

Tekoälyn vaikutusten eettinen arviointi

Ennakointiakatemia 6.11.2024

Taina Kalliokoski

sosiaalietiikan tutkijatohtori

ETAIROS-hanke, Teologinen tiedekunta, Helsingin yliopisto

”Ihmiset ovat aina olleet huonoja arvioimaan teknologisten keksintöjen vaikutuksia luonnonympäristöön, ihmiselämään ja yhteiskuntaan. Itsesäätely ei ole onnistunut ehkäisemään aikaamme vaivaavaa ekologista kriisiä.”

-Akateemikko Ilkka Niiniluoto, Mielipide HS 24.10.2024 ”Ennakkosäätelyn heikkous pönkittää teknojättien valtaa”

Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic

15 MINUTE READ



This image was generated by OpenAI's image-generation software, Dall-E 2. The prompt was: "A seemingly endless view of African workers at desks in front of computer screens in a printmaking style." TIME does not typically use AI-generated art to illustrate its stories, but chose to in this instance in order to draw attention to the power of OpenAI's technology and shed light on the labor that makes it possible. Image generated by Dall-E 2/OpenAI

6.11.2024

Tekoälyn materiaaliset ja moraaliset edellytykset

- Laadukas data **edellyttää datatyötä**
- jota tehdään usein vähemmän kuin ideaaleissa työoloissa (Perrigo 2023).
- Datan luokittelun perusteet voivat olla vinoutuneita (esim. Berg & Valaskivi 2023)
- Kuka omistaa koulutusdatan?
- Tekoälyavusteisia työkaluja ei voi käyttää suostumatta datan lähteeksi.

taina.kalliokoski@helsinki.fi



a Unicom, which uses artificial intelligence in its operations. TAO LIANG / XINHUA VIA GETTY IMAGES

As Use of A.I. Soars, So Does the Energy and Water It Requires

Generative artificial intelligence uses massive amounts of energy for computation and data storage and millions of gallons of water to cool the equipment at data centers. Now, legislators and regulators – in the U.S. and the EU – are starting to demand accountability.

BY DAVID BERREBY • FEBRUARY 6, 2024

Tekoälyn materiaaliset edellytykset

- **Laskentakapasiteetti**
 - raaka-aineet konfliktialueilta
 - makeaa vettä jäähdyttämiseen
- **Energiantarve kaksinkertaistuu** tällä hetkellä **joka 100. päivä** (Stollenwerk 2024),
 - syynä tekoälyn lisääntynyt käyttö ja käyttötavat (Jones 2024)
- **Onko ekologisille kustannuksille hyväksyttävät perusteet?** Jakautuvatko loppukäyttäjälle helpon tekoälyn kehittämiskustannukset reilusti?

Tekoälyn käytön vaikutukset organisaatioon ja yhteiskuntaan

- **Hyödyt**
 - Millaisia hyödyntekeviä vaikutuksia tekoälyllä halutaan saada aikaan organisaatiossa ja yhteiskunnassa laajemmin?
 - Realistisuus
 - Oikeutus
 - Taloudelliset, sosiaaliset, terveydelliset, moraaliset kustannukset?
- **Haitat tai riskit**
 - Mitä saadaan, mitä menetetään? Millaisia trade-offeja hyväksytään?
 - terveys – yksityisyys,
 - tehokkuus – turvallisuus,
 - helppous – kyvykkyys

Kuva: Kelsey He, Unsplash.



Esim. Generatiivisen tekoälyn käytön vaikutukset

	LYHYT AIKAVÄLI	PITKÄ AIKAVÄLI
AIOTTU	Kirjoittaminen nopeutuu; työ helpottuu; Hyvän tekstin/kuvien tuottaminen eri tarkoituksiin on mahdollista yhä useammalle	Taloudellinen kannattavuus paranee
EI-AIOTTU	Työntekijöiden kokemus oman työn arvosta ja ylpeys aikaansaannoksista vähenee; Tekstit muuttuvat selkeämmiksi ja siten luettavimmiksi	Uusien ammattilaisten ja asiantuntijoiden kyky ponnistella monimutkaisten aiheiden parissa heikkenee; tekstit, kuvat ja esitykset yhdenmukaistuvat
ENNAKOITU	Työvoiman tarve vähenee	Ihmisten vapaa-aika lisääntyy, taloudellinen niukkuus lisääntyy, kun vanhat tulonjakomallit eivät toimi
ENNAKOIMATON 6.11.2024	<u>Valeuutiset ja huijaukset lisääntyvät,</u> energiankulutus kasvaa	Koulutuksen arvostus heikkenee ("tekoäly hoitaa" "tekoälytutkinto"); Tiedonvälityksen luotettavuus rapautuu; Eriarvoisuus lisääntyy

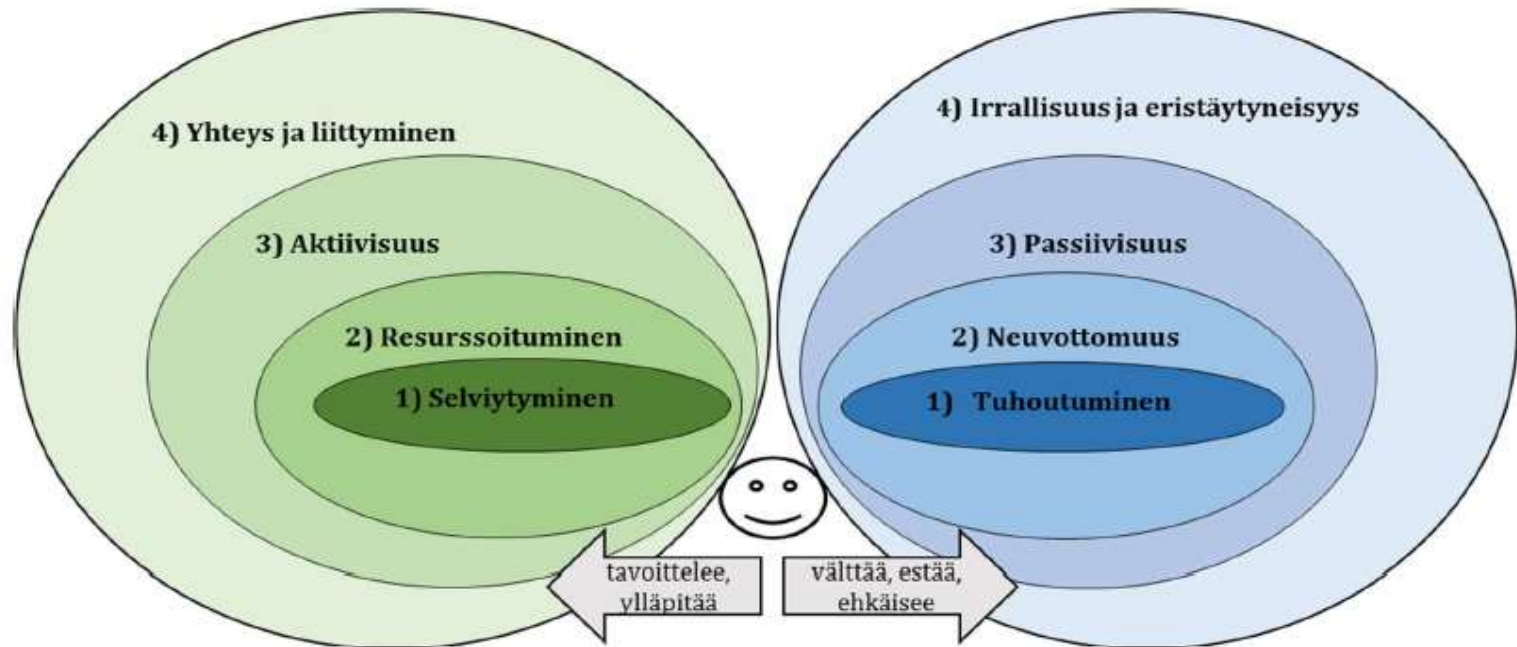
6.11.2024

MITEN ENNAKOIDA TEKOÄLYN VAIKUTUKSIA IHMISELÄMÄÄN?

taina.kalliokoski@helsinki.fi

7

Hyvä elämä edellyttää toimijuutta, perustavien hyvien edistämistä ja ei-hyvien estämistä. (Kalliokoski 2020)



IHMISEN PERUSHYVÄT JA PERUSTAVAT EI-HYVÄT

(SMITH 2015)

(KALLIOKOSKI 2020)

Selviytyminen



I fyysinen
selviytyminen,
terveys, turvallisuus,
mielihyvä

tuhoutuminen,
sairaus,
vammautuminen,
kuolema;
turvattomuus;
mielihyvän puute,
kärsimys

Tuhoutuminen



Kuvat: Feng Shan, Jilbert Ebrahimi, Unsplash

Resursoituminen



II Todellisuutta
koskeva tieto



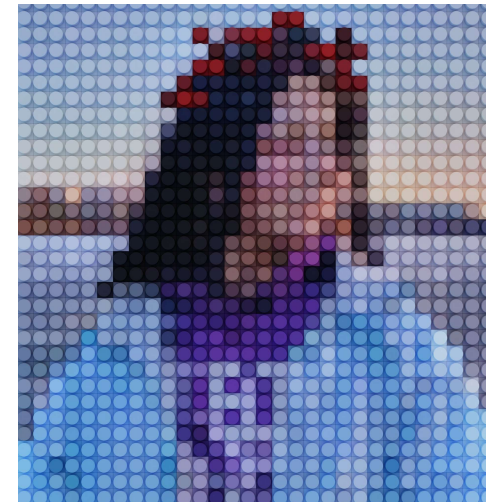
III Eheä, tunnustusta
saanut identiteetti

Neuvottomuus

Ei tietoa
todellisuudesta;
Epätodet käsitykset



Epäyhtenäinen,
tunnustamaton
identiteetti



Kuvat: Kiana Bosman; Sep; Jorve Saavedra; Vadim Bogulov, Unsplash

6.11.2024

taina.kalliokoski@helsinki.fi¹⁰

Aktiivisuus



IV Tarkoituksellisen
toimijuuden
harjoittaminen

Passiivisuus



Toimijuuden
hajoaminen,
toimimattomuus

Kuvat:, Arunas Naujokas, Ulrich Derboven, Unsplash

6.11.2024

taina.kalljokoski@helsinki.fi

ETAIROS

Yhteys ja liittyminen



**V Moraalinen
vahvistuminen**



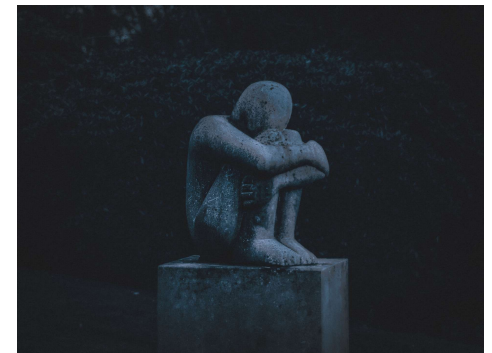
**VI Sosiaalinen
kuuluminen,
rakkaus**

Irrallisuus ja eristäytyminen

**Moraalinen
tietämättömyys,
vääristymät
vahvistumattomuus,
epämoraalinen elämä**



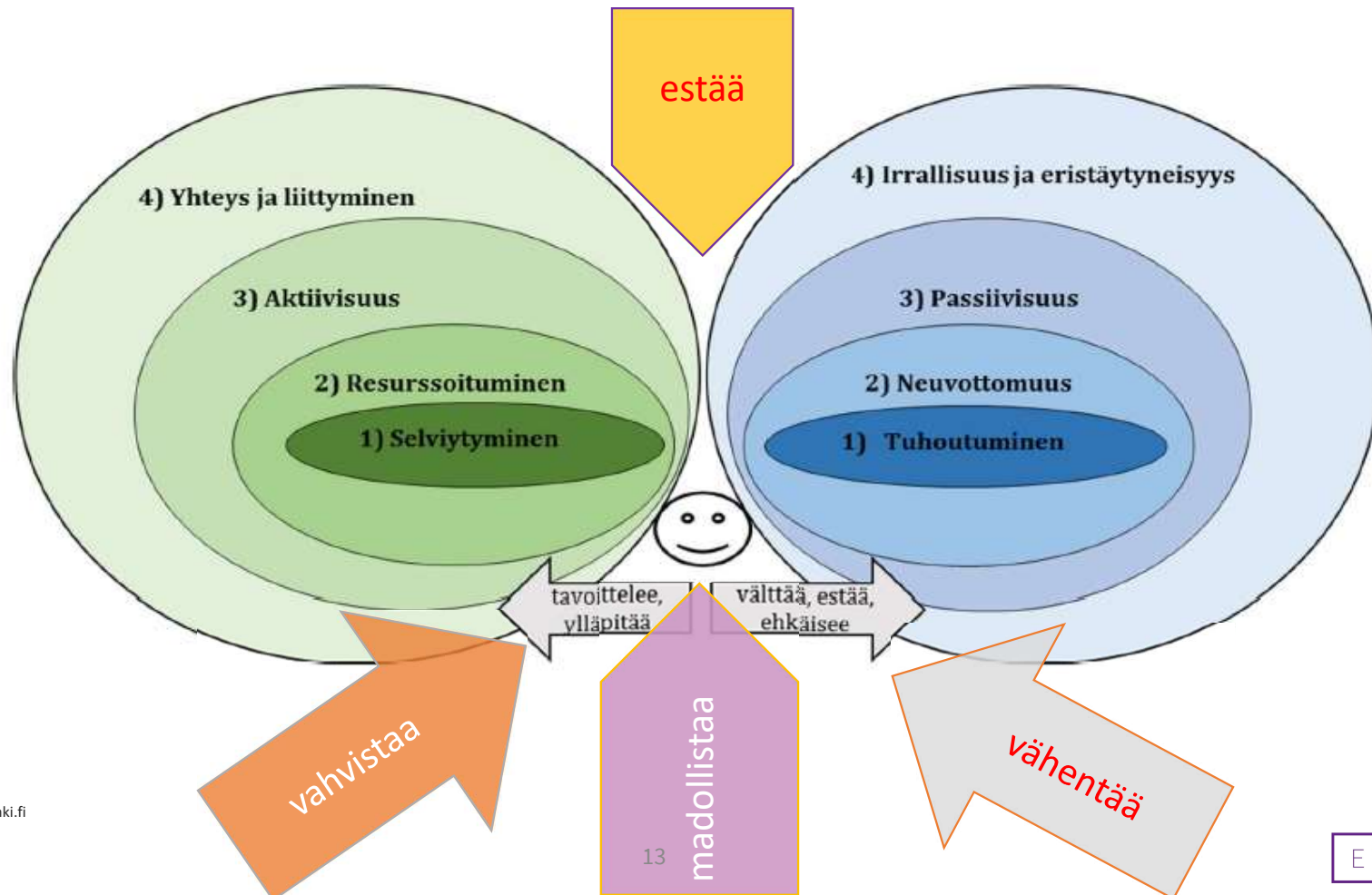
**Sosiaalinen
vieraantuminen,
yksinäisyys,
rakkaudellisten
suhteiden puute, viha
välinpitämättömyys,
pelko**



12

Kuvat: Li An Lim, Callum Skelton, Artem Kniaz and K Mitch Hodge, Unsplash

Miten tekoälysovellukset vaikuttavat perushyvien tavoitteluun?



Tekoälylukutaidon, teknisen tekoälyosaamisen kouluttamisen ohella on harjoitettava



eettistä harkintakykyä ja
moraalista mielikuvitusta.

Kuva: Kyle Glenn, Unsplash.

Lähteet

Berg, Anton & Valaskivi, Katja (2023) . Berg , A & Valaskivi , K 2023 , ' Dataistunut uskonto : Kaupalliset kuvantunnistuspalvelut uskontoa ”tunnistamassa” ', Teologinen Aikakauskirja 128 :2.

ETAIROS-hanke. www.etairos.fi

Hallamaa, Jaana & Taina Kalliokoski (2023). Vastuunkantaminen sosioteknisen systeemin epäonnistumisessa. Teologinen Aikakauskirja 128:2.

Jones, Florence (2024). AI's energy-hungry data centres bring sustainability challenges 4.4.2024. <https://www.sustainableviews.com/ais-energy-hungry-data-centres-bring-sustainability-challenges-f8c3bdae/>

Kalliokoski, Taina (2020). Yhteisöllisyyden rajat yhteistoiminnan ja ihmisen perushyvien näkökulmasta. Diss. Helsingin yliopisto. <http://hdl.handle.net/10138/314726>

Perrigo, Billy 18.1.2023. Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. <https://time.com/6247678/openai-chatgpt-kenya-workers/>

Smith, Christian (2015). To Flourish or Destruct? A Personalist Theory of Human Goods, Motivations, Failure, and Evil. Chicago: Chicago University Press.

Stollenwerk, Felix (2024). Open Source LLMs for Sweden, the Nordic countries and Europe. Keynote luento AI Day 22.10.2024.

Unesco, Ethics of Artificial Intelligence <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics>