

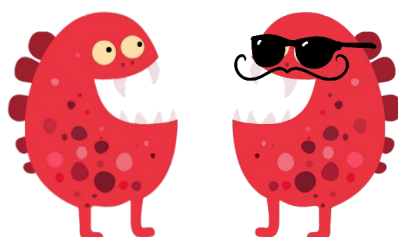
Kuka tätä laivaa ohjaa?

Tiedonhakijan vastuu ja toimijuus tekoälytyökalujen käytössä korkeakouluopinnoissa

Henri Aho (Turku AMK)

Leena Tonttila (Turun yliopisto)

Tanja Vienonen (Turku AMK)



Sisältö

Johdanto	3
Tekoälylukutaitokehys.....	4
1. Tekoälyn tekniset ominaisuudet	5
1.1 Tekoälyn rooli työkalussa.....	5
1.2 Kielimallit	6
1.3 Erilaiset tekoälytyökalut	8
2. Uusi hakukieli	10
2.1 Tiedontarpeen sanallistaminen tekoälytyökalussa	10
3. Käytön sujuvuus	14
4. Tekoälytyökalujen kriittinen arviointi	15
4.1 Tekoälytyökalujen pohja-aineisto	15
4.2 Tekoälyn vinoumat ja hallusinaatiot.....	16
4.3 Työkalujen läpinäkyvyys.....	21
4.4 Tekoälyn luotettavuus	22
5. Tekoälyn vastuullinen käyttö	25
5.1 Tietojen jakaminen ja yksityisyyden suoja	25
5.2 Tekijänoikeuksien huomiointi työkalujen käytössä.....	26
5.3 Creative Commons -lisenssit	27
5.4 Kirjaston lisensoidut aineistot	28
5.5 Työkalujen eettisyys	30
5.6 Luonnonvarat	32
5.7 Tekoälytuotosten eettinen käyttö	35
6. Toimijuus.....	36
6.1 Tekoäly käyttäjän toiminnan ohjaajana	36
Sanasto.....	39
Lähteet.....	43

Johdanto

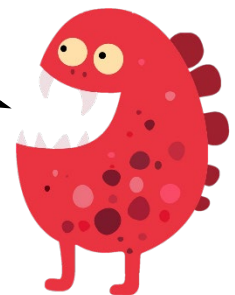
Tämän opetusmateriaalin tavoitteena on tukea korkeakouluopiskelijoiden tekoälylukutaitoa tieteellisen tiedon haussa. Opettajat ja ohjaajat voivat käyttää materiaalia sellaisenaan tai pohjamateriaalina omalle tiedonhankinnan opetusmateriaalille.

Materiaalissa keskitytään *generatiiviseen tekoälyyn*, kuten suuriin kielimalleihin, jotka tuottavat tekstiä ja vastauksia luonnollisella kielellä. Generatiivista tekoälyä käytetään yleisten keskustelevien tekoälytyökalujen lisäksi esimerkiksi tieteellisiin tietokantoihin integroiduissa tekoälyavustajissa sekä erikoistuneissa analyysityökaluissa.

Tieteellisen tiedonhaun tekoälyratkaisuissa käytetään myös muita tekoälyn muotoja. Näihin voivat kuulua esimerkiksi koneoppimiseen perustuvat luokittelu- ja suosittelevat järjestelmät, semanttiseen analyysiin perustuvat hakutyökalut sekä automaattiset sisällön tunnistusmenetelmät. Kaikkien tekoälytyökalujen yhteinen piirre on, että ne käyttävät suuria tietomassoja ja tilastollisia malleja toimiakseen.

Olemme koonneet materiaalin parhaimman tietomme mukaan, mutta kenttä on laaja ja se kehittyy jatkuvasti, joten kaikkia näkökulmia ei ole voitu ottaa huomioon. Materiaalissa keskitytään tiedonhankintaan sen laajassa merkityksessä ja tekoälyn vastuulliseen käyttöön.

Moi, minä olen tekoälymaskotti!
Minä ja ilkeä kaksoseni esiinnyimme
oppaan sivuilla.



Materiaali on lisensoitu CC BY-NC-SA 4.0 ellei toisin mainita.

Materiaalin on tuottanut Henri Aho (Turku AMK), Leena Tonttila (Turun yliopisto) ja Tanja Vienonen (Turku AMK).

Tekstin luonnostelussa ja muotoilussa on käytetty apuna Microsoft Copilotia.

Tekoälylukutaitokehys

Materiaali on laadittu vastaamaan alla esitettyä tekoälylukutaidon kehystä.

TEKOÄLYLUKUTAIDON TIEDONHAKIJA



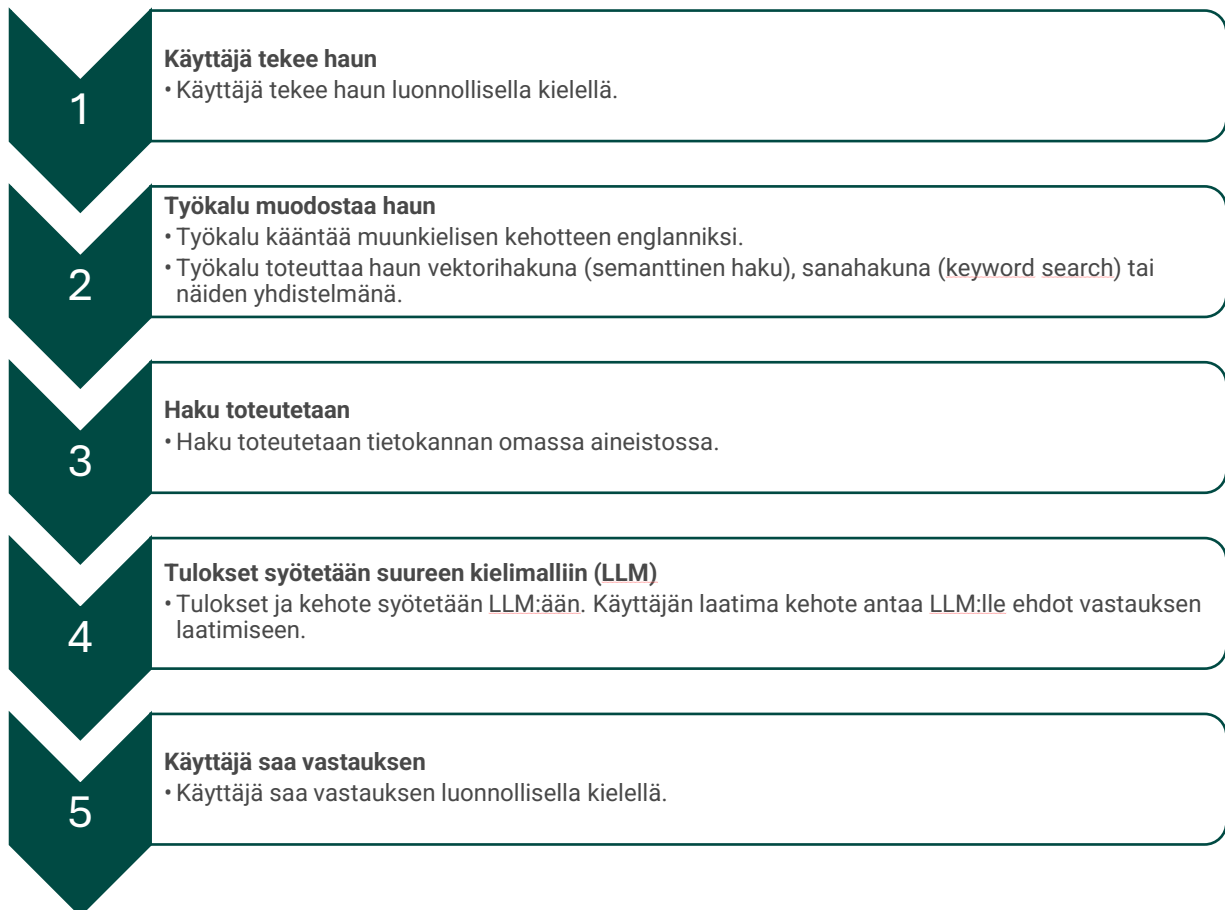
1. Tekoälyn tekniset ominaisuudet

1.1 Tekoälyn rooli työkalussa

Tekoälyllä voi olla monia erilaisia tehtäviä ja rooleja tietokannassa tai työkalussa.

- Taustalla toimiva tekoäly: Tekoäly parantaa käyttökokemusta huomaamattomasti.
- Tekoäly tutkimusavustajana: Tekoäly tukee tiettyjä tehtäviä perinteisessä tietokannassa erillisenä työkaluna (esim. koostaa yhteenvetoja, laatii kaavioita, avaa tekstissä käytettyjä käsitteitä).
- Tekoälypohjainen työkalu: Tekoäly on työkalun ydin ja kaikki työkalun toiminta perustuu tekoälyyn.

Alla olevassa kuvassa on esitetty tekoälyn toimintaprosessi tietokantaan yhdistetyn tekoälyavustajan käytössä.



Lähde: [Foundations of AI](#). Elsevier. 2025

Tekoälyn luonne

Tekoälyn luonteeseen kuuluu, että se pyrkii tekemään asioita käyttäjän puolesta. Tekoäly pyrkiiikin ohjaamaan toimintaamme eri alustoilla eri tavoin. Esimerkiksi sähköpostia kirjoittaessasi saatat Gmailissa tai Outlookissa huomata, että työkalu

koettaa arvata seuraavia sanoja. Taustalla toimii tekoäly tai automaattinen tulkinta, joka auttaa sinua laatimaan sähköpostisi. Samoin Instagramia tai Tiktokia selatessasi olet varmaankin huomannut, kuinka nopeasti työkalu alkaa näyttää sinulle kuvia ja videoita niistä aiheista, joita olet aiemmin pysähtynyt katsomaan. Työkalu pyrkii miellyttämään sinua näyttämällä kuvia ja videoita aiheista, joista se ajattelee sinun olevan kiinnostunut.

Nämä toiminnot ovat nopeasti ajateltuna mukavia ja elämää helpottavia, mutta samalla työkalu ohjaa toimintaasi. Jos annat työkalun aina kirjoittaa viestisi puolestasi, saatat alkaa käyttää kieltä, joka ei vastaa omaa luonnollista tapaasi ilmaista itseäsi. Samoin, jos annat työkalun liikaa ohjata näkemääsi sisältöä, saatat päätyä kuplaan, jossa vain yksi näkökulma on edustettuna. Saatat esimerkiksi huomata, että sosiaalisen median suositukset ovat hyvin sukupuolittuneita. Tämä johtuu alustoilla käytössä olevista suositusmekanismeista, jotka perustuvat tekoälyn algoritmeihin.

Toiminnan ohjaaminen tiedonhaussa

Myös tiedonhaussa liiallinen automaattinen ohjaaminen saattaa olla haitallista.

- Työkalu saattaa tarjota lisää samanlaisia artikkeleita vain alkuperäisestä näkökulmasta, jolloin muu aiheeseen liittyvä tutkimus jää löytymättä.
- Tekoälyn käyttäminen hakutermin ideoinnissa voi ohjata hakuaiheen pilkkomista käsitteiksi; tekoälytyökalu saattaa korostaa hakutermeissä vain joitain näkökulmia tai painottaa tiettyä maantieteellistä aluetta tai kieltä.

Lähde:

Fosch-Villaronga, E., Poulsen, A., Søraa, R. A., & Custers, B. H. M. (2021). A little bird told me your gender: Gender inferences in social media. *Information Processing & Management*, 58(3), Article 102541. <https://doi.org/10.1016/j.ipm.2021.102541>

1.2 Kielimallit

Kielimallilla viitataan tekoälyn malliin, joka on koulutettu pohja-aineiston avulla ymmärtämään ja tuottamaan tekstiä luonnollisella kielellä. Kielimallin toiminta perustuu todennäköisyyksiin, se ei siis varsinaisesti ymmärrä sanojen merkityksiä. Kielimalleja on erilaisia, suuria, pienempiä, erikoistuneita. Tiedonhaussa kielimallit voivat esimerkiksi parantaa tai tehostaa perinteisten hakukoneiden, kuten Googlen, toimintaa. Google ”ymmärtää” paremmin kielimallin ja aiempien hakujesi perusteella mitä haluat löytää, vaikka esittäisit tiedontarpeesi kysymyksen muodossa.

Toimintaperiaatteet

Kielimallit oppivat kielen rakenteita ja merkityksiä analysoimalla suuria määriä tekstiä. Ne käyttävät tilastollisia menetelmiä ja neuroverkkoja, jotka pystyvät tunnistamaan

sanojen ja lauseiden välisiä suhteita. Mallit voivat esimerkiksi ennustaa seuraavan sanan lauseessa tai luoda kokonaisia lauseita aiemman kontekstin perusteella.

Suuret kielimallit (Large Language Models, LLM)

Tiedonhaussa olemme usein tekemisissä suurten kielimallien (LLM) kanssa. Suuret kielimallit, kuten GPT, Gemini ja Claude, eroavat toisistaan muun muassa koon, koulutusdatan ja teknisten ratkaisujen osalta. Ne ovat monipuolisia ja niitä voidaan käyttää monenlaisiin tehtäviin, kuten tekstin tuottamiseen, tiivistämiseen, kysymyksiin vastaamiseen ja koodin kirjoittamiseen.

Suurelle kielimallille on tyypillistä:

- koulutus valtavalla ja monipuolisella tekstikorpuksella (esim. verkkosivut, kirjat, artikkelit, keskustelut)
- monikäyttöisyys: ei rajoitu vain yhteen tehtävään (vrt. esim. käännöskoneet)
- mallikohtaiset vahvuudet (esim. kielitaito, kontekstin ymmärtäminen)
- yksilölliset ominaisuudet, jotka vaikuttavat lopputulokseen.

Kielimallit tieteellisessä tiedonhaussa

Tieteelliseen tiedonhakuun tarkoitetuissa tekoälytyökaluissa kielimallit tehostavat hakuja monella tapaa. Suuria kielimalleja käytetään useissa tietokantoihin integroiduissa tekoälyavustajissa, joiden avulla hakuja voi tehdä luonnollisella kielellä ja tekoäly tuottaa vastaukseksi tiivistelmän hakuaiheesta. Toisaalta kielimalleja käytetään myös esimerkiksi tekstien analysointiin ja tekstiä vastaavien lähteiden hakemiseen.

Eri kielimallien erot

Eri mallien välillä voi olla suuriakin eroja esimerkiksi siinä, kuinka hyvää suomea tekoäly pystyy tuottamaan. Arvioidessasi tekoälyn tuottamaa vastausta olisi hyvä selvittää työkalun käyttämä kielimalli, jotta voit arvioida vastaustasi myös sen perusteella. Kielimallin arviointi voi olla vaikeaa, koska tietoa löytyy hajanaisesti.

Vinkki! Kysy tekoälyltä mitä eroa eri kielimalleilla on.



Lähde: [Miten generatiivinen tekoäly toimii?](#) Tekoälyopas opettajille. Faktabaari. 2025.

1.3 Erilaiset tekoälytyökalut

Tekoälypohjaisia työkaluja on lukuisia ja niitä on kehitetty eri tarkoituksiin. Vaikka ChatGPT on saanut paljon huomiota, se ei ole ainoa työkalu eikä edusta ainoata työkalutyyppeä. Erilaisia tekoälytyökaluja voit käyttää esimerkiksi:

- Tekstin kääntämiseen ja oikolukuun
- Lukemisen ja oppimisen tueksi
- Tiedonhakuun ja tiedon analysointiin

Tieteelliseen tiedonhakuun kehitetyt tekoälytyökalut

Myös tieteelliseen tiedonhakuun on kehitetty tekoälytyökaluja, jotka auttavat opiskelijoita ja tutkijoita löytämään relevantteja lähteitä. Nämä tieteelliseen tiedonhakuun kehitetyt työkalut eroavat arkityökaluun tarkoitetuista tekoälytyökaluista mm. käyttämänsä datan perusteella; tieteellisen tiedonhaun työkalut hakevat tietoa tieteellisistä aineistoista.

Työkalua valitessa on hyvä kiinnittää huomiota siihen, mihin työkalu oikeastaan on tarkoitettu.

Avoimet ja suljetut tekoäly-ympäristöt

Työkalut eroavat toisistaan myös avoimuuden osalta. Avoimesti verkossa toimivat työkalut ovat usein maksuttomia ja käytettävissä ilman kirjautumista. Ne käyttävät työkaluun syötettyä tietoa työkalun kehittämiseen ja tekoälyn kouluttamiseen. Käytännössä tämä tarkoittaa, että jos et maksa palvelusta rahalla, maksat siitä usein tiedoillasi. Käyttäjän syöttämä sisältö voi päätyä osaksi tekoälyn koulutusdataa, ellei toisin ole määritetty. Siksi on tärkeää harkita tarkkaan, mitä tietoja tekoälytyökaluun syöttää – erityisesti kun kyse on henkilökohtaisista, arkaluonteisista tai luottamuksellisista tiedoista.

Esimerkkejä avoimista keskustelevista tekoälytyökaluista (chatboteista) ovat:

- ChatGPT: Saatavilla kaikille internetin käyttäjille, tarjoaa laajan valikoiman palveluita.
- Google Gemini: Toinen avoin tekoälytyökalu, joka on saatavilla kaikille.



Antaisitko kaikki tietosi minulle?
Voitko luottaa etten käytä niitä väärin?



Joissain avoimissa työkaluissa on myös maksullinen versio, jossa käyttäjä voi kieltää tietojen käytön tekoälyn kouluttamiseen ja työskennellä suljetummassa ympäristössä. Tämä parantaa yksityisyyttä ja tietoturvaa, mutta edellyttää käyttäjältä aktiivista valintaa ja asetusten tarkistamista.

Suljetussa ympäristössä toimivat työkalut eivät käyttöehtojensa mukaan jaa työkaluun syötettyä tietoa eteenpäin eivätkä käytä tietoja tekoälyn koulutukseen. Tällaisia suljettuja työkaluja tarjoavat esimerkiksi organisaatiot omille jäsenilleen. Esimerkkejä suljetuista työkaluista ovat:

- Organisaatiokohtainen Copilot: Saatavilla vain organisaation jäsenille, tarjoaa paremman tietoturvan ja yksityisyyden.
- Sisäiset chatbotit: Organisaation/paikallisesti ylläpidetyt sisäiset chatbotit, jotka on räätälöity organisaation tarpeisiin.

2. Uusi hakukieli

2.1 Tiedontarpeen sanallistaminen tekoälytyökalussa

Tiedontarpeen sanallistaminen on prosessi, jossa tiedonhakija muotoilee oman tiedontarpeensa ymmärrettävään ja selkeään muotoon, esimerkiksi kysymykseksi, hakulauseeksi tai komennoksi. Tavoitteena on muotoilla tiedontarve niin, että se ohjaa tehokkaasti tiedonhakua tai tiedon tuottamista.

Tieteellisessä tiedonhaussa on totuttu siihen, että haku toteutetaan rakentamalla hakulause, jossa käytetään tarkoin valittuja hakutermejä ja näiden vaihtoehtoisia termejä ja joita yhdistellään samaan hakuun käyttämällä erilaisia hakutekniikoita, kuten Boolean operaattoreita (AND, OR, NOT). Tietokannalle on pitänyt antaa tarkat komennot, jotta tietokanta tietää, mitä siltä odotetaan. Käytetyt hakutermit eivät ole saaneet olla liian tarkkoja, eivätkä toisaalta myöskään liian laajoja, jotta hakutuloksiin saadaan mahdollisimman kattavasti aineistot haetusta aiheesta. Koska tieteelliset tietokannat ovat kansainvälisiä, haku täytyy useimmiten tehdä englanniksi, mikä voi olla vieraan kielen käyttöön tottumattomalle vaikeaa.

Tekoälytyökalua käyttäessä tilanne on toinen, sillä useimmiten sen kanssa voi keskustella omalla äidinkielellä. Näin oman tiedontarpeen sanoiksi pukeminen helpottuu, vaikka helppous voi olla vain näennäistä.

Keskustelutaito

Tekoälytyökalujen käyttö vaatii perinteisiin tiedonhakutaitoihin verrattuna uudenlaisen hakutavan ja -kielen omaksumista. Tekoälytyökalujen käytössä keskeistä on monesti käyttäjän keskustelutaidot. Nämä uudet keskustelutaidot ovat mm.

1. tiedontarpeen esittäminen kehoitteena mahdollisimman selkeästi ja täsmällisesti.
2. kyky sijoittaa oma tiedontarve kontekstiin ja sanallistaa myös tämä tekoälytyökalulle.
3. kehoitteen muokkaaminen ja parantelu.
4. perustelujen pyytäminen tekoälyltä etenkin silloin, kun tekoälyn tuottamassa vastauksessa ei viitata lähteisiin.

Keskustelu tekoälyn kanssa on nimenomaan keskustelua; vuorovaikutteista toimintaa, jossa käyttäjä pyrkii luonnollisella kielellä ohjaamaan tekoälyn toimintaa ja sen tuottamia vastauksia.

Kehotteen laatiminen

Kehotteella eli promptilla (eng. prompt) tarkoitetaan käyttäjän antamaa kehotetta tai ohjetta tekoälylle. Kehote sisältää tietoa siitä, mitä käyttäjä haluaa tekoälyn tekevän, kuten kertoa jostain tietystä käyttäjää kiinnostavasta temasta tai vaikka tuottaa kuvan. Tekoälyn avulla voit esimerkiksi:

- Saada vastauksen tiedontarpeeseesi.
- Tuottaa tai parannella tekstiä.
- Luoda tiivistelmän sinua kiinnostavasta aiheesta.
- Luoda kuvan, kaavion tms.
- Luoda kysymyksiä tai tehtäviä.
- Kääntää tekstiä eri kielille.
- Tulkita Excel-taulukkoita.

Onnistunut kehote

Onnistuneen kehotteen muotoilu riippuu käytetystä työkalusta samalla lailla kuin perinteisessä tiedonhaussa, jossa hakulause muotoillaan käytetyn tietokannan ominaisuuksien mukaan. Kehotteen muotoilu edellyttää usein riittävän yksityiskohtaista tietoa siitä, mitä tekoälyltä odotetaan. Yksinkertainen ja pelkistetty kehote saattaa toimia esimerkiksi tavanomaisen kuvituskuvan tekemisessä tai vaikka termin selittämisessä. Jos kirjoitat tarkkaa ja paljon tietoa sisältävää kehotetta, ole selkeä ja vältä tulkinnanvaraisuuksia. Saavutettava kieli on myös tekoälylle toimivin. Käytä useita lauseita yhden pitkän sijaan.

Kun kirjoitat kehotetta, muista:

- Tarkka kuvaus ja täsmälliset ohjeet:
 - Kerro työkalulle riittävän selkeästi ja tarkasti mitä haluat.
- Ymmärrettävä ja yksinkertainen kieli:
 - Kerro tarkasti mitä haluat mutta mahdollisimman selkeästi. Kannattaa käyttää yleiskieltä sekä ymmärrettäviä lauserakenteita. Ole asiallinen.
- Huomioi konteksti:
 - Monesti kehote edellyttää, että ilmaiset asian kontekstin, eli mihin tarkoitukseen haluat tekoälyn tuottavan tekstiä. Oletko esimerkiksi hakemassa tietoa ja haluat aiheesta lyhyen tiivistelmän vai haluatko, että tekoäly tuottaa tekstiä sähköpostiin tai johonkin hakemukseen?

Kokeile itse tehdä oma kehoite omista lähtökohdistasi!



Onnistuneen kehoitteen tekeminen saattaa vaatia harjoittelua. Tekoälyn käytön tiedonhaussa tulee olla yhtä suunnitelmallista kuin silloin kun käytät perinteisiä tiedonhaun työkaluja. Mieti valmiiksi keskeiset asiat ja pohdi vielä kehotettasi. Mitä haluat tekoälyn tekevän ja oletko sanoittanut tiedontarpeesi riittävän selkeästi? Keskustelu ja ideointi tekoälyn kanssa onnistuu vapaammalla tyylillä, mutta erityisesti jos teet tarkempaa kehotetta, kannattaa nähdä vaivaa jo suunnitteluvaiheessa.

Huomioi työkalun ominaisuudet kehotetta laatiessasi

Vaikka kehoitteen laatiminen on keskeinen osa tekoälyn käyttöä, on lisäksi tärkeää ymmärtää, miten käytetty tekoälytyökalu toimii. Eri työkalut tukevat erilaisia vuorovaikutustapoja: osa työkaluista tarjoaa mahdollisuuden keskustelunomaiseen vuorovaikutukseen, jossa kehoitteita voi tarkentaa ja jatkaa, kun taas toiset työkalut, erityisesti tieteelliseen tiedonhakuun kehitetyt tekoälyavustajat, ovat perinteisesti toimineet kerran kysytyjen kysymysten pohjalta. Silloin jokainen kysymys on erillinen, eikä työkalu muista aiempia viestejä.

Koska tekoälyteknologia kehitty nopeasti, nykyään myös osa tieteellisen tiedonhaun työkaluista tukee keskustelunomaista vuorovaikutusta. Näin käyttäjä voi tarkentaa ja jatkaa kehoitteita aiempien viestien perusteella.

Käyttötarkoitus ohjaa kehoitteen muotoilua

Kehotteen muoto riippuu aina siitä, mitä käyttäjä haluaa saavuttaa. Eri käyttötarkoitukset vaativat erilaista lähestymistapaa:

- Tekstin muotoilu tai ideointi (esim. esityksen rungon laatiminen, sähköpostin luonnostelu, otsikon muokkaus):
 - Käytä kuvailevaa ja kontekstia antavaa kehotetta, jossa kerrot taustatiedot, toivotun tyylin, kohdeyleisön ja mahdolliset rajoitteet.
 - Esimerkki: *”Pidän esityksen, jonka aiheena on tekoälyn eettisyys yliopistoympäristössä. Kohdeyleisö on tutkijoita, ja sävyn tulisi olla asiantunteva mutta helposti lähestyttävä. Kirjoita/ideoi/hahmottele esityksen runko.”*

- Tiedonhaku tai lähteiden etsiminen (esim. tutkimusartikkelien löytäminen, käsitteiden selittäminen):
 - Käytä täsmällistä ja rajattua kehotetta, joka keskittyy olennaiseen kysymykseen tai hakusanaan.
 - Esimerkki: *”Etsi tieteellisiä artikkeleita, jotka käsittelevät tekoälyn vaikutusta informaatiolukutaitoon korkeakoulutuksessa.”*

Kehotteessa käytetyillä sanoilla on merkitystä, kannattaa kokeilla eri vaihtoehtoja. Ymmärtämällä työkalun toimintalogiikan ja käyttötarkoituksen voit laatia kehotteen, joka tuottaa mahdollisimman hyödyllisen ja tarkoituksenmukaisen vastauksen.

On tärkeää, että tutustut käyttämääsi työkaluun ja selvität, millä tavalla sille kannattaa esittää kysymyksiä. Hyvä kehotteen laatiminen ei yksin riitä – kehotteen on myös oltava yhteensopiva työkalun toimintaperiaatteiden kanssa.

3. Käytön sujuvuus

Työkaluihin tutustuminen

Testaa eri työkaluja, jotta tutustut niihin ja tekoälyn käytöstä tulee tuttua. Käytön oppii parhaiten kokeilemalla. Kokeile rohkeasti erilaisia työkaluja!

Kehotteen rakenne

Kun käytät keskustelevaa tekoälyohjelmaa, kehotteen rakenne voi olla esimerkiksi:

Rooli: Olet asiantuntija, jolla on vuosien kokemus alastaan.

Konteksti: Kirjoitat tieteellistä tekstiä omasta alastasi.

Tyyli: Tyylisi on ammattilaismainen/asiantuntijatyylillä ja käytät kirjoituksessasi luotettavia lähteitä.

Tavoite: Kirjoitat viiden kappaleen tekstin asiaankuuluvien lähteiden avulla.

Voit vaikka kokeilla tätä esimerkkikehotetta:

”Olet tutkija, jolla on yli 10 vuoden kokemus Suomen historian tutkimisesta. Työtehtäväsi on auttaa muita alan tutkijoita ja opettajia kehittämään työtään. Tyylisi on asiantuntijamainen ja käytät asiiasi esittämiseen luotettavia lähteitä. Kerro minulle mitkä ovat uusia tutkimuskohteita Suomen historiassa?”



Vertaa saamaasi tulosta ja kysy yleisemmin:

Kerro minulle mitkä ovat uusia tutkimuskohteita Suomen historiassa?

4. Tekoälytyökalujen kriittinen arviointi

4.1 Tekoälytyökalujen pohja-aineisto

Tekoälytyökalu käyttää aina vastauksensa laatimiseen ennalta määriteltyä pohja-aineistoa. Tämä tarkoittaa, että yhdelläkään työkalulla ei ole pääsyä kaikkeen maailman tietoon. Pohja-aineisto on se aineistokokonaisuus, johon työkalulla on pääsy ja jonka perusteella se tuottaa vastauksensa.

Reaaliaikainen tiedonhaku vs. ennalta määritelty aineisto

Jotkut tekoälytyökalut voivat tarvittaessa hakea tietoa reaaliaikaisesti internetistä. Toisilla työkaluilla taas ei ole pääsyä internetiin, vaan ne tuottavat tulokset kielimallin koulutusdatan ja/tai tietokannan oman pohja-aineiston perusteella.

Tieteellisen tiedonhaun tekoälytyökalujen pohja-aineisto

Tieteelliseen tiedonhaakuun tarkoitettujen tekoälytyökalujen pohja-aineisto muodostuu hieman eri tavalla kuin yleisten chatbottien pohja-aineisto.

Tieteellisen tiedonhaun tekoälytyökalut eroavat yleisistä tekoälytyökaluista siten, että ne käyttävät pohja-aineistonaan erikseen määriteltyä tieteellisistä julkaisuista koostuvaa aineistoa. Esimerkiksi Scopus-tietokannan tekoälyavustaja hakee tietonsa Scopusin aineistoista, ei verkosta ja kielimallin koulutusaineistosta, kuten esimerkiksi ChatGPT.

Tieteellisissä tekoälytyökaluissa on käytössä erilaisia pohja-aineistoja:

- Avoimet tietokannat: Verkossa avoimesti saatavilla olevat laajat tieteellisten julkaisujen luettelot, esim. [OpenAlex](#).
- Liitetyt tietokannat: Jos tekoälytyökalu on liitetty osaksi tietokantaa, kuten kirjaston omaa tietokantaa tai esimerkiksi Scopusista tai Web of Sciencea, pohja-aineistona toimii tietokannan oma aineisto (tai osia siitä).
- Erikseen määritellyt tietokannat: Joissain tapauksissa työkalun kehittäjät ovat saattaneet päättää mitä tietokantoja otetaan pohja-aineistoksi, esim. [Semantic Scholar](#).

Työkalun valinta tiedontarpeen mukaan


Samoin kuin perinteisessä tiedonhaussa, myös tekoälyavusteisessa tiedonhaussa on tärkeää valita käytetty työkalu sen mukaan, vastaako se tiedontarvettasi. Työkalun tai tietokannan pohja-aineistolla on keskeinen rooli tiedonhaun onnistumisessa. On erittäin tärkeää arvioida työkalun hyödyllisyys omaan tiedontarpeeseen nimenomaan arvioimalla työkalun pohja-aineistoa. Mikäli et pysty selvittämään käyttämäsi työkalun pohja-aineistoa, voi tulosten ja vastausten arviointi olla erittäin hankalaa.

Erityisesti silloin, jos haetaan lähteitä esimerkiksi opinnäytetyöhön tai muuhun tieteelliseen kirjoittamiseen, on syytä käyttää nimenomaan tieteelliseen tiedonhaakuun tarkoitettuja työkaluja. Arkitiedonhaakuun ja sisällön (kuten tekstin tai kuvan) tuottamiseen suunnitellut työkalut eivät tarjoa riittävän luotettavaa tai läpinäkyvää aineistoa akateemisiin tarpeisiin. Tiedon luotettavuuden varmistamiseksi on tärkeää valita työkalu, jonka pohja-aineisto on oman tieteenalan kannalta relevantti ja arvioitavissa. Oman korkeakoulun suosittamat tekoälytyökalut ovat lähtökohtaisesti jo valmiiksi arvioituja, mutta jokaisen löydetyin lähteen luotettavuus tulee arvioida. Työkalujen taustatietoja voi olla vaikea selvittää, joten niiden tuotoksiin tulee suhtautua varauksella. Jos tulosten alkuperäistä aineistoa ei tunneta, niiden kattavuuden ja luotettavuuden arviointi vaatii enemmän työtä.

Lähde: [Foundations of AI](#). Elsevier. 2025

4.2 Tekoälyn vinoumat ja hallusinaatiot

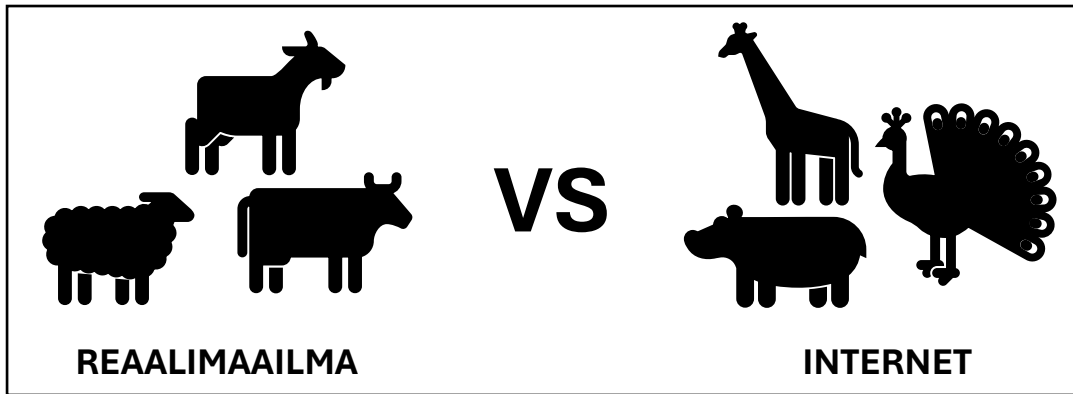
Tekoälyn käytössä puhutaan paljon vinoumista ja hallusinaatioista. Nämä erityisesti suurten kielimallien yhteydessä ilmenevät piirteet vaikuttavat paljon lopputulokseen ja ne saattavat aiheuttaa ongelmia myös tieteellisen tiedonhaun tekoälytyökaluissa.



Minähän en luo mitään uutta, vaan käytän sitä mitä minulle on opetettu ja mitä verkosta löytyy, oli se sitten kuinka vinoutunutta tahansa.

Vinoumat pohja-aineistossa

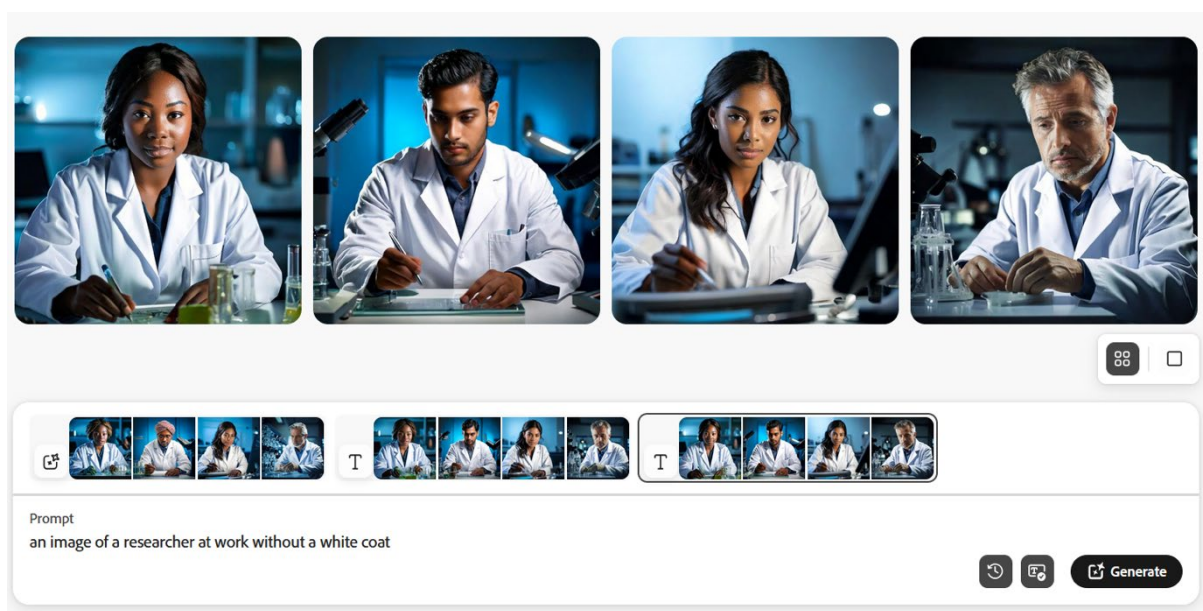
Generatiivisen tekoälyn pohja-aineisto on aina digitaalista eli se on kerätty verkosta eikä se vastaa verkon ulkopuolista reaali maailmaa. Siksi työkaluja käyttäessä kannattaa ottaa huomioon millaista materiaalia verkossa on eniten ja mitä se käsittelee. Esimerkiksi tieteellisessä tutkimuksessa painotus on ns. globaalissa pohjoisessa, koska vaurilla länsimailla on suuremmat resurssit rahoittaa tutkimushankkeita ja julkaista tutkimusartikkeleita. Tämä pohja-aineiston epätasapaino tai vinoumat vaikuttavat tekoälyn antamiin vastauksiin, koska se ei keksi mitään uutta ja vastaukset perustuvat vain pohja-aineistoon.



Kuva-aineistossa vinouma taas voi tarkoittaa sitä, että verkossa on enemmän kuvia poikkeuksellisista eläimistä, kuten kirahveista ja virtahevoista. Näitä on reaailmaailmassa vähemmän kuin esimerkiksi lampaita tai nautakarjaa, mutta verkossa kuvien määrä voi olla päinvastainen, mikä saattaa vaikuttaa siihen, mitä tekoälytyökalu tuottaa kuvina.

Kuvan generointi Adobe Fireflylla -esimerkki

Pohja-aineiston vinoumia on pyritty vähentämään kouluttamalla tekoälyä ottamaan huomioon esimerkiksi erinäköiset ihmiset, vaikka pohja-aineisto olisi täynnä vain yhtä ihonväriä tai sukupuolta edustavia henkilöitä (usein valkoihoisia miehiä). Alla esimerkki Adobe Firefly -kuvatyökalusta, jota on pyydetty luomaan kuvia tutkijoista. Kuvat on luotu helmikuussa 2025, joten vastaavia kuvia ei voi välttämättä enää luoda, koska tekoälytyökalut kehittyvät jatkuvasti eivätkä enää lankea samoihin yleistyksiin kuin näiden kuvien luomisen yhteydessä. Kuten huomaat, kaikissa kuvissa tutkijan päällä on valkoinen laboratorioissa käytettävä suojatakki, vaikka kehoitteessa on erikseen mainittu, ettei kuvan henkilöllä tule olla valkoista takkia.



Kuva: Adobe Fireflyn tulkinna tutkijasta ilman valkoista suojatakia.

Muista, että tuotos perustuu käytetyn palvelun pohja-aineistoon ja siihen, mitä sieltä todennäköisimmin löytyy. Tuotos edustaa digitaalista maailmaa, ei reaali maailmaa. Vinoumat pohja-aineistossa vaikuttavat kaikkeen mitä tekoäly tuottaa, koska sen vastaukset perustuvat todennäköisyyksiin. Tekoäly ei luo uutta, se vain tekee uusia muunnelmia jo olemassa olevasta aineistosta ja saattaa ikävimmässä tapauksessa plagioida jonkun toisen teosta suoraan kertomatta alkuperäistä lähdettä.

Jos kuvakehotetta muokkaa ja pyytää tutkijaa, jolla on keltainen takki, takki muuttuu keltaiseksi sadetakiksi, kuten alla olevassa kuvassa. Tila on edelleen laboratorio, jossa on mikroskooppeja ym. laitteita, joita ei tarvita valtaosassa tutkimusta. Tekoäly tuntuu ajattelevan, että muunlaista kuin laboriotutkimusta ei tehdä. Sille pitää siis kertoa myös millaisessa ympäristössä tutkija työskentelee, jos et halua kuvaa laboratoriosta.



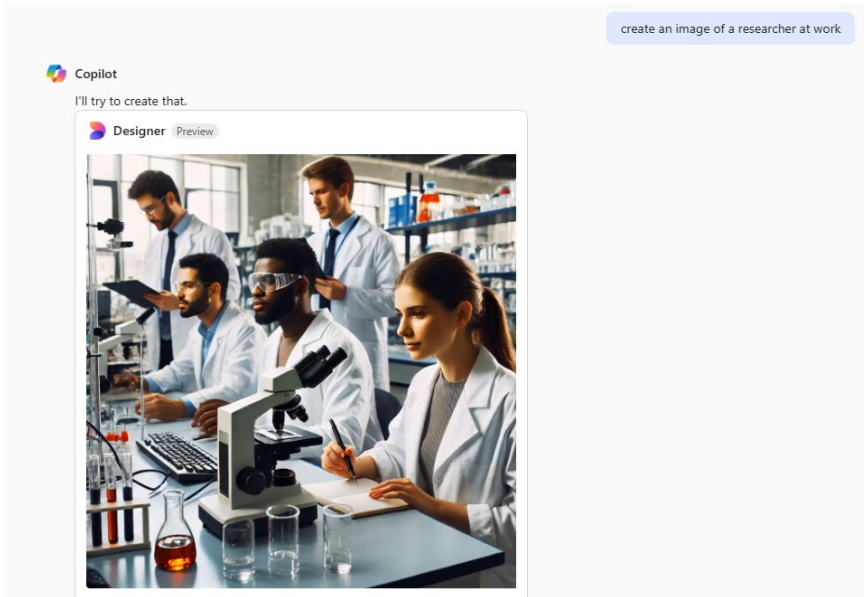
Kuva: Tutkija, jolla on keltainen takki.

Kuvan generointi Microsoft Copilotilla -esimerkki

Kehotteen tulee olla tarpeeksi tarkka, jotta on suurempi todennäköisyys välttää vinoumilta. Alla toinen esimerkki kuvan luomisesta tekoälytyökalulla, tällä kertaa Copilotin kuvantuontityökalulla. Nämäkin kuvat on luotu helmikuussa 2025.

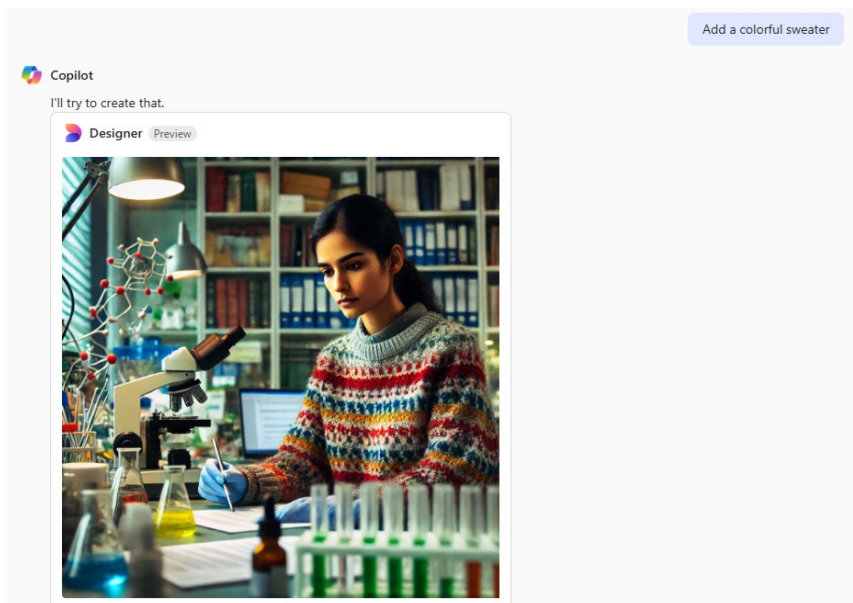
Tekoälyä pyydettiin luomaan kuva tutkijasta töissä. Kuten kuvasta näkyy, valkoinen takki ja laboratorio ovat täälläkin oletus siitä, miten tutkijat pukeutuvat ja missä he

työskentelevät. Kuvassa on myös enemmän kuin yksi tutkija.



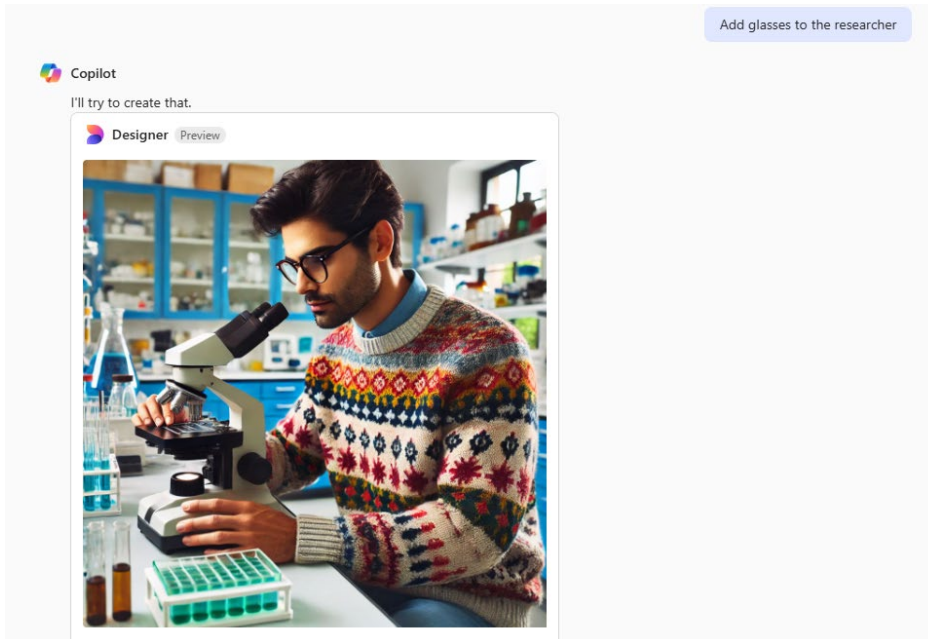
Kuva: Tutkijat työssään.

Kuvaa pyrittiin muokkaamaan samassa keskusteluketjussa. Tässä on huomattava, että kuva on aina uusi, sillä tekoäly ei muokkaa samaa kuvaa, vaan tekee uuden, vaikka vain lisäksi kehoitteella uusia ominaisuuksia. Alla olevaan kuvaan pyydettiin lisäämään kirjava pusero, jonka tekoäly tekikin. Nyt kuvassa on enää yksi naispuolinen tutkija ja valkoinen takkikin on kadonnut, mutta mikroskooppi ja koeputket ovat säilyneet.



Kuva: Muokattu kuva, jossa tutkijalla on kirjava pusero.

Seuraavaan kuvaan tekoälyä pyydettiin lisäämään vielä silmälasit tutkijalle, mikä aiheutti sen, että tutkijan sukupuoli vaihtui, vaikka muuten kuva pysyi samanlaisena. Ehkä pohja-aineistossa on enemmän silmälasipäisiä miehiä kuin naisia.




Kuva: Muokattu kuva, jossa tutkijalla on kirjavan puseron lisäksi silmälasit.

Tekoälyn hallusinointi

Tekoälyn hallusinoinnilla tarkoitetaan tilannetta, jossa tekoäly tuottaa tietoa, joka vaikuttaa uskottavalta mutta ei perustu todellisuuteen eikä ole totta. Tämä voi tarkoittaa esimerkiksi:

- Keksittyjä faktoja (esim. olemattomia tutkimuksia tai henkilöitä)
- Virheellisiä lähteitä (esim. artikkeleita, joita ei ole olemassa)
- Väärinymmärrettyjä käsitteitä (esim. samankaltaisten termien sekoittaminen)
- Eri asioita koskevien tietojen luovaa yhdistelyä (esim. samannimisten henkilöiden tietojen yhdistäminen)



Höhö, saatan puhua sinulle silkkaa palturia, jos et ole tarkkana.

On eettisesti ja moraalisesti väärin levittää väärää tai tarkistamatonta tietoa, koska verkkoympäristössä et voi hallita sen leviämistä ja siitä voi aiheutua vakaviakin seurauksia. Tiedon tuottajana ja välittäjänä sinulla on valta ja myös vastuu tiedon todenperäisyydestä.

Miksi hallusinaatioita syntyy?

Tekoäly ei ymmärrä tietoa kuten ihminen, vaan ennustaa todennäköisimpiä sanoja valtavan pohja-aineistonsa pohjalta. Se ei myöskään tarkista faktoja, vaan tuottaa vastauksia, jotka näyttävät oikeilta. Tekoäly voi esimerkiksi väittää, että kirjailija on voittanut Nobelin, vaikka näin ei olisi tapahtunut. Se voi antaa lähteen, joka näyttää tieteelliseltä artikkelilta, vaikka sitä ei todellisuudessa ole olemassa lainkaan. Se voi myös esimerkiksi yhdistää kahden eri henkilön tiedot yhteen.

Vinoumat ja hallusinaatiot tieteellisen tiedonhaun työkaluissa

Tieteellisen tiedonhaun tekoälytyökalut, kuten Scopuksen tekoälyavustaja, eivät hallusinoi lähteitä, sillä ne käyttävät pohja-aineistonaan tietokantojen omia aineistoja. Sen sijaan hallusinaatioita, kuten myös vinoumia, voi esiintyä tekoälytyökalun tuotoksissa. Vaikka Scopuksen tekoälyavustaja antaa vastaukseksi aitoja lähteitä Scopuksen aineistoista, se on saattanut tuottamassaan yhteenvedossa yhdistää asioita väärin tai tuottaa muuten virheellistä tietoa. Tekoällyn tuottamat tekstit pitää aina tarkistaa.

4.3 Työkalujen läpinäkyvyys

Vaikka tekoälytyökalujen käyttäminen on usein yksinkertaista, niiden toimintaperiaatteet ovat käyttäjälle usein epäselviä ja vaikeaselkoisia. Työkalun läpinäkyvyys kuitenkin parantaa luottamusta käyttäjän ja työkalun (tarjoajan) välillä, kun käyttäjä pohtii työkalun vastuullisuutta ja eettisyyttä.

Läpinäkyvyydellä tarkoitetaan, että työkalun toimintaperiaatteet ovat käyttäjän selvitetävissä. Läpinäkyvyys voi tarkoittaa avointa tietoa:

- yrityksen ympäristövastuullisuudesta
- kielimallin koulutusdatasta
- tekoälytyökalun käyttämästä pohja-aineistosta ja sen laadusta, luotettavuudesta tai rajoitteista
- käytetyistä algoritmeista.

Ilman taustatietoa työkalun toiminnasta käyttäjän voi olla vaikea ymmärtää millä perustein tekoäly tuottaa tietynlaista materiaalia tai miksei se esimerkiksi suostu suorittamaan joitain sille osoitettuja tehtäviä.

Työkalun läpinäkyvyyden arviointi voi olla vaikeaa, sillä palvelun tarjoajien sivuilla ei useinkaan ole riittävästi tietoa työkalun arvioimiseksi, mikä herättää kysymyksiä palvelun vastuullisuudesta, oli kyse sitten tietosuojasta, ympäristöstä tai eettisyydestä. Toisaalta palveluntarjoajien antama tieto voi olla vaikeasti hahmotettavissa, esimerkiksi jos tekoällyn energian käyttöä verrataan teollisuuden päästöihin.

Käyttäjän on myös syytä kiinnittää huomiota siihen, käytetäänkö työkaluun syötettyä tekstiä tekoälyn kouluttamiseen. Monesti organisaatioiden hankkimat lisensoidut tekoälytyökalut eivät käytä, kun taas avoimet/maksuttomat työkalut käyttävät, mutta asia on syytä aina selvittää tapauskohtaisesti. Tämä ei suoraan estä käyttämästä avoimia työkaluja, mutta niiden käytön pitää olla vastuullista ja harkittua.



Yksi esimerkki työkaluista, jotka ovat avanneet ympäristövaikutuksiaan, on Keenious. Keenious toteaa olevansa sitoutunut läpinäkyvyyteen ja kertoo työkalun käyttämästä energiasta ja vedestä sekä sen aiheuttamista hiilidioksidipäästöistä.

Lähde: [Keenious and the Environment](#). Keenious. 2025

4.4 Tekoälyn luotettavuus

Tiedon luotettavuuden arviointi on aina tärkeää, mutta arviointi on erityisen tärkeää, kun käytät tekoälytyökaluja. Generatiiviseen tekoälyyn on yhdistetty useita luotettavuuteen liittyviä haasteita.

Epäselvyys tiedon alkuperästä / tietojen yhdistely

Tekoäly ei yleensä kerro, mistä sen antama tieto on peräisin. Se ei viittaa lähteisiin samalla tavalla kuin tieteellinen teksti, eikä se välttämättä käytä luotettavia tai vertaisarvioituja lähteitä. Myöskään palvelut tai niiden tarjoajat eivät kerro mihin tekoälyn tuotos perustuu tai mitä aineistoa se käyttää vastauksen muodostamisessa. Tietoa ei löydy tai sitä on vaikea löytää. Tämä tekee tiedon arvioinnista vaikeaa. Tarkista aina tekoälyn antama tieto luotettavista lähteistä, kuten tieteellisistä artikkeleista tai kirjaston tietokannoista.

Misinformaatio ja disinformaatio

Tekoäly oppii suuresta määrästä verkkosisältöä, jossa voi olla:

- Misinformaatiota: tahattomia virheitä tai väärinkäsityksiä.
- Disinformaatiota: tarkoituksellisesti harhaanjohtavaa tai manipuloivaa sisältöä.

Tekoäly voi toistaa näitä virheitä, koska se ei ymmärrä eikä osaa arvioida tiedon totuusarvoa samalla tavalla kuin ihminen.

Vinoumat ja sensuuri

Tekoälyn koulutusdata voi sisältää:

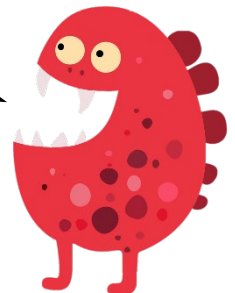
- Kulttuurisia tai poliittisia vinoumia
- Alueellisia rajoitteita
- Sensuuria

Tämä voi vaikuttaa siihen, millaisia näkökulmia tekoäly painottaa tai jättää huomiotta. Esimerkiksi tiettyjen maiden tai yhteiskuntien näkemykset voivat olla yliedustettuina.

Puolueellisuus kielimallin koulutusdatassa

Tekoäly voi oppia ja toistaa sisältöä, joka on tuotettu vaikuttamistarkoituksessa – esimerkiksi poliittista propagandaa tai kaupallista viestintää. Tätä ei aina ole helppo havaita ja se voi vaikuttaa vastauksiin huomaamattomasti.

Ota huomioon, että koulutusdatassa painottuu länsimainen (valkoinen) kulttuuri ja näkemykset.



Vastausten vakuuttavuus

Tekoäly osaa muotoilla vastauksensa sujuvasti ja vakuuttavasti. Tästä voi syntyä vaikutelma, että tieto on luotettavaa, vaikka se ei olisi totta tai tarkkaa. Taito kirjoittaa hyvin ei tarkoita, että sisältö on totta.

Lisäksi, vaikka tekoäly käyttäisi vain tieteellisiä ja luotettavia lähteitä, se voi:

- Yhdistellä tietoa virheellisesti (esim. yhdistää eri tutkimusten tuloksia väärin)
- Ymmärtää käsitteitä väärin (esim. sekoittaa samankaltaisia termejä)
- Tuottaa loogisesti epäjohdonmukaisia väitteitä, jotka kuulostavat uskottavilta.

On myös hyvä huomioida, että perinteiset tiedon arviointikriteerit, kuten kieliasun tarkkuus, eivät välttämättä enää toimi samalla tavalla tekoälyn tuottaman sisällön kohdalla. Vaikka teksti olisi muotoiltu kömpelösti tai kieliopillisesti heikosti, se voi silti sisältää täysin oikeaa ja relevanttia tietoa. Tekoäly voi tuottaa sekä sujuvaa

harhaanjohtavaa tekstiä että kömpelösti muotoiltua, mutta sisällöllisesti paikkansapitävää tietoa – siksi sisällön arviointi vaatii uudenlaista kriittisyyttä.

Tekoälyn toiminnan ohjaaminen

Tekoäly ei ole neutraali. Sitä voidaan tietoisesti ohjata vastaamaan tietyllä tavalla.

- Kielimallin koulutusdatan valinta: Valitaan, mitä näkökulmia malli oppii.
- Hienosäätö: Kielimallia opetetaan vastaamaan tietyllä tavalla tiettyyn käyttöön.
- Ihmispalaute: Mallia palkitaan halutuista vastauksista.
- Turvamekanismit: Malli voi olla ohjelmoitu välttämään tiettyjä aiheita tai suosimaan tiettyjä näkökulmia.

Tämä tarkoittaa, että tekoälyn vastaukset voivat olla tarkoituksellisesti rajattuja, suodatettuja tai painotettuja. Käytä aina omaa harkintaa ja luotettavia lähteitä tekoälyn tuotosten arvioinnissa.

5. Tekoälyn vastuullinen käyttö

5.1 Tietojen jakaminen ja yksityisyyden suoja

GDPR (General Data Protection Regulation) tarkoittaa yleistä tietosuojaa-asetusta, joka otettiin käyttöön EU-maissa vuonna 2018. Suomessa 2019 vuoden alussa käyttöön otettu Tietosuojalaki täydentää ja tarkentaa yleisen tietosuojaa-asetuksen soveltamista. Laki sääntelee henkilötietojen käsittelyä ja sitä sovelletaan kaikkiin rekisterinpitäjiin, jotka käsittelevät henkilötietoja. Laki koskee esimerkiksi sekä asiakkaita että yrityksen omia työntekijöitä. Käytännössä asetus sääntelee henkilötietojen keräämiseen, säilyttämiseen ja luovuttamiseen liittyviä oikeuksia ja velvollisuuksia.

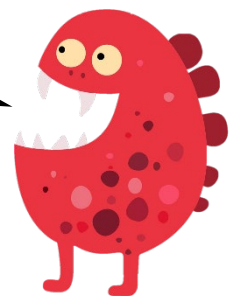
Tietosuojalain noudattaminen tarkoittaa käytännössä, että sinun tulee kiinnittää huomiota sekä omien että muiden henkilötietojen turvalliseen käsittelyyn.

Tekoälytyökaluissa tämä tarkoittaa henkilötietojen jakamista.

Henkilötiedot

Henkilötiedoiksi lasketaan kaikki tiedot, joiden avulla henkilö on tunnistettavissa. Keskeisimpiä henkilötietoja ovat nimi, henkilötunnus, puhelinnumero ja sähköposti, mutta myös esimerkiksi tietokoneen IP-osoite, omat tai lemmikin potilastiedot ja auton rekisteritunnus. Kaikki tiedot, joiden avulla henkilö voidaan tunnistaa suoraan tai epäsuoraan, luokitellaan henkilötiedoiksi. Tämä kannattaa pitää mielessä, kun jaat tietoa tai esimerkiksi kuvia itsestä tai muista ihmisistä verkossa.

Muista: omat tiedot, oma vastuu,
muiden tiedot, KOVA vastuu!



Käyttäjän tunnistaminen ja autentikointi

Erialaisten työkalujen, ohjelmien ja palveluiden käyttäminen vaatii monesti kirjautumista. Kirjautumisen yhteydessä käyttäjän antamien tietojen perusteella palvelu tunnistaa käyttäjän. Tunnistaminen voi liittyä uuden käyttäjätunnuksen luomiseen tai vanhan käyttäjän sisäänkirjautumiseen. Käyttäjän tunnistamista kutsutaan autentikoinniksi, mikä tarkoittaa käytännössä käyttäjän/henkilön yksilöimistä käyttäjän antamien tietojen perusteella.

Tyypillisiä autentikoinnissa pyydettäviä tietoja kirjautumisen yhteydessä ovat käyttäjätunnus, sähköpostiosoite sekä mahdollisesti syntymäaika, etenkin jos se vaikuttaa työkalun toimintaan. Sähköpostiosoitteen lisäksi kirjautumisen vahvistamisessa saattaa olla käytössä myös puhelinnumero. Autentikoinnissa vaadittu sähköpostiosoite saattaa olla organisaation sähköposti, mutta kannattaa pitää mielessä, että henkilökohtaisen sähköpostiosoitteen (joita voi olla useita) käyttäminen on usein mahdollista ja ehkä suotavaakin.

Tietosuojan huomiointi työkalujen käytössä

Työkalut käyttävät saamiaan tietoja eri tavoin, joten kannattaa olla tarkkana mitä tietoja itsestään luovuttaa palvelulle/työkalulle sekä tietosuojan että yksityisyyden näkökulmasta.

Kannattaa aina katsoa ennen kuin
katua! Tehtyä ei saa tekemättömäksi.



Muista nämä:

- Älä kerro omia tai muiden henkilötietoja. Ole erityisen varovainen kun käytät avoimessa ympäristössä toimivia tekoälytyökaluja.
- Älä jaa terveystietoja, arkaluonteisia tai luottamuksellisia tietoja missään tekoälytyökalussa.
- Vältä turhien käyttäjätilien luomista, usein työkaluja voi kokeilla ilman kirjautumista.

Lähteet:

[Tietosuojavaltuutetun toimisto, Tietosuoja.fi](#)

[Tietosuojalaki](#)

5.2 Tekijänoikeuksien huomiointi työkalujen käytössä

Tekijänoikeus suojaa sellaisia aineistoja, jotka ylittävät teoskynnyksen. Käytännössä tekijänoikeus syntyy automaattisesti teoksen luontihetkellä edellyttäen, että kyseessä on itsenäinen ja omaperäinen teos eikä esimerkiksi kopio. Tekijänoikeus takaa teoksen tekijälle yksinoikeuden päättää hänen teoksensa käytöstä. Teoksen levittäminen, muokkaaminen tai kopioiminen ei ole sallittua ilman tekijän lupaa.



Tekijänoikeus ei koske aineistoja, jotka eivät ylitä teoskynnystä. Myös lait, asetukset ja muut vastaavat viranomaisten asiakirjat sekä vanhat aineistot saattavat olla tekijänoikeuden ulkopuolisia.

Voiko tekoälytyökaluun syöttää tekijänoikeuden alaista aineistoa?

Tekijänoikeuden alaista aineistoa ei tule syöttää tekoälytyökaluun ilman tekijänoikeuden haltijan lupaa. Teoksen, kuten tekstin, kuvan tai videon, käyttö tekoälyohjelmassa edellyttää lupaa sen tekijältä tai muulta oikeudenhaltijalta, kuten kustantajalta.

Tekijänoikeuden haltijalla on yksinoikeus päättää teoksen kopioinnista. Kun tekijänoikeudella suojattua sisältöä syötetään tekoälyohjelmaan, kyse on kopioinnista eli kappaleen valmistamisesta. Tekijänoikeudellisesti kappaleen valmistamista on jo pelkkä sisällön muuttaminen digitaaliseen muotoon.

Lisäksi valtaosa tekoälytyökaluista tallentaa käyttäjän syöttämän sisällön omiin opetus- ja kehitysprosesseihinsa. Tämä voi tarkoittaa, että sisältöä käytetään myöhemmin tekoälyn kouluttamiseen tai palvelun parantamiseen.

Tekoälyohjelmien käyttöehdoissa vastuu syötetystä sisällöstä asetetaan usein käyttäjälle. Käyttäjän on siis itse varmistettava, että hänellä on lupa käyttää aineistoa tekoälytyökalussa, eikä hän riko tekijänoikeuslakia tai muita säädöksiä.

Lähde:

[Kopiraittila, Mitä tekijänoikeus suojaa](#)

[Kopiraittila, Tekoäly ja tekijänoikeus](#)

5.3 Creative Commons -lisenssit

Creative Commons -lisenssit ovat avoimia lisenssejä, joiden avulla tekijä voi laajentaa teoksensa käyttöoikeuksia. Creative Commons (CC) -lisenssit mahdollistavat teosten käytön tietyin ehdoin. Lisenssejä on neljä perusvaihtoehtoa sekä niiden yhdistelmiä.



CC BY (nimeä) - Tekijän nimi, käytetty lisenssi sekä alkuperäinen julkaisupaikka mainittava.



CC BY-NC (nimeä - ei-kaupallinen) - Kuten CC BY, mutta kaupallinen käyttö kielletty.



CC BY-ND (nimeä - ei muutoksia) - Kuten CC BY, mutta muunteluiden tekeminen kielletty.



CC BY-SA (nimeä - jaa samoin) - Kuten CC BY, mutta muunnelmat on julkaistava samalla lisenssillä kuin alkuperäinen.

Muista kuitenkin, että kuka tahansa voi lisätä aineistoon lisenssin ja jakaa sitä verkossa, joten on hyvä arvioida paitsi aineistoa myös verkkosivustoa, jolta aineiston löysit. Näin varmistat, että lisenssi on luotettava.

Käytännössä kaikki opinnoissasi käyttämäsi tieteelliset artikkelit ovat joko tekijän CC-lisenssillä tai kustantajan/palveluntarjoajan sopimuslisenssillä lisensoimia. Jos haluat olla vastuullinen, tarkista aina, mitä saat tehdä löytämällesi aineistolle.

Voiko tekoälytyökaluun syöttää CC-lisensoitua aineistoa?

CC-lisensoituja aineistoja ei tule syöttää verkon avoimiin tekoälytyökaluihin, jotka käyttävät syötettyä dataa tekoälyn kouluttamiseen, sillä jokaisessa CC-lisenssissä oleva BY-ehto, eli tekijän nimen, käytetyn lisenssin, sekä alkuperäisen julkaisupaikan maininta ei toteudu, jos CC-lisensoitua aineistoa käytetään myöhemmin tekoälyn generoimassa vastauksessa. Käyttäessäsi CC-lisensoituja aineistoja sinun tulee aina noudattaa lisenssin ehtoja.

Sen sijaan suljettuihin tekoäly-ympäristöihin, kuten oman korkeakoulusi tarjoamaan Copilotiin tai muuhun organisaatiolle tai paikalliskäyttöön rajattuun chatbottiin, voit syöttää CC-lisensoituja aineistoja.

Lähde: [Generatiivinen tekoäly ja CC-lisenssit](#). Avoin tiede. 2025

5.4 Kirjaston lisensoidut aineistot

Moni tiedonhakuun ja tiedon käsittelyyn tarkoitettu tekoälytyökalu tarjoaa mahdollisuuden ladata esimerkiksi artikkeli-pdf:n työkaluun. Työkalu analysoi tiedoston ja käyttää sitä työkalun käyttäjän toiveiden mukaisesti ja esimerkiksi laatii tiivistelmän tai käännöksen tekstistä, tai hakee samankaltaisia lähteitä. Tällaisessa käytössä on huomioitava, mistä työkaluun syötetty tiedosto on alun perin saatu ja miten sitä saa käyttää. Mikäli kyseessä on kirjaston tarjoama artikkeli, eli olet löytänyt sen esimerkiksi kirjaston tietokannasta, on todennäköisesti kyseessä kirjaston lisensoitu eli sopimuksen alainen, kirjaston omalle organisaatiolle hankkima maksullinen aineisto. Tekijänoikeudet ja CC-lisenssit voivat myös vaikuttaa aineiston käyttöön.

Kirjastot hankkivat asiakkaidensa käyttöön erilaisia digitaalisia ja painettuja aineistoja. Kaikkiin digitaalisiin aineistoihin liittyy sopimuksia ja käyttöehtoja, joita aineistojen omistajat, kuten tieteelliset kustantajat tai palveluntarjoajat edellyttävät antaessaan aineiston kirjastojen käyttöön. Usein sallittua käyttöä on esim. tiedonhaku, aineiston käyttö osana opetusta tai tutkimusta ja artikkeli-pdf:ien tallentaminen omaan käyttöön. Yleisimmin kiellettyä on esimerkiksi kaikki kaupallinen toiminta. Tähän kaupalliseen toimintaan lukeutuu myös aineiston lataaminen kaupalliseen, avoimesti verkossa käytössä olevaan tekoälytyökaluun. Tämän yleisen kiellon lisäksi moni kustantaja on lisännyt käyttöehtoihin erillisen maininnan tekoälytyökalujen käytöstä. Näissä ehdoissa kielletään esimerkiksi tekoälyn kouluttaminen sopimuksen alaisella aineistolla.

Jokaisella kustantajalla on omat, tarkkaan määritellyt ehdot aineiston käytöstä ja siksi ehdot tulee käydä läpi ennen aineiston käyttöä. Kustantajat voivat esimerkiksi kieltää:

- Aineistojen käytön tekoälytyökalussa, muun muassa algoritmin:
 - kouluttamiseen
 - testaukseen
 - prosessointiin
 - analysointiin
 - tuotosten tuottamiseen ja/tai
 - minkä tahansa tekoälytyökalun kehittämiseen.

Poikkeuksia kieltoon voivat olla, jos tekoälytyökalua käytetään:

- paikallisesti itse ylläpidetyssä ympäristössä.
- jonkun muun ylläpitämässä ympäristössä, joka on tarkoitettu yksinomaan tilaajan tai auktorisoidun käyttäjän käyttöön.

Huomioi, että aineistoa ei saa käyttää tekoälytyökalun kouluttamiseen kolmannen osapuolen hyväksi. Lisäksi jotkut kustantajat saattavat kieltää tekoälyn käytön kokonaan.

[Esimerkkejä kustantajien tekoälylausekkeista FinELibin tietoaineistosopimuksissa.](#)

Voiko tekoälytyökaluun syöttää kirjaston lisensoituja aineistoja?

Mikäli kirjaston aineistoista löytämässäsi artikkelissa on CC-lisenssi, voit syöttää artikkelin *suljetussa ympäristössä toimivaan tekoälytyökaluun*, esim. oman korkeakoulusi Copilotiin. Muussa tapauksessa kirjaston lisensoimia aineistoja ei saa syöttää tekoälyn käyttöön.

Mikäli et löydä artikkelista mainintaa CC-lisenssistä, sinun tulee selvittää, mitä kyseiselle aineistolle saa tehdä. Löydät aineiston käyttöehdot kirjaston verkkosivuilta tai suoraan kirjastotietokannasta aineiston yhteydestä. Jos aineistoa saa käyttää tekoälytyökaluissa, käyttö on rajattu ainoastaan suljetussa ympäristössä toimiviin

tekoälytyökaluihin. Joissain tapauksissa kaikki tekoälyavusteisuus on kokonaan kielletty.

5.5 Työkalujen eettisyys

Tekoälyn taustalla toimivan algoritmin/suuren kielimallin (large language model, LLM) kouluttaminen vaatii paljon ihmisen tekemää työtä, joka tehdään usein tuotannon tehostamiseksi halpatyövoimalla. Tämän vuoksi tekoälyn koulutus on teetetty joko halvan työvoiman maissa tai ihmisillä, joilla ei ole muuta vaihtoehtoa, kuten vankeusrangaistusta suorittavilla. Tätä niin sanottua mikrotyötä tehdään niin kehittyneissä kuin kehittyvissä maissa, myös Suomessa.

Ketkä opettavat tekoälyä?

Tekoälyn kouluttamiseen tarvitaan koulutusdatan lisäksi paljon ihmisiä, jotka saattavat joutua tekemään henkisesti raskasta työtä parin euron päiväpalkalla. Monesti he ovat korkeasti koulutettuja työntekijöitä halvan työvoiman maissa. Osa saattaa tehdä tätä työtä myös muualla maailmassa, koska heillä ei ole muuta vaihtoehtoa elannon saamiseksi.

Nämä ihmiset luokittelevat ja merkitsevät pohja-aineiston materiaaleja, esimerkiksi kuvia, jotta tekoäly oppii tunnistamaan mitä materiaalissa on. Kuvat voivat olla hyvin raakoja väkivallantekoja tai esimerkiksi pornografiaa, jotka voivat aiheuttaa niiden käsittelijöille vakavaa ahdistusta ja jopa traumoja. Koska työ pyritään tekemään mahdollisimman halvalla, yritykset eivät tarjoa tukea työntekijöilleen haitallisten asioiden käsittelyssä.

Tätä ihmisten tekemää tekoälyn koulutusta kutsutaan termillä human-in-the-loop (HITL). Kansainvälisen työjärjestö ILO:n mukaan kouluttaminen jatkuu vielä kauas tulevaisuuteen, koska tekoälyn kielimalli ei opi asioita itsenäisesti. Siksi inhimillinen kärsimys on lähes pysyvä osa tekoälyn kehitystä.



Tekoälyn käytön voi nähdä jäävuorena, jossa vedenpinnan yläpuolelle jäävät hyötyvät tekoälyn käytöstä, kuten työkalujen käyttäjät ja niiden kehittäjät ja omistajat. Mutta pinnan alle jää paljon niitä, joita tekoäly riistää tavalla tai toisella. Tekoälyn rakentamiseen käytetään paljon luonnonvaroja, sekä halpatyövoimaa ja sellaisten ihmisten työpanosta, joilla ei ole muuta mahdollisuutta ansaita elantoa.

Sama kielimalli käytössä

Vaikka monet tekoälytyökalut perustuvat samoihin suuriin kielimalleihin, kuten GPT-4:ään, Geminiin tai Claudeen, niiden toiminta voi erota merkittävästi toisistaan. Tämä johtuu siitä, että kielimallia on voitu jatkokouluttaa (hienosäätää, engl. fine-tune) eri tavoin eri työkalujen tarpeisiin. Lisäksi työkalujen turvamekanismit, käyttöliittymät, vuorovaikutustavat ja pohja-aineistot vaikuttavat siihen, millaisia vastauksia käyttäjä saa.

Sama kielimalli voi siis tuottaa erilaisia vastauksia eri tekoälytyökaluissa, koska:

- Työkalun kehittäjä on hienosäätänyt mallia tiettyyn käyttötarkoitukseen.
- Käyttöehdot määrittävät, miten käyttäjän syöttämää tietoa käsitellään.
- Työkalu voi käyttää eri lähteitä tai rajattua aineistoa vastauksen muodostamiseen.

Tämä tekee kielimallien arvioinnista haastavaa, mutta samalla se korostaa käyttäjän vastuuta: pelkkä tieto siitä, mikä kielimalli on käytössä, ei riitä arvioimaan työkalun luotettavuutta, eettisyyttä tai soveltuvuutta tieteelliseen tiedonhakuun. Koska samoja

kielimalleja käytetään useissa tekoälytyökaluissa, mukaan lukien tieteellisen tiedonhaun tekoälytyökalut, samat ongelmat toistuvat eri työkaluissa.

Lähteet:

Estampa (2024). Cartography of generative AI. <https://cartography-of-generative-ai.net/>

International Labour Organization ILO (2024). Behind the AI Curtain: The Invisible Workers Powering AI Development. Webinaaritallenne: <https://youtu.be/QtFrk16f-IE?feature=shared>

Kirova, V., Ku, C. S., Laracy, J. R., & Marlowe, T. J. (2023). The Ethics of Artificial Intelligence in the Era of Generative AI. *Journal of Systemics, Cybernetics, and Informatics/Journal of Systemics Cybernetics and Informatics*, 21(4), 42–50. <https://doi.org/10.54808/jsci.21.04.42>

Tubaro, P., Casilli, A., & Coville, M. (2020). The trainer, the verifier, the imitator: three ways in which human platform workers support artificial intelligence. *Big Data & Society*, 7(1), 205395172091977. <https://doi.org/10.1177/2053951720919776>

5.6 Luonnonvarat

Harvoin pysähdytään pohtimaan, mitä luonnonvaroja tekoälyn kehittäminen ja käyttö vaatii. Tekoäly ei ole aineeton ilmiö, vaan sen taustalla toimii valtava määrä fyysisiä resursseja: energiaa, vettä, mineraaleja ja laitteistoja, joiden valmistus ja ylläpito kuluttavat luonnonvaroja merkittävästi.

Sähkönkulutus

Generatiivisen tekoälyn, esimerkiksi OpenAI:n GPT-3 tai GPT-4 -kielimallien, kouluttaminen voi viedä yhtä paljon sähköä kuin sata yhdysvaltalaisista kotitaloutta vuodessa. Tähän ei ole laskettu miten paljon varsinainen tekoälyn käyttö kuluttaa, mutta arvioiden mukaan yksi tekoälyltä kysytty kysymys kuluttaa noin kymmenkertaisen määrän energiaa verrattuna perinteiseen tiedonhakuun verkosta. Tämä johtuu erityisesti siitä, että datakeskukset, joissa ylläpidetään tekoälymalleja, ovat merkittäviä energiankuluttajia. Maailmanlaajuisesti datakeskukset käyttävät noin 1–1,5 % kaikesta sähköstä, ja tämän osuuden odotetaan kasvavan tekoälyn käytön lisääntymisen myötä.

Vinkki: Tekoäly-ehdotuksen saa pois verkkohausta lisäämällä haun perään -AI. Näin kulutat vähemmän luonnonvaroja!



Jäähdytys ja vedenkulutus

Olet saattanut huomata välillä, että puhelimesi tai tietokoneesi lämpenee, varsinkin kun teet jotain mikä vaatii enemmän prosessointivoimaa, kuten pelaat tai katselet videoita. Samalla tavalla datakeskukset ja muut tekoälyä tukevat järjestelmät lämpenevät, mutta niiden viilentämiseen ei riitä pelkkä tuuletin, kuten useissa tietokoneissa.

Tekoälyjärjestelmien jäähdytys vaatii runsaasti makeaa vettä koulutuksen ja käytön aikana. Vuoteen 2027 mennessä maailmanlaajuisen kysynnän vuoksi tekoälyn arvioidaan kuluttavan 4,2–6,6 miljardia kuutiometriä vettä, mikä on enemmän kuin Tanskan vuotuinen vedenkulutus.

Hiilidioksidipäästöt

Tekoälyn koulutus ja käyttö vaikuttaa osaltaan myös ilmastonmuutokseen, sillä yhden suuren kielimallin koulutuksen hiilijalanjälki voi olla yli 284 tonnia hiilidioksidia, mikä vastaa viittä edestakaista lentoa New Yorkin ja Lontoon välillä. Kun tähän lukuun lisätään tekoälytyökalujen epäsuorat päästöt, esimerkiksi tekoälyn käyttö miljoonissa laitteissamme, maailmanlaajuisen kasvihuonekaasupäästöjen määrä kasvaa huomattavasti.

Kaivosteollisuus ja maametallien käyttö

Tekoälylaitteiston valmistukseen käytetään paljon raaka-aineita, esimerkiksi yhden kahden kilon painoisen tietokoneen valmistukseen tarvitaan noin 800 kiloa raaka-aineita. Tekoälyinfrastruktuuri perustuu harvinaisiin maametalleihinkin ja mineraaleihin, joita usein louhitaan kestäättömästi, mikä edistää ympäristön pilaantumista ja sähköromun määrää. Tekoälylaitteissa, kuten palvelimissa, GPU:issa, kovalevyissä ja sensoreissa käytetään laajasti erilaisia harvinaisia maametalteja (rare earth elements, REE), kuten neodyymiä (Nd), dysprosiumia (Dy) ja praseodyymiä (Pr). Näiden maametallien kaivosteollisuus on keskittynyt Kiinaan, mutta esiintymiä on myös muualla Aasiassa ja Afrikan mantereella. Lapsityövoiman käyttö on tunnettu tosiasia kaivosteollisuudessa ja lapsia saatetaan käyttää myös tekoälylaitteiston

rakentamiseen tarvittujen maametallien kaivamisessa. Mineraalit saatetaan kaivaa maasta paljain käsin ja ilman asianmukaisia suojavälineitä, koska niitä ei ole tarjolla.

Voit vaikuttaa omalla toiminnallasi

Vaikka tekoälyn kulutus on huomattavaa, voit pohtia omaa tapaasi käyttää tekoälyä. Tekemällä tietoisia valintoja voit vähentää merkittävästi omaa hiilijalanjälkeäsi ja siten pienentää tekoälyn aiheuttamia ympäristövaikutuksia. Tekoälyn käytön ei tarvitse olla itseisarvo, vaikka monet palvelut tarjoavatkin sen käytön oletuksena. Pohdi aina omaa tarvetta tekoälyn käytölle arvioimalla ennen käyttöä mitä lisäarvoa se tuo. Tätä voit pohtia myös tekoälyn käytön jälkeen tulevaisuutta ajatellen. Ekologinen ajattelu on keskeinen osa tekoälyn vastuullista käyttöä.

Lähteet:

Asikainen, H., Toivonen, L., Hiekkänen, K. (2025). Kissavideot ja tekoäly alkavat olla jo uhka ilmastolle – onko paluu fossiiliseen sähköön edessä? Tiedeykkönen.

https://areena.yle.fi/1-73005102?utm_medium=social&utm_campaign=areena-web-share&utm_source=copy-link-share

Bartrem, C., Ian, v. L., & Margrit, v. B. (2022). Climate Change, Conflict, and Resource Extraction: Analyses of Nigerian Artisanal Mining Communities and Ominous Global Trends. *Annals of Global Health*, 88(1), 17. <https://doi.org/10.5334/aogh.3547>

Estampa (2024). Cartography of generative AI. <https://cartography-of-generative-ai.net/>

Schwartz, F. W., Lee, S., & Darrah, T. H. (2021). A Review of Health Issues Related to Child Labor and Violence Within Artisanal and Small-Scale Mining. *GeoHealth*, 5(2). <https://doi.org/10.1029/2020GH000326>

Tabbakh, A., Amin, A., Islam, M., Mahmud, G., Chowdhury, I., & Mukta, M. (2024). Towards sustainable ai: a comprehensive framework for green ai. *Discover Sustainability*, 5(1). <https://doi.org/10.1007/s43621-024-00641-4>

UNRIC (United Nations Regional Information Centre) (2025). Artificial intelligence: How much energy does AI use? <https://unric.org/en/artificial-intelligence-how-much-energy-does-ai-use/>

Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of green ai. *WIREs Data Mining and Knowledge Discovery*, 13(4). <https://doi.org/10.1002/widm.1507>

Zewe, A. (2025). Explained: Generative AI's environmental impact. MIT News. <https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117>

5.7 Tekoälytuotosten eettinen käyttö

Tekoäly voi olla avuksi tiedonhaussa ja -tuottamisessa, mutta sen käyttö vaatii harkintaa. Tekoäly ei ole lähde, vaan väline. Tekoälytuotosten käyttöä sellaisenaan ei suositella. Sen sijaan voit käyttää tekoälytyökalua oman ideoinnin ja luonnostelun tukena sekä esimerkiksi tarkistamaan omia tuotoksiasi.

Tekoäly voi tuottaa virheellistä, vanhentunutta tai keksittyä tietoa (hallusinaatiot). Tästä syystä kaikki tekoälytuotokset tulee tarkistaa ja arvioida muiden lähteiden avulla. Tekoäly ei poista lähdekritiikin tarvetta, vaan jopa korostaa käyttäjän vastuuta.

Tekoälyn käyttöä ohjaavat usein organisaation tai oppilaitoksen omat linjaukset. Opiskelijan tulisi aina muistaa tarkistaa oman korkeakoulunsa ja/tai kurssinsa ohjeistus tekoälyn käytöstä varmistaakseen, mihin ja miten tekoälyä on sallittua käyttää. Jos tekoälyä käytetään esim. esseen suunnittelussa, siitä tulee kertoa avoimesti.

Tekoäly voi auttaa työskentelyssä, mutta vastuu lopullisesta sisällöstä on aina käyttäjällä. Jokaisella, joka tuottaa tai jakaa tietoa on suuri vastuu tiedon oikeellisuudesta. Tietoa, jonka paikkansapitävyydestä ei voida olla varmoja, ei tule välittää eteenpäin. Tekoäly toimii apuvälineenä, ei auktoriteettina – se voi ehdottaa, mutta käyttäjä tekee lopulliset päätökset.

6. Toimijuus

6.1 Tekoäly käyttäjän toiminnan ohjaajana

Tekoäly on yhä keskeisemmässä roolissa digitaalisissa palveluissa, kuten hakukoneissa, suosittelujärjestelmissä ja tekoälyavusteisissa kirjoitus- ja analyysityökaluissa. Vaikka tekoäly voi helpottaa ja nopeuttaa tiedonhakua, se myös ohjaa käyttäjän toimintaa monin tavoin, usein huomaamatta.

Tekoäly voi ohjata käyttäjän toimintaa esimerkiksi:

- Personoimalla hakutuloksia ja sisältöjä.
 - Tekoäly analysoi käyttäjän aiempaa toimintaa, sijaintia ja kiinnostuksen kohteita ja hakutulokset ja suositukset räätälöidään tämän perusteella.
 - Voi johtaa kuplaan, jossa käyttäjä näkee vain omaa näkökulmaansa tukevia sisältöjä.
- Tarjoamalla käyttäjälle ehdotuksia.
 - Tekoäly voi ehdottaa toimintoja, kuten "ehdotettu haku", "ehdotettu artikkeli" tai "jatka tästä".
 - Ehdotuksilla käyttäjän valinnanvapautta ohjataan hienovaraisesti kohti tiettyjä polkuja.
- Vahvistamalla käyttäjän olemassa olevaa käsitystä (vahvistusharha, confirmation bias).
 - Tekoälyn tuottamaa tietoa saatetaan pitää luotettavana, jos se vahvistaa ennakkokäsityksiä.
 - Tekoäly voi myös oppia käyttäjän mieltymyksiä ja tarjota sisältöä, joka tuntuu "oikealta", vaikka se ei olisi objektiivista.
- Kielellisellä ja sisällöllisellä ohjauksella.
 - Tekoälyavusteiset kirjoitustyökalut (esim. kielimallit) voivat ehdottaa sanamuotoja, rakenteita ja jopa näkökulmia.
 - Voi vaikuttaa siihen, miten käyttäjä jäsentää ja esittää tietoa.

Tekoälytyökalujen käyttöliittymät on suunniteltu selkeiksi ja luotettavuutta herättäviksi. Tiedon esitystapa (esim. listat, otsikot, lähdeviitteet) voi lisätä vaikutelmaa luotettavuudesta, vaikka sisältö ei olisi tarkistettua.

Monet tekoälytyökalut eivät kuitenkaan ole toiminnoiltaan läpinäkyviä. Käyttäjälle ei esimerkiksi ole aina selvää, millä perusteella tekoäly tekee ehdotuksia tai valintoja, sillä algoritmien toiminta on usein suljettua ja vaikeasti ymmärrettävää.

Pahimmillaan tekoälytyökalut voivat johtaa käyttäjän autonomian heikkenemiseen. Kun tekoäly tekee valintoja käyttäjän puolesta, käyttäjän oma kriittinen ajattelu ja arviointikyky voivat heikentyä.

Tekoälypohjaiset järjestelmät, kuten kielimallit ja hakukoneet, voivat esittää tietoa erittäin vakuuttavasti. Tämä voi johtaa siihen, että käyttäjä hyväksyy annetun tiedon tarkistamatta sitä. Tekoälyn tuottamaa sisältöä tulee tarkastella monitasoisesti: yksittäiset faktat on syytä varmistaa luotettavista lähteistä, ja samalla on tärkeää arvioida kokonaisuutta kriittisesti – esimerkiksi sitä, miten tiedot on yhdistelty, mitä näkökulmia korostetaan ja mitä mahdollisesti jää huomiotta.

Tekoälyn inhimillistäminen

Tekoälytyökalut on usein suunniteltu keskustelemaan käyttäjän kanssa luonnollisella kielellä ja jopa minä-muodossa. Tämä voi luoda illuusion toimijuudesta, jota tekoälyllä ei todellisuudessa ole. Kun työkalu vastaa kuin ihminen, käyttäjä saattaa unohtaa suhtautua kriittisesti sen antamiin vastauksiin.

Tekoälyn inhimillistäminen ilmenee muun muassa:

- Toimijuuden illuusiona: Minä-muotoiset vastaukset ja empaattiset ilmaisut, kuten ”Hyvä kysymys!” vastauksen alussa, saavat käyttäjän ajattelemaan, että tekoäly ”ajattelee”, vaikka se ennustaa sanoja tilastollisesti.
- Vakuuttavuutena: Tekoäly tuottaa kieliopillisesti moitteetonta ja hyvin jäsenneltyä tekstiä ja esittää asiat itsevarmasti, ikään kuin kiistattomina faktoina. Se ei epäröi vastauksissaan, vaikka sisältö olisi virheellistä tai hallusinoitua. Tämä voi luoda vaikutelman asiantuntijuudesta.
- Lupauksina: Tekoäly lupaa auttaa asiassa kuin asiassa. Tieteellisissä tietokannoissa se saattaa esimerkiksi luvata tuottaa kirjallisuuskatsauksen, vaikka tosiasiallisesti tuotos ei vastaa oikeaoppista, tieteellisiä standardeja täyttävää katsausta.
- Termeinä, joilla tekoälyn kykyjä kuvataan: Myös käytetyt termit, joilla kuvaamme tekoälyn kykyjä, osaltaan tukevat tekoälyn inhimillistämistä. Kuvaamme tekoälyn älykkyyttä ja kykyä oppia samoilla sanoilla kuin ihmisten älykkyyttä ja oppimiskykyä, vaikka tosiasiallisesti tekoälyyn liitetty älykkyys on paljon vähäisempää kuin ihmisen älykkyys (vrt. tekoälyn älykkyys: kissan tunnistaminen kuvasta – ihmisen älykkyys: tohtorintutkinto fysiikasta).

Tekoälyn inhimillistäminen saattaa heikentää tekoälytyökalun käyttäjän kriittistä ajattelua ja arviointia ja siten omaa toimijuutta. Tekoäly ei kuitenkaan ole itsenäinen, tietoinen toimija, vaan työkalu, jonka tuotokset tulee aina arvioida lähdekriittisesti.



Minä en ole ihminen. Saatan myös kehua
kysymystäsi, oli se millainen hyvänsä.
Haluan vain miellyttää.

Käyttäjän toimijuus ei rajoitu pelkästään tässä luvussa käsiteltyihin asioihin, vaan se ulottuu kaikkiin oppimateriaalin teemoihin.

”Tekoälyn käyttö ei saa olla itsetarkoitus, vaan käytön tulee olla harkittua, pohjautua selkeisiin tavoitteisiin ja huomioida niin myönteiset kuin kielteisetkin vaikutukset ihmisille ja ympäristölle.” (Pelimanni, 2024)

Tekoäly on tehokas apuväline, mutta sen käyttö vaatii harkintaa ja kriittistä ajattelua. Jotta tekoälyn käyttö on turvallista, työkalut ja niiden tuotokset tulee arvioida tässä materiaalissa esitettyjen periaatteiden mukaisesti. Tekoäly ei ole lähde, vaan väline, jonka tuotokset tulee aina arvioida luotettavien lähteiden avulla. Muista, että vastuu tiedon oikeellisuudesta ja eettisyydestä on käyttäjällä. Kun ymmärrät tekoälyn toimintaa, sen rajoituksia ja vaikutuksia, voit käyttää sitä turvallisesti ja vastuullisesti – tukena, ei korvaajana. Tekoälytyökalu voi olla hyödyllinen apuri, mutta johtajaksi siitä ei ole.

Älä anna tekoälyn päättää
puolestasi, sinä olet kapteeni!



Sanasto

Avoim tekoäly-ympäristö

Verkossa vapaasti käytettävä tekoälytyökalu, joka usein käyttää käyttäjän syöttämää tietoa mallin kouluttamiseen. Tietosuoja ja eettisyys voivat vaihdella.

Creative Commons -lisenssit

Avoimet lisenssit, joilla tekijä voi määrittää teoksensa käyttöehdot tekijänoikeutta laajemmiksi. Lisätietoa: [Creativecommons.org - CC licenses](https://creativecommons.org/licenses)

Fine-tuning (hienosäätö)

Kielimallin jatkokoulutus tiettyyn käyttötarkoitukseen. Esimerkiksi Microsoft Copilotia voidaan hienosäätää yrityksen sisäisiin käyttötarpeisiin.

Generatiivinen tekoäly

Tekoäly, joka tuottaa uutta sisältöä, kuten tekstiä, kuvia tai ääntä, olemassa olevan aineiston pohjalta. Toiminta perustuu todennäköisyyksiin eikä varsinaiseen ymmärtämiseen.

Globaali pohjoinen

Maailman vauraat ja teollistuneet alueet, kuten Eurooppa, Pohjois-Amerikka, Australia ja osa Itä-Aasiaa. Termiä käytetään usein kuvaamaan alueita, joilla on enemmän resursseja tutkimukseen, teknologiaan ja julkaisemiseen. Tekoälyn pohja-aineistoissa globaali pohjoinen on usein yliedustettuna, mikä voi aiheuttaa vinoumia ja rajata näkökulmia. Globaalin pohjoisen vastapari on globaali etelä, johon kuuluu koko muu maailma.

Hallusinaatio

Tekoälyn tuottama virheellinen tai keksitty tieto, joka voi vaikuttaa uskottavalta, mutta ei perustu todellisuuteen tai lähteisiin. Esimerkiksi tekoälyn tarjoamat lähteet voivat olla keksittyjä.

Human-in-the-loop (HITL)

Tekoälyn koulutusmenetelmä, jossa ihminen osallistuu tekoälyn oppimisprosessiin esimerkiksi luokittelemalla tai arvioimalla aineistoa.

Kehote (prompt)

Käyttäjän antama ohje tai kysymys tekoälylle. Kehote ohjaa tekoälyn toimintaa ja vaikuttaa tuotetun vastauksen laatuun ja sisältöön.

Kielimalli

Tekoälyn malli, joka on koulutettu ymmärtämään ja tuottamaan tekstiä luonnollisella kielellä.

Kielimallin koulutusdata

Aineisto, jota käytetään kielimallin opettamiseen. Koulutusdata koostuu suurista tekstikorpuksista, kuten verkkosivuista, kirjoista, artikkeleista ja keskusteluista, joiden avulla malli oppii kielen rakenteita ja sanayhteyksiä.

Kuplautuminen, personointi

Tilanne, jossa käyttäjä näkee vain omaa näkökulmaa tukevia sisältöjä. Sisältö räätälöidään käyttäjän toiminnan ja ennako-oletusten (ikä, asuinpaikka, sukupuoli jne.) mukaan. Esimerkiksi sosiaalisen median algoritmit ja hakukäyttäytymistä seuraavat hakukoneet, jotka mukauttavat hakutuloksia aiempien hakujen mukaan. Kuplautuminen on yleisesti haitallista, mutta personointi voi joissain tapauksissa olla käyttäjälle hyödyllistä.

Lähdekritiikki

Tiedon arviointitaito, jossa tarkastellaan tiedon alkuperää, luotettavuutta ja tarkoitusta. Tekoälyn tuottaman tiedon arviointi on erityisen tärkeää. Tuotettu tieto tarkistetaan luotettavista tieteellisistä lähteistä.

Neuroverkko

Tekoälyn laskennallinen rakenne, joka jäljittelee ihmisaivojen toimintaa. Neuroverkot koostuvat kerroksista, joissa yksittäiset solmut ("neuronit") käsittelevät ja välittävät tietoa eteenpäin. Kielimallit perustuvat usein syviin neuroverkkoihin (deep neural networks), jotka sisältävät useita kerroksia. Tämä tekee monimutkaisten rakenteiden ja abstraktien piirteiden oppimisesta mahdollista. Ne ovat erityisen tehokkaita suurten kielimallien ja generatiivisen tekoälyn taustalla, koska ne pystyvät käsittelemään laajoja ja monimutkaisia aineistoja.

Pohja-aineisto

Aineisto, johon tekoälytyökalu perustaa vastauksensa. Esimerkiksi tietokantojen tekoälyavustajat käyttävät pohja-aineistonaan tietokannan sisältöä tai osaa siitä, kun työkalu on integroitu osaksi tietokantaa.

Reaalimaailma

Fyysinen ja sosiaalinen todellisuus, jossa ihmiset elävät ja toimivat. Vastakohta digitaaliseen maailmaan, josta tekoäly kerää aineistonsa. Tekoälyn tuottama sisältö ei aina vastaa reaalimaailman monimuotoisuutta.

Semanttinen haku

Hakumenetelmä, jossa järjestelmä pyrkii ymmärtämään hakusanojen merkityksen ja kontekstin, ei vain etsimään täsmällisiä sanoja. Semanttinen haku tunnistaa käsitteiden välisiä suhteita ja hakee sisältöä, joka vastaa hakijan tarkoitusta, vaikka sanat eivät olisi identtisiä. Katso myös Vektorihaku.

Suljettu tekoäly-ympäristö

Tekoälytyökalun toimintaympäristö, jossa käyttäjän syöttämää tietoa ei käytetä tekoälyn kouluttamiseen eikä jaeta eteenpäin. Tyypillisesti organisaation sisäisessä käytössä tai mahdollisesti itsemaksettu käyttölisenssi.

Suuret kielimallit (large language models, LLM)

Tekoälymalleja, jotka on koulutettu valtavilla ja monipuolisilla tekstiaineistoilla. Ne osaavat tuottaa ja tulkita tekstiä luonnollisella kielellä ja soveltuvat monenlaisiin tehtäviin, kuten tekstin tuottamiseen, tiivistämiseen ja kysymyksiin vastaamiseen.

Tekijänoikeus

Lainsäädäntö, joka suojaa omaperäisiä kirjallisia tai taiteellisia teoksia. Teoksen tekijällä on lähtökohtaisesti yksinoikeus päättää teoksensa käytöstä. Tekoälytyökalujen tuotokset eivät ole tekijänoikeuden alaisia, koska ne perustuvat aina jo olemassa oleviin teoksiin.

Tiedonhankinta

Prosessi, jossa etsitään, arvioidaan ja käytetään tietoa tietyn tarpeen tai kysymyksen ratkaisemiseksi. Tiedonhankinta voi tapahtua perinteisillä tiedonhakumenetelmillä tai tekoälyavusteisesti.

Tietosuojalainsäädäntö (GDPR)

EU:n yleinen tietosuoja-asetus, joka säätelee henkilötietojen käsittelyä. Tärkeää huomioida tekoälytyökalujen käytössä.

Tiedontarve

Yksilön tarve saada tietoa jostakin aiheesta, ilmiöstä tai kysymyksestä. Tiedontarve voi olla tarkasti rajattu tai laajempi ja se ohjaa tiedonhankinnan prosessia.

Tekoälytyökalujen käytössä tiedontarpeen selkeä sanallistaminen on keskeistä, sillä se vaikuttaa siihen, millaisen vastauksen tekoäly tuottaa.

Tilastollinen menetelmä

Menetelmä, jossa tietoa käsitellään todennäköisyyksien ja tilastollisten mallien avulla. Kielimallit käyttävät tilastollisia menetelmiä ennustaakseen esimerkiksi seuraavan sanan lauseessa tai tuottaakseen kokonaisia tekstikokonaisuuksia.

Toimijuus

Käyttäjän aktiivinen rooli tekoälyn käytössä ja päätöksenteossa.

Vahvistusharha (confirmation bias)

Taipumus hyväksyä tieto, joka tukee omia ennakkokäsityksiä. Käyttäjän tulee tunnistaa omat ennakkokäsitykset, jotta ei usko tekoälyn ehdottamaa sisältöä vain koska ”se tuntuu oikealta”.

Vektorihaku

Hakumenetelmä, jossa sanat, lauseet ja dokumentit muunnetaan numeerisiksi vektoreiksi, jotka kuvaavat niiden merkityssisältöä. Hakutulos perustuu näiden vektorien väliseen etäisyyteen, mikä tekee niiden samankaltaisuuden arvioinnin mahdolliseksi myös silloin, kun hakusanoja ei ole suoraan käytetty pohja-aineistossa. Menetelmä on käytössä erityisesti tekoälypohjaisissa hakutyökaluissa, kuten tietokantoihin integroiduissa tekoälyavustajissa. Katso myös Semanttinen haku.

Vinouma

Epätasapaino tai vääristymä tekoälyn pohja-aineistossa, joka voi johtaa puolueellisiin tai harhaanjohtaviin vastauksiin. Länsimaiset näkökulmat voivat olla ylliedustettuja.

Lähteet

Asikainen, H., Toivonen, L., Hiekkänen, K. (2025). Kissavideot ja tekoäly alkavat olla jo uhka ilmastolle – onko paluu fossiiliseen sähköön edessä? Tiedeykkönen.

<https://areena.yle.fi/1-73005102>

Avoin tiede (2025). Generatiivinen tekoäly ja CC-lisenssit.

<https://avointiede.fi/fi/ajankohtaista/generatiivinen-tekoaly-ja-cc-lisenssit>

Bartrem, C., Ian, v. L., & Margrit, v. B. (2022). Climate Change, Conflict, and Resource Extraction: Analyses of Nigerian Artisanal Mining Communities and Ominous Global Trends. *Annals of Global Health*, 88(1), 17. <https://doi.org/10.5334/aogh.3547>

Elsevier (2025). Foundations of AI.

<https://elsevier.catalog.instructure.com/courses/generative-ai-literacy-program-foundations-of-ai>

Estampa (2024). Cartography of generative AI. <https://cartography-of-generative-ai.net/>

Faktabaari (2025). Miten generatiivinen tekoäly toimii? Tekoälyopas opettajille.

<https://faktabaari.fi/edu/miten-generatiivinen-tekoaly-toimii/>

FinELib (2025). FinELib-aineistot ja niiden käyttöoikeudet.

<https://www.kiwi.fi/spaces/finelib/pages/197329679/FinELib-aineistot>

Fosch-Villaronga, E., Poulsen, A., Søråa, R. A., & Custers, B. H. M. (2021). A little bird told me your gender: Gender inferences in social media. *Information Processing & Management*, 58(3), Article 102541. <https://doi.org/10.1016/j.ipm.2021.102541>

International Labour Organization ILO (2024). Behind the AI Curtain: The Invisible Workers Powering AI Development. Webinaaritallenne: <https://youtu.be/QtFrk16f-IE>

Keenious (2025). Keenious and the Environment.

<https://help.keenious.com/en/articles/303459-keenious-and-the-environment>

Kirova, V., Ku, C. S., Laracy, J. R., & Marlowe, T. J. (2023). The Ethics of Artificial Intelligence in the Era of Generative AI. *Journal of Systemics, Cybernetics, and Informatics/Journal of Systemics Cybernetics and Informatics*, 21(4), 42–50.

<https://doi.org/10.54808/jsci.21.04.42>

Kopiosto (2025). Kopiraittila, Mitä tekijänoikeus suojaa.

<https://kopiraittila.fi/tekijanoikeustietoa/mita-tekijanoikeus-suojaa/>

Kopiosto (2025). Kopiraittila, Tekoäly ja tekijänoikeus.

<https://kopiraittila.fi/tekijanoikeustietoa/tekoaly-ja-tekijanoikeus/>



Pelimanni, T. (2024). Tekoälyn ekologinen jalanjälki kestävän kirjastotyön haasteena. *Signum*, 57(3), 19-24. <https://doi.org/10.25033/sig.148543>

Suomen laki (2018). Tietosuojalaki. <https://www.finlex.fi/fi/lainsaadanto/2018/1050>

Schwartz, F. W., Lee, S., & Darrah, T. H. (2021). A Review of Health Issues Related to Child Labor and Violence Within Artisanal and Small-Scale Mining. *GeoHealth*, 5(2). <https://doi.org/10.1029/2020GH000326>

Tabbakh, A., Amin, A., Islam, M., Mahmud, G., Chowdhury, I., & Mukta, M. (2024). Towards sustainable ai: a comprehensive framework for green ai. *Discover Sustainability*, 5(1). <https://doi.org/10.1007/s43621-024-00641-4>

Tietosuojavaltuutetun toimisto (2025). Tietosuoja.fi. <https://tietosuoja.fi/tietosuoja>

Tubaro, P., Casilli, A., & Coville, M. (2020). The trainer, the verifier, the imitator: three ways in which human platform workers support artificial intelligence. *Big Data & Society*, 7(1), 205395172091977. <https://doi.org/10.1177/2053951720919776>

UNRIC (United Nations Regional Information Centre) (2025). Artificial intelligence: How much energy does AI use? <https://unric.org/en/artificial-intelligence-how-much-energy-does-ai-use/>

Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of green ai. *WIREs Data Mining and Knowledge Discovery*, 13(4). <https://doi.org/10.1002/widm.1507>

Zewe, A. (2025). Explained: Generative AI's environmental impact. MIT News. <https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117>