doi.org/10.1177/2056305120903408>