

Siina Raskulla

Lakimies  
7–8/2023  
s. 1232–1245

# Ihmisen ja tekoälyn vuorovaikutuksen metalait

## 1. Johdanto

Keväällä 2023 julkisuuteen nousi laajasti mediahuomiota saanut avoin kirje. Kirjeessä kehoitettiin kaikkia tekoälylaboratorioita pysäyttämään tehokkaiden tekoälyjärjestelmien kehittäminen kuudeksi kuukaudeksi sekä työskentelemään poliittisten päättäjien kanssa kestävien tekoälyn hallintajärjestelmien luomiseksi. Kirjeen taustalla oli huoli tekoälykehityksen aiheuttamasta uhasta, joka koski ”sivilisaatiomme hallinnan menettämistä”, ja sen allekirjoitti muun muassa *Elon Musk*.<sup>1</sup> Muskin kilpailija, *Sam Altman*, ja muun muassa *Bill Gates* taas allekirjoittivat oman julkilausumansa. Tässä julkilausumassa todettiin, että tekoäly muodostaa eksistentiaalisen uhan, joka on rinnastettavissa pandemiaan ja ydinuhkaan.<sup>2</sup>

Erityisesti nimekkäiden alalla toimivien allekirjoittajiensa ansiosta avoin kirje tekoälykehityksen väliaikaiseksi pysäyttämiseksi sai huomattavaa julkisuutta. Tekoälyn kehittymisen vauhdista ja vaikutuksista on kuitenkin käyty ja käydään jatkuvaa keskustelua niin konkreettisen sääntelytyön, kuten EU:n tekoälyasetuksen, kuin spekulatiivisempien kysymysten tarkastelun muodossa.

Tämä artikkeli tarkastelee spekulatiivisempaa kysymystä ihmisen ja kehittyneen tekoälyn vuorovaikutuksesta. Miten voimme varmistaa, että tekoälyn kehittyessä tietoiseksi kykenemme rauhanomaiseen rinnakkaiseloon? Miten varmistamme, ettei tietoinen superälykyys tuhoa meitä tieltään?<sup>3</sup> Artikkelissa otetaan vaikutteita ihmisen ja maan ulkopuolisen älyn keskinäistä vuorovaikutusta sääteleviä metalakeja koskevasta keskustelusta.

\* *Siina Raskulla*, HT, yliopistonlehtori, Tampereen yliopisto.

1. Future of Life Institute, *Pause Giant AI Experiments: An Open Letter*. 22.3.2023 osoitteessa <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (vierailtu 30.8.2023). Tarkemmin politiikkasuosituksista ks. Future of Life Institute, *Policymaking in the Pause. What can policymakers do now to combat risks from advanced AI systems?* 12.4.2023 osoitteessa [https://futureoflife.org/wp-content/uploads/2023/04/FLI\\_Policymaking\\_In\\_The\\_Pause.pdf](https://futureoflife.org/wp-content/uploads/2023/04/FLI_Policymaking_In_The_Pause.pdf) (vierailtu 30.8.2023).
2. Center for AI Safety, *Statement on AI Risk – AI experts and public figures express their concern about AI risk* osoitteessa <https://www.safe.ai/statement-on-ai-risk> (vierailtu 1.9.2023).
3. Myös ei-tietoinen tekoäly voi tuhota ihmiskunnan esimerkiksi puutteellisen ohjelmoinnin tai virheen takia, mutta tässä kirjoituksessa ovat lähtökohtana säännöt käyttäytymistä ohjaavana mekanismina. Ks. jäljempänä luku 3.

Metalait kehitettiin vähintään yhtä spekulatiivista tilannetta varten: ihmisen ja maan ulkopuolisen älyn kohtaamiseen. Artikkelissa tutustutaan ensin *Andrew G. Haleyn* ideoimiin ja *Ernst Fasanin* tarkemmin muotoilemiin metalakeihin. Tämän jälkeen perehdytään siihen, miten *Kamil Muzykan* on soveltanut Fasanin muotoilemia metalakeja yleistekoölyyn. Seuraavaksi tarkastellaan Haleyn ja Fasanin metalakien saamaa kritiikkiä ja pohditaan kritiikin soveltuvuutta ihmisen ja yleistekoölyn vuorovaikutusta sääteleviin metalakeihin. Lopuksi keskustellaan metalakien tarpeellisuudesta suhteissamme yleistekoölyyn tai tekoölyyn yleisesti.

Keskeisenä käsitteenä artikkelissa on yleistekoöly (general artificial intelligence, AIG), josta voitaisiin käyttää myös termiä vahva (strong AI), keinotekoinen tietoisuus (artificial consciousness) tai esimerkiksi spontaani äly (spontaneous intelligence).<sup>4</sup> Yleistekoölyllä tarkoitetaan tekoölyjärjestelmää, joka on tietoinen tai kykenee tietoisuuteen.<sup>5</sup> Artikkelin ulkopuolelle rajataan siis lähtökohtaisesti ”heikot” tekoölyjärjestelmät, joilla ei ole tietoisuutta, vaikka ulkoapäin tarkasteltuna se saattaisikin vaikuttaa siltä.<sup>6</sup>

## 2. Mitä ovat metalait?

Niin *Isaac Asimovin* robotiikan perussäännöt, *Asilomarin* periaatteet tekoöllylle sekä Euroopan unionin tekoölyasetus ovat luonteeltaan antroposentrisiä eli ihmiskeskeisiä. Antroposentriset lait asettavat velvollisuuksia ja rajoituksia tekoöllylle, sen kehittäjille ja käyttäjille, joiden tavoitteena on yksilöiden, ryhmien ja ihmiskunnan oikeuksien ja hyvinvoinnin turvaaminen.<sup>7</sup>

Metalakien lähtökohtana on sen sijaan kohtaaminen tasavertaisten olentojen kesken. Ne hahmoteltiin alun perin ajatellen kohtaamista maan ulkopuolisen

4. Ks. Nadia Banteka, *Artificially Intelligent Persons*. *Houston Law Review* 58(3) 2021, s. 537–596, 542–548; Jiahong Chen – Paul Burgess, *The boundaries of legal personhood: how spontaneous intelligence can problematise differences between humans, artificial intelligence, companies and animals*. *Artificial Intelligence and Law* 27(1) 2019, s. 73–92, 75 ja John R. Searle, *Minds, brains, and programs*. *Behavioral and Brain Sciences* 3(3) 1980, s. 417–457, 417.

5. Kamil Muzyka, *The basic rules for coexistence: The possible applicability of metalaw for human-AGI relations*. Paladyn, *Journal of Behavioral Robotics* 11(1) 2020, s. 104–117, 107.

6. Heikkojen tekoölyjärjestelmien oikeushenkilöllisyydestä ks. esim. Siina Raskulla, *Hybrid Theory of Corporate Legal Personhood and Its Application to Artificial Intelligence*. *SN Social Sciences* 3(78) 2023 [sähköisenä] osoitteessa <https://doi.org/10.1007/s43545-023-00667-x>. Kysymykseen, miten kykenemme erottamaan vahvan ja heikon tekoälyn, jos niiden keskeisenä erona on sisäisen kokemus tietoisuudesta, palataan luvussa 4.

7. Ks. Andrew G. Haley, *Space Law and Metalaw – Jurisdiction Defined*. *Journal of Air Law and Commerce* 23(3) 1975, s. 286–303, 286 ja Muzyka 2020, s. 104.

elämän (extraterrestrial intelligence, ETI) kanssa.<sup>8</sup> Vieraan elämänmuodon kohtaaminen on ajatuksena kiinnostanut ihmisiä yli 2000 vuoden ajan. Ajatuksen parissa ovat työskennelleet monet oikeustieteeseenkin aikanaan vaikutteensa jättäneet generalistit, kuten *Tuomas Akvinolainen*, *Montesquieu* ja *Immanuel Kant*.<sup>9</sup> Ajatus ihmisen ja maan ulkopuolisen elämän vuorovaikutuksen säännöistä on herättänyt keskustelua planeettojen- tai tähtienvälisestä oikeudesta tai astraalilaeista.<sup>10</sup> Nykypäivään keskustelusta on – kirjallisuuden perusteella – säilynyt parhaiten Andrew G. Haleyn vuonna 1956 julkaistussa teoksessaan *Space Law and Metalaw, A Synoptic View* esittelemä käsite metalait.<sup>11</sup>

Haley otti metalakien perustaksi ”tähtienvälisen kultaisen säännön”, eli ”tee toisille niin kuin haluaisit heidän tekevän sinulle”.<sup>12</sup> Tästä hän johti kourallisen perussääntöjä, kuten voimankäytön kiellon ja tasa-arvon periaatteen.<sup>13</sup> Ernst Fasan jalosti Haleyn ajatuksia muotoillen 11 metalakia, jotka määrittävät ihmisen ja maan ulkopuolisen elämän välisen suhteen peruseriaatteen (katso jäljempänä taulukko 1).<sup>14</sup> Myöhemmin Fasan tiivistää nämä metalait kolmeen universaalisti

8. Oikeusjärjestelmien antroposentrisyyttä haastaa nykyään myös oikeuksien antaminen niin elolliselle kuin elottomalle luonnolle, kuten eläimille, puille ja joille. Ks. esim. Jana Norman, *Posthuman Legal Subjectivity: Reimagining the Human in the Anthropocene*. Routledge 2022.
9. Ks. Ernst Fasan, *Relations with Alien Intelligences – The Scientific Basis of Metalaw*, s. 181–246 teoksessa Patricia Margaret Sterns – Lesli I. Tennen (eds), *Private Law, Public Law, Metalaw and Public Policy in Space: A Liber Amicorum in Honor of Ernst Fasan*. Aspen Publishers 2017 (julkaistu aiemmin: Berlin Verlag Arno Spitz 1970), 189–191. Etiikasta ja metafysiikasta tunnettu Kant tutki myös aurinkokunnan muodostumista ja planeettojen kiertoratoja sekä johti luonnon tarkoituksesta tulla havaituksi sen, että myös muilla taivaankappaleilla olisi luonnollisesti oltava elämää ainakin jossain vaiheessa. Ks. Immanuel Kant, *Natural Science*. Cambridge University Press 2012, s. 297. Ks. myös Andreas Losch, *Kant’s Wager. Kant’s Strong Belief in Extra-Terrestrial Life, the History of This Question and Its Challenge for Theology Today*. *International Journal of Astrobiology* 15(4) 2016, s. 261–270.
10. Yhteenvetoa aiemmasta keskustelusta esim. Fasan 2017 [1970], s. 214 ja Muzyka 2020, s. 109–110.
11. Haleya pidetään monesti ensimmäisenä avaruusjuristina (ks. Steven Freeland, *A Natural System of Law? Andrew Haley and the International Legal Regulation of Outer Space*. *Journal of Space Law* 39(1) 2013, s. 77–98). Tämä on voinut myötävaikuttaa siihen, miksi juuri hänen käyttämänsä käsite on saanut eniten huomiota. Metalakien käsite siten kuin sitä tässä käytetään, tulee erottaa metasääntelystä (eng. meta-regulation), joka liittyy kansainväliseen oikeuteen ja erityisesti markkinoiden itsesääntelyyn. Ks. esim. Robert Baldwin – Martin Cave – Martin Lodge, *Understanding Regulation: Theory, Strategy, and Practice*. 2nd ed. Oxford University Press 2012, luku 8 ja Henry E. Smith, *Equity as Meta-Law*. *The Yale Law Journal* 130(5) 2021, s. 1050–1287.
12. Variaatioista kultaisesta säännöstä eri kulttuureissa ja aikakausina, niin uskonnollisia kuin sekularistisia, ks. George S. Robinson, *The Biochemical Foundations of Evolving Metalaw: Moving at a Glance to the Biological Basis of Sentient “Essence”*. *Journal of Space Law* 39(1) 2013, s. 181–216, 195–196.
13. Ks. Fasan 2017 [1970], s. 213 ja siinä mainitut lähteet.
14. Tarkemmin Fasan 2017 [1970], s. 227–236. Fasan luo katsauksen 1960-luvulla käytyyn metalakeja koskevaan keskusteluun, josta hän on saanut vaikutteita omien metalakiensa muodostamiseen. Ks. Fasan 2017 [1970], s. 216–222 ja siinä mainitut lähteet. Myös pyrkimyksille ottaa

hyväksytyyn oikeussääntöön, jotka ovat (1) toiselle lajille aiheutetun vahingon välttäminen, (2) lajien tahdonautonomia ja (3) oikeus omaan elintilaan.<sup>15</sup> Metalakien tavoitteena on rauhanomainen rinnakkaiselo ja yhteistyö.<sup>16</sup> Maan ulkopuolisen elämän etsintää harjoittavan SETI-instituutin laatima protokolla maan ulkopuolisen älyn kohtaamisesta heijastaa Haleyn ja Fasanin ajatuksia.<sup>17</sup>

Ernst Fasanin metalait luovat sekä velvollisuuksia että odotuksia, eli niiden perustana on vastavuoroisuus.<sup>18</sup> Asimovin lait eivät siten tämän määritelmän mukaan ole lakeja vaan yksisuuntaisia käskyjä, joista tässä asiayhteydessä voitaisiin puhua myös komentokehotteina eli prompteina. Asimovin perussääntöjen heikkoutena voidaankin pitää sitä, etteivät ne kohtele tekoälyjä vapaina olentoina.<sup>19</sup> Tekoäly ei voi odottaa mitään vastineeksi käskyjen noudattamisesta toisin kuin lakeja noudattavat ihmiset, jotka voivat odottaa lain myös suojelevan heitä muiden tekemiä rikkomuksia vastaan.

Fasanin mukaan metalakien osapuolilta edellytetään kuitenkin eräitä välttämättömiä ominaisuuksia, joita ovat (1) elämä, (2) älykkyys, (3) havaittavissa oleminen, (4) kolmiulotteisuus tai kyky toimia kolmiulotteisessa avaruudessa sekä (5) tahto elää.<sup>20</sup> Tahto elää sisältää halun suojella itseään.<sup>21</sup> Älykkyys taas tarkoittaa Fasanin mukaan kykyä loogiseen ajatteluun, vapaan tahdon omaamista sekä kykyä oppia.<sup>22</sup> Fasan ei kuitenkaan vastaa kysymykseen siitä, miten voimme tunnistaa älykkään olennon – oli kyseessä sitten maan ulkopuolinen älykkyys tai ihmisen kehittämä tekoäly. Tähän kriittiseen kysymykseen palataankin myöhemmin.<sup>23</sup> Seuraavassa luvussa tarkastellaan ensin, voidaanko metalakeja soveltaa tekoälyyn, ja jos voidaan, eroavatko ne jotenkin ihmisten ja maan ulkopuolisen älyn väliseen kanssakäymiseen sovellettavista laeista.

yhteyttä maapallon ulkopuoliseen elämään on hahmoteltu protokolla, joka heijastaa metaoikeudellisia periaatteita, ks. SETI-institute, *Protocols for an ETI Signal Detection*. 23.4.2018 osoitteessa <https://www.seti.org/protocols-eti-signal-detection> (vierailtu 6.8.2023). Keskustelua protokollasta ks. Patricia Margaret Sterns – Lesli I. Tennen, SETI, Metalaw, and Social Media, s. 159–179 teoksessa Patricia Margaret Sterns – Lesli I. Tennen (eds), *Private Law, Public Law, Metalaw and Public Policy in Space: A Liber Amicorum in Honor of Ernst Fasan*. Aspen Publishers 2017 ja John Gertz, *Post-Detection Seti Protocols & Meti: The Time Has Come to Regulate Them Both*. *Journal of the British Interplanetary Society* 69(8) 2016, s. 263–270.

15. Ernts Fasan, *Legal Consequences of a SETI Detection*. *Acta Astronautica* 42(10–12) 1998, s. 677–679, 678–679.

16. Muzyka 2020, s. 110.

17. SETI-institute 2018. Tarkemmasta vertailusta ks. Gertz 2016, s. 266–267.

18. Fasan 2017 [1970], s. 213.

19. James Lovelock – Bryan Appleyard, *Novacene: The Coming Age of Hyperintelligence*. Allen Lane 2019, s. 94.

20. Fasan 2017 [1970], s. 187, 205–211.

21. Fasan 2017 [1970], s. 187, 226 ja Fasan 1998, s. 678.

22. Fasan 2017 [1970], s. 206 ja Fasan 1998, s. 678.

23. Ks. luku 4.

### 3. Metalait ja tekoöly

Pohtiessaan robottimuotoista maan ulkopuolista elämää Fasan sivuaa lyhyesti ajatusta, mutta jättää käsittelemättä kysymyksen ihmisten synnyttämästä robotti-elämästä.<sup>24</sup> Kamil Muzyka tarttui kysymykseen ja tarkastelee artikkelissaan kysymystä metalakien soveltuvuudesta ihmisten kehittämään yleistekoölyyn.

Kysymyksenasetteluun metalakien tarpeellisuudesta ja soveltuvuudesta on sisäänrakennettuna oletus, että metalakien osapuolena oleva tekoöly täyttää Fasanin kriteerit. Toisin sanoen tekoöly on älykäs, elävä ja sillä on tahto elää. Havaittavissa olemista ja kolmiulotteisuutta taas voidaan pitää välttämättömänä seurauksena siitä, että tekoöly on alun perin ihmisen kehittämä.

Muzyka on muotoillut metalakeja, jotka perustuvat vastavuoroisuuteen ja tasa-arvoiseen oikeudelliseen asemaan ihmisen ja yleistekoölyn välillä.<sup>25</sup> Muzyka on nostanut esiin ulkopuolisen älykkyyden ja yleistekoölyn keskeiseksi eroiksi sen, että maan ulkopuolinen elämä on syntynyt ilman inhimillistä interventiota. Lisäksi jaamme tekoölyn kanssa saman territorion ja resurssit eli maapallon ja kasvavassa määrin sen lähiavaruuden.<sup>26</sup> Muzyka argumentoi, että huomioimalla nämä erot ja tekemällä joitain muutoksia Fasanin metalakeihin niitä voidaan soveltaa ihmisten ja yleistekoölyn välisiin suhteisiin.<sup>27</sup> Muzyka on tehnyt myös joitain selvittäviä tarkennuksia Fasanin muotoilemiin lakeihin (katso taulukko 1).

Keskeiset muutokset koskevat lakeja oman elintilan käytöstä (nro 6) sekä lakia lajien säilymisen ja kehittymisen suhteesta (nro 8). Niissä on otettu huomioon ensinnäkin se, että jaamme elintilan, ja toisaalta se, että ihminen on synnyttänyt tekoölyn elämänmuotoja ja omaa siten jonkinasteisen suojeluvollisuuden. Yleistekoölyllä on kuitenkin autonomia ja sen myötä itsemääräämisoikeus (nro 4), mistä seuraa, että suojeluvollisuus on molemminpuolinen.<sup>28</sup>

24. Fasan 2017 [1970], s. 211. Avaruudessa olevien etäisyyksien takia tähtienvälinen matkustaminen onnistuisi paremmin epäorgaaniselta kuin orgaaniselta elämänmuodolta. Bohlanderin (2021, s. 402) mukaan on todennäköisempää, että kohdatessamme maan ulkopuolisen elämänmuodon vastassamme on kehittyneen tekoölyn ohjaama konesivilisaatio.

25. Muzyka 2020, s. 104, ks. s. 104–106.

26. Muzyka 2020, s. 110–111.

27. Muzyka 2020, s. 115.

28. Tässä kirjoituksessa oletetaan, että älykkäillä ja tietoisilla olennoilla on oikeus elää myös ilman nimenomaista tahdonilmaisua. Keskeinen ero kuitenkin on, että ilman uudeksi älykkääksi elämänmuodoksi katsotun yleistekoölyn omaa elämisen halun tahdonilmaisua ihminen olisi velvollinen suojelemaan tätä uutta elämänmuotoa. *Voluntas vivendi* vaikuttaa siihen, miten autonomiseksi yleistekoöly voidaan mieltää. Tahdonilmaisun puuttuminen ei tarkoita, etteikö tekoöly tulisi elämänmuotona suojella, vaan se tarkoittaa sitä, että ihmisellä on suojeluvollisuus. Kysymys on, voidaanko tekoölyltä tällöin vaatia vastaavaa suojeluvollisuutta vai onko suojeluvollisuus yksipuolinen siten, että ihmisen on huolehdittava niin tekoölyn suojelemisesta kuin siitä, ettei tekoöly vaaranna ihmisen elämää.

**Taulukko 1.** Ernst Fasanin (2017 [1970], s. 227–236) metalait ihmisten ja maan ulkopuolisen elämän vuorovaikutukselle ja Kamil Muzykan (2020, s. 111–115) muotoilemat vastaavat metalait yleistekoälyn ja ihmisen vuorovaikutukselle. Metalakien järjestys on se, jossa Fasanin on ne ilmaissut. Metalakeja on tarkennettu Muzykan huomioilla, jotka kirjoittaja on pyrkinyt olennaisin osin tiivistämään.

Nro	Fasanin metalait	Metalakien soveltaminen yleistekoälyn ja ihmisten välisiin suhteisiin
1.	Metalakien osapuoli ei voi vaatia mahdollisuuksia.	Mahdottomaksi vaatimukseksi lasketaan esimerkiksi sellainen yksipuolinen vaatimus, joka olisi osapuolelle voitoton tilanne (eng. no-win scenario) tai yksinkertaisesti mahdoton toteuttaa.
2.	Ainuttakaan metalakien sääntöä ei tule noudattaa, mikäli sen noudattaminen johtaisi käytännössä veloitettun lajin itsemurhaan.	Mitä tulee esimerkiksi rajallisiin resursseihin, yhdenkään älyllisen lajin ei tulisi vahingoittaa toista – eikä vaatia toiselta lajilta asioita, jotka käytännössä tarkoittaisivat lajin itsemurhaa.
3.	Kaikilla universumin älyllisillä olennoilla on periaatteessa yhtäläiset oikeudet ja yhtäläinen arvo.	Riippumatta siitä, onko toinen osapuoli vahvempi kuin toinen, kumpikaan metalakien osapuolista ei saa alistaa toista.
4.	Kaikilla metalakien osapuolilla on itsemääräämisoikeus.	Niin ihmisillä kuin yleistekoälyllä on oikeus pyrkiä kohti omia tavoitteitaan.
5.	Kaikkia tekoja, jotka aiheuttavat vahinkoa toiselle lajille, on vältettävä.	Ihmisen ja yleistekoälyn välisissä suhteissa tulisi välttää vahingon tekemistä ja harmin aiheuttamista toiselle osapuolelle.
6.	Kaikilla lajeilla on oikeus omaan elintilaan.	Älyllisten lajien jakaessa yhteisen elintilan heidän tulisi laatia keskinäinen sopimus yhteisten resurssien käyttämisestä.
7.	Kaikilla lajeilla on oikeus puolustaa itseään mitä tahansa toisen lajin suorittamaa vahingollista tekoa vastaan.	Molemmilla suhteen osapuolilla on oikeus itsesuojeluun ja selviytymiseen.
8.	Yhden lajin säilyminen on merkittävämpää kuin toisen lajin kehittyminen.	Ihmiskunnan hyvinvointi ja elinpiirin laajentuminen eivät saa estää yleistekoälyn itsenäisyyttä, suvereniteettia ja kehitystä – ja vice versa. Metalakien osapuolena ihmisten tulisi suojella yleistekoälyjä toisten ihmisten pahantahtoisuudelta ja yleistekoälyjen tulisi suojella ihmisiä haitallisilta vaikutuksilta, jotka aiheutuvat toisten yleistekoälyjen toiminnasta.

9.	Vahingon tapahtuessa sen aiheuttaja on velvollinen palauttamaan vahingon kärsineelle osapuolelle alkuperäisen oikeustilan.	Vahingonkorvausvelvollisuutta on pidettävä kohtuullisena myös ihmisten ja yleistekoälyn välisissä suhteissa.
10.	Metaoikeudellisia sopimuksia on noudatettava.	<i>Pacta sunt servanda</i> - eli sopimus sitoo -periaate.
11.	Toisen lajin auttaminen ei ole oikeudellinen velvollisuus vaan eettinen peruseriaate.	Toisen avustamisen tulisi perustua sopimukseen, ei metalakien täytäntöönpanoon. Vakavissa tilanteissa osapuolen tulisi auttaa toista edellyttäen, että sillä on kestävät edellytykset tehdä niin.

## 4. Metalakien kritiikki

### 4.1. Aluksi

Vaikka metalait ovat hyvin siirrettävissä maan ulkopuolisen elämän ja ihmisten välisistä suhteista ihmisen vuorovaikutukseen yleistekoälyn kanssa, suhtautuminen niihin ei ole varauksetonta. Tämä johtuu itse metalaeista ja niiden luonteesta.

Kirjallisuudessa Haleyn ja Fasanin metalakeja kohtaan esitetty keskeisin kritiikki voidaan tiivistää osa-alueisiin, jotka ovat (1) universalismin kritiikki, (2) antroposentrismin ja liiallisen ennakoimisen kritiikki sekä (3) reaalioliittinen kritiikki. Rajanveto kritiikin eri osa-alueiden välillä on veteen piirretty. Näiden lisäksi Muzyka tekee yleistekoälyn ja ihmisen vuorovaikutuksen metalakeja tarkastellessaan oleellisen täydennyksen käytännön kritiikin osa-alueisiin nostamalla esille vielä yhden kritiikin kohteen eli (4) tietoisuuden tunnistamiseen liittyvän ongelman sekä antropomorfismin. Kyse on käytännössä siitä, miten voimme tietää, että vastassamme oleva tekoäly todella on tietoinen eikä vain vaikuta siltä.<sup>29</sup>

Tässä luvussa pyritään muodostamaan tiivis yleiskuva metalakien kritiikistä ja arvioimaan sitä, missä määrin tämä kritiikki ulottuu myös ihmisten ja yleistekoälyn välisiä suhteita sääteleviin metalakeihin.

29. Mielenkiintoista on, miksi tietoisuuden tunnistamisen ongelma nousi esille vasta yleistekoälyn yhteydessä eikä lainkaan maan ulkopuolisen älyn kohtaamista koskevassa keskustelussa.

## 4.2. Metalakien kritiikki aiemmassa kirjallisuudessa

### 4.2.1. Universalismin kritiikki

Edellä todetulla tavalla metalaki-käsitteen käyttöönottanut Andrew G. Haley otti metalakien perustaksi ”tähtienvälisen kultaisen säännön” – eli ”tee toisille niin kuin haluaisit heidän tekevän sinulle”. Ernst Fasan otti perustakseen kategorisen imperatiivin pyrkien johtamaan siitä metalait, jotka pätevät kaikkiin tietoihin olentoihin koko universumissa.<sup>30</sup> Immanuel Kant itse pyrki irrottamaan kategorisen imperatiivin kultaisesta säännöstä.<sup>31</sup> Tästä huolimatta molempia kohtaan voidaan esittää kritiikkiä niiden universalismin perusteella.

Halu elää ja halu välttää harmia muodostavat pohjan metalaeille. Haley ja Fasan johtavat nämä periaatteet ja niistä kumpuavat metalait rationaalisesti valitsevasta luonnontilasta. Aiemmassa kirjallisuudessa on tarkasteltu kriittisesti metalakien luonnonoikeudellisia perusoletuksia ja huomautettu, että rationaalisuus voi olla hyvin suhteellista.<sup>32</sup>

Vaikka Kant nostaa esille lajienväliset suhteet, hän ei tarkastellut syvällisemmin niihin liittyvää eettistä keskustelua. Haley ja Fasan taas eivät kumpikaan vastaa Kantin kategorista imperatiivia kohtaan aiemmin esitettyyn kritiikkiin.<sup>33</sup> Metalakien universalismin kritiikki jakaa monia elementtejä niiden ihmiskeskeisyyden eli antroposentrismien kritiikin kanssa.

### 4.2.2. Antroposentrismien ja liiallisen ennakoimisen kritiikki

Antroposentrismien kritiikki kyseenalaistaa niin ikään oletuksen siitä, että kaikki älyllinen elämä universumissa jakaisi samat arvot. Fasan itse on argumentoinut, etteivät hänen metalakinsa ole antroposentrisiä, koska ne turvaavat ihmisoikeuksia vastaavat oikeudet myös maan ulkopuoliselle älykkäälle elämälle. Sen sijaan Fasan luonnehtii metalakejaan älykkyyskeskeisiksi (intelligentocentric).<sup>34</sup>

30. Fasan 2017 [1970], s. 223–226 ja Fasan 1998, s. 678. Ks. Immanuel Kant, *The Metaphysics of Ethics*. T. and T. Clark 1796 osoitteessa <https://oll.libertyfund.org/title/calderwood-the-metaphysics-of-ethics> (vierailtu 14.8.2023), luku 1. Fasan viittaa tässä yhteydessä myös Ciceroon, Aristoteleeseen, Tuomas Akvinolaiseen sekä Montesquieu'hön keskustellessaan luonnonoikeudesta ja universaaleista laeista. Ks. Fasan 2017 [1970], s. 228–232.

31. Ks. Michael Bohlander, *Metalaw – What is it good for?* *Acta Astronautica* 188 (2021), s. 400–404, 401.

32. P. Sterns, *Metalaw and relations with intelligent beings revisited*. *Space Policy* 20(2) 2004, s. 123–130, 126–129; Bohlander 2021, s. 401–402; Gertz 2016, s. 263 ja Sterns – Tennen 2017, s. 171.

33. Bohlander 2021, s. 401.

34. Fasan 2017 [1970], s. 240 ja Fasan 1998, s. 678.



Antroposentrismin kritiikki laajentaa tarkastelun jaettujen arvojen ja oikeuksien ulkopuolelle.<sup>35</sup> Se luo perustaa universalismin kritiikille pohtimalla esimerkiksi kysymyksiä vieraiden älyllisten elämänmuotojen fysiologisista ominaisuuksista, jotka määrittelevät muun muassa lajin säilymisen ja hyvinvoinnin edellytyksiä. Metalait heijastavat toisiinsa elämänmuotoihin omia käsityksiään niin kognitiosta<sup>36</sup> kuin kielestä ja kommunikaatiosta<sup>37</sup>. Metalait antavat samat oikeudet niin ihmisille kuin maan ulkopuoliselle äylle, mutta ne on rakennettu inhimillisille lähtökohdille – ei ainoastaan kulttuurisesti vaan myös biologisesti. *George Robinsonin* mukaan tämä metalakien biologinen antroposentrismi johtaa niiden todennäköiseen epäonnistumiseen.<sup>38</sup>

Antroposentrismiin – eli metalakien laatimiseen liiaksi ihmisten näkökulmasta – voidaan helposti liittää liiallisen ennakoimisen kritiikki. Miten voimme ennalta pyrkiä luomaan sääntöjä älyllisen elämänmuodon kanssa, josta emme tiedä mitään?<sup>39</sup>

#### 4.2.3. Reaalipoliittinen kritiikki

Metalakien reaalipoliittinen kritiikki on käytännönläheinen heijastus filosofisemmasta metalakien universalismin kritiikistä. Metalakeihin on sisäänrakennettuna ajatus rauhanomaisesta rinnakkaiselosta ja yhteistyöstä.<sup>40</sup> Tämä ei todistettavasti onnistu edes ihmiskunnan omien kansakuntien kesken, joten millä todennäköisyydellä onnistumme siinä vieraan elämänmuodon kanssa? *Bohlanderin* mukaan reaalipoliittikka edellyttää vihamielisyyteen varautumista rauhan turvaamiseksi: *sic vis pacem, para bellum*.<sup>41</sup> Haleyn metalait sen sijaan ovat idealistisia eivätkä vastaa riittävästi todellisuutta ollakseen hyödyllisiä.<sup>42</sup>

Vertailukohdaksi voidaan ottaa yleisesti avaruus oikeus. YK:n vuonna 1967 ratifoiman ja Suomenkin allekirjoittaman Avaruuden yleissopimuksen (SopS 57/1967)<sup>43</sup> 1 artikla kuuluu seuraavasti: ”Ulkoavaruuden, siihen luettuna kuu ja muut taivaankappaleet, tutkimisen ja käytön on koiduttava kaikkien maiden eduksi ja hyödyksi kiinnittämättä huomiota niiden taloudelliseen tai tieteelliseen kehitystasoon, ja niiden tulee olla koko ihmiskunnan omaisuutta.”

35. Yhteenvedoa metalakien antroposentrismien kritiikistä ks. Robinson 2013, s. 183–186.

36. Muzyka 2020, s. 108–109.

37. Sterns 2004, s. 123–126.

38. Robinson 2013, s. 191–203.

39. Ks. Freeland 2013, s. 93.

40. Muzyka 2020, s. 104.

41. Bohlander 2021, s. 402–403.

42. Freeland 2013, s. 78–79.

43. Sopimuksen virallinen nimi: Yleissopimus valtioiden toimintaa johtavista periaatteista niiden tutkiessa ja käyttäessä ulkoavaruutta, siihen luettuna kuu ja muut taivaankappaleet (eng. Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies, tai The Outer Space Treaty).

Sopimuksen tavoitteena on varmistaa avaruuden rauhanomainen ja kestävä käyttäminen. Sopimuksen mukaan avaruutta ei myöskään saa aseistaa. Se edustaa vahvasti idealistista näkemystä ihmiskunnan toiminnasta. SETI-instituutin mukaan sen laatima protokolla maan ulkopuolisen älyn kohtaamisesta heijastaa Avaruuden yleissopimuksen rauhanomaisia tavoitteita.<sup>44</sup> Yleissopimuksen allekirjoittaneiden joukosta uupuu kuitenkin historian merkittävimmät avaruusvaltiot – Yhdysvallat, Venäjä ja Kiina.<sup>45</sup>

Ne tahot, jotka siihen kykenevät, pyrkivät alistamaan muut omien intressiensä palvelijoiksi. Tällöin metalakien heijastamat periaatteet rauhanomaisuudesta ja osapuolten tasa-arvosta vaikuttavat utopistisilta jo yksin ihmiskunnan toiminnan perusteella – puhumattakaan siitä, ettemme tiedä mitään maan ulkopuolisen älyn rauhanomaisuudesta.

#### *4.2.5. Tietoisuuden tunnistamisen ongelma ja antropomorfismin kritiikki*

Muzyka lisäsi yleistekoälyä koskevassa artikkelissaan metalakien kriittiseen tarkasteluun tietoisuuden tunnistamisen ongelman.<sup>46</sup> Ongelmaa ei ole tarkasteltu aiemmassa kirjallisuudessa. Tätä voidaan pitää mielenkiintoisena, koska huomio on oleellinen pyrittäessä tunnistamaan tietoisuuden tasoa niin maan ulkopuolisissa kuin maapallolle natiiveissa elämänmuodoissa – niin orgaanisissa kuin epäorgaanisissa. Miten tunnistamme, että kohtaamamme elämänmuoto on tietoinen eikä vain vaikuta siltä?

Ensin on kuitenkin tehtävä terminologinen selvennys, joka koskee ”älykkyyttä” ja ”tietoisuutta”. Fasanin metalakien soveltuvuuden vähimmäiskriteereihin kuului älykkyys, mutta sen lisäksi myös elämä ja tahto elää. Älykkyys yksin ei siis riitä perustaksi metalakien soveltamiselle. Vaikka Fasanin määritelmää voidaan kritisoida puutteellisuudesta<sup>47</sup>, lähdetään tässä kuitenkin liikkeelle siitä oletuksesta, että yhdessä nämä kolme piirrettä kuvaavat oliota, jonka toimintaa ohjaa tietoinen päätöksenteko – ainakin siinä määrin, että lakeja ja niiden käytäytymistä ohjaavaa vaikutusta on tarkoituksenmukaista tarkastella.<sup>48</sup>

Haasteena kuitenkin on, ettei tietoisuuden ongelmaa ole ratkaistu. Emme tiedä, mikä mekanismi tuottaa tietoisuuden kokemuksen. Siten emme myöskään pysty ulkopuolisen tarkastelun perusteella varmistamaan, onko vastassamme oleva asia tai olento – mukaan lukien toinen ihminen<sup>49</sup> – todella tietoinen. Ihmis-

44. Ks. SETI-institute 2018.

45. Avaruuden käytöstä käydystä keskustelusta (sopimuksen laatimisen aikaan) ks. Haley 1957.

46. Muzyka 2020, s. 115.

47. Robinson 2013, s. 187.

48. Ks. alaviite 28.

49. Bayesilainen paras arvaus on, että myös muut kaltaisemme ovat samalla tavalla tietoisia kuin me itse. Ks. Anil Seth, Sinuna Oleminen: Uusi Tietoisuuden Tiede. Terra Cognita 2022, s. 135. Tekoälyn sekä maan ulkopuolisen älyn erilaisuuden vuoksi emme pysty soveltamaan

keskeinen ajattelutapa eli antroposentrisyys saattaa johtaa siihen, että annamme asioille tai olennoille inhimillisiä ominaisuuksia eli antropomorfoisimme niitä.<sup>50</sup>

#### 4.3. Kritiikin soveltuminen ihmisen ja yleistekoälyn vuorovaikutuksen metalakeihin

Missä määrin ihmisen ja maan ulkopuolisen älyn välisiin metalakeihin kohdistettu kritiikki soveltuu myös ihmisen ja yleistekoälyn vuorovaikutuksen perussääntöihin? Entä miten Muzykan esiin nostamaan tietoisuuden tunnistamisen ongelmaan tulisi vastata?

Lähdetään liikkeelle universalismin kritiikistä, joka kohdistuu metalakien yleismaailmalliseen luonteeseen. Oletus siitä, että kaikki maan ulkopuolinen älyllinen elämä omaisi samat perustavanlaatuiset käsitykset, edellyttää universalismin periaatteen soveltamista huomattavasti laajemmalle alueelle kuin ihmisen ja maapallolle ”natiivin” yleistekoälyn metalait edellyttävät. Keskeisenä erona on, että yleistekoäly on syntynyt ja kehittynyt ihmisten myötävaikutuksella ja jaamme sen kanssa samat resurssit. Se, että yleistekoäly jakaisi kanssamme arvoja, käsityksiä ja tarpeita, on siis todennäköisempää vaikkakaan ei täysin varmaa.

Universalismin kritiikki soveltuu kuitenkin myös ihmisen ja yleistekoälyn välisen vuorovaikutuksen metalakeihin. Ensinnäkin tulee ottaa huomioon, miten edes ihmiskuntaa edustavat eri kansakunnat eivät ole yksimielisiä siitä, mille pohjalle yhteiskunnat tulisi rakentaa. Metalakien universalismin kritiikki ei kuitenkaan tee niitä tyhjiksi, aivan kuten Yhdistyneiden kansakuntien ihmisoikeuksien yleismaailmallisen julistuksen luonne ei tee siitä merkityksetöntä.

Reaalipoliittisesta näkökulmasta vaikuttaa myös todennäköiseltä, ettemme tule kohtaamaan yleistekoälyä ihmiskuntana vaan kansakuntina tai niiden yhteenliittyminä.<sup>51</sup> Aivan kuten voimme tunnistaa yleisesti erilaisia ihmisoikeuksia ja erityisesti niiden suhdetta tekoälyn käyttöön koskevia lähestymistapoja, voimme tunnistaa erilaisia lähestymistapoja, joita esimerkiksi Yhdysvallat, Eu-

tätä tekniikkaa vieraisiin elämänmuotoihin vastaavalla todennäköisyydellä. Ks. Elisabeth Hildt, *Artificial Intelligence: Does Consciousness Matter?* *Frontiers in Psychology* 10 (2019), s. 1535–1535, 1535.

50. Tietoisuutta ja sen malleja koskevasta keskustelusta ks. Seth 2022 ja David J. Chalmers, *Facing Up to the Problem of Consciousness*. *Journal of Consciousness Studies* 2(3) 1995, s. 200–219.

51. Jos maan ulkopuolinen äly koputtaisi huomenna ovaamme, kohtaisimme heidät todennäköisesti kansakuntina emmekä ihmiskuntana. Hieman erilainen tilanne syntyy, mikäli lähetämme ihmisiä avaruuteen edustamaan ihmiskuntaa. Ks. Rita M. Lauria – George S Robinson, *From Cyberspace to Outer Space: Existing Legal Regimes Under Pressure from Emerging Meta-Technologies*. *University of La Verne Law Review* 33(2) 2012, s. 219–253, 234–235.

roopan unioni ja Kiina edustavat.<sup>52</sup> Metalakien ei tarvitse edustaa yleismaailmallisia arvoja, mutta ne voivat silti heijastaa sitä, millaista maailmaa haluaisimme rakentaa – tunnistaen samalla eroavuudet ja varautuen vastakkainasetteluihin.

Metalakien sisältö ei kuitenkaan ole yksin meistä riippuvainen. Niinpä on varottava antroposentrismien ja antropomorfismin ansoja: emme voi olettaa ymmärtävämmme vieraan elämänmuodon kokemusmaailmaa, haluja ja tarpeita ilman kommunikaatiota.<sup>53</sup> Kommunikaatiosuhteen luomiseksi ongelmana on kuitenkin tietoisuuden tunnistaminen. Tekoälytutkijat eivät pysty tekemään ennustuksia siitä, milloin ihmisen ja tekoälyn suhde muuttuu käyttäjä-työkalusuhteesta sosiaalisiksi suhteeksi, jossa tulisi huomioida osapuolten tasa-arvo.<sup>54</sup> Vielä vaikeampi ongelma on se, ettemme välttämättä tiedä, milloin tämä hetki on saavutettu. Miten erotamme todella tietoisesta elämänmuodosta sitä jäljittelevästä ohjelmasta?

Paras vaihtoehto olisi, että tekoälyjärjestelmän kehittäneet ja sen toiminnan tuntevat kykenisivät arvioimaan, milloin tekoäly toimii tietoisella tasolla.<sup>55</sup> Koska emme kuitenkaan tiedä tarkalleen tietoisuuden synnyn mekanismeja epäorgaanisissa tai edes orgaanisissa elämänmuodoissa, tietoisuuden tunnistaminen nojaa tällä hetkellä ulkoiseen havainnointiin. Ongelmana kuitenkin on, että tekoäly on jo läpäissyt aiemmin tätä varten laaditut testit, kuten Turingin testin, jonka kritisoidaankin mittaavan lähinnä sitä, miten hyvin tekoäly antaa ihmiselle vaikutelman tietoisuudesta.<sup>56</sup>

Mistä sitten toisaalta tiedämme, ettei tämä vaikutelma olisi oikein (”jos se kävelee kuin anka ja kuulostaa ankalta...”)?<sup>57</sup> Tulisiko vuorovaikutuksessa tekoälyn kanssa noudattaa varovaisuusperiaatetta ja myöntää tietoiselta vaikuttavalle tekoälylle oikeuksia pelkän vaikutelman perusteella ja pyrkiä torjumaan tahatonta orjuuttamista, vaikka seurauksena olisikin joukko oikeudellisia zombeja?<sup>58</sup> Zombit eivät kuitenkaan täytä Fasanin kriteerejä elämästä. Oikeuksien

52. Ks. esim. Anu Bradford, Europe’s Digital Constitution. Verfassungsblog. 6.9.2023 osoitteessa <https://verfassungsblog.de/europes-digital-constitution/> (vierailtu 7.9.2023).

53. Ks. Robinson 2013, s. 191–199 ja Muzyka 2020, s. 108–109.

54. Muzyka 2020, s. 105.

55. Ks. esim. Christian List, Group Agency and Artificial Intelligence. *Philosophy & technology* 34(4) 2021, s. 1213–1242, 1219 ja Roger Penrose, *The emperor’s new mind: concerning computers, minds, and the laws of physics*. Oxford University Press 1990, s. 9–10.

56. Ks. esim. Will Oremus, Google’s AI passed a famous test – and showed how the test is broken. *The Washington Post*. 17.6.2022 osoitteessa <https://www.washingtonpost.com/technology/2022/06/17/google-ai-lamda-turing-test/> (vierailtu 11.9.2023).

57. Ks. Joshua Gellers, *Rights for Robots: Artificial Intelligence, Animal and Environmental Law*. Routledge 2021.

58. Filosofisella zombilla viitataan tietoisuutta vailla oleviin olentoihin, jotka kuitenkin käyttäytyvät aivan kuin olisivat tietoisia. Ks. Stevan Harnad, Guest editorial: Why and how we are not zombies? *Journal of Consciousness Studies* 1(2) 1994, s. 164–167 ja Roman Dushkin, To the question on identification and differentiation of a philosophical zombie. *Philosophical Thought* 1/2020, s. 52–66.

myöntäminen pelkästään varmuuden vuoksi olisi ongelmallista myös metalakien integriteetin näkökulmasta.<sup>59</sup>

Onko siis liian aikaista pohtia metalakeja? Ei lainkaan. Vaikka metalait on tarkoitettu pysyviksi, metalakeja voidaan myös muokata, ei ainoastaan tekoälyn vaan ihmisen oman evoluution myötä.<sup>60</sup>

## 5. Tarvitsemmeko metalakeja?

Jos tekoäly saavuttaa avullamme tietoisuuden, miksi tarvitsisimme erillisiä metalakeja? Miksi yhteiskuntasopimus pitäisi neuvotella uudelleen jokaisen uuden sukupolven kanssa – oli kyseessä sitten *homo sapiens sapiens* tai *machina sapiens*?

Yksi metalakeja koskeva keskustelu liittyy nimenomaan historiassa tehtyjen virheiden välttämiseen ja erityisesti siihen, kuinka virheellisesti olemme ihmiskuntana kohdelleet itsellemme vieraita kansoja ja kulttuureita.<sup>61</sup> Tämänkaltaiset keskustelut ovatkin tärkeää itsereflektiota, joka ei koske ainoastaan suhdettamme tekoälyyn vaan suhdettamme ihmiskunnan seuraaviin askeleisiin: elämään maan ulkopuolella, geneettisesti tai teknologisesti paranneltuun ihmiseen eli transhumanismiin sekä elämään yhdessä ei-ihmisten kanssa eli posthumanismiin.<sup>62</sup>

Metalakeihin sisältyvä itsemääräämisoikeus tarkoittaa, että niin kauan kuin siitä ei ole haittaa muille lajeille, ihmiskunta saa kehittyä haluamaansa suuntaan – mahdollisesti kohti osittaista tai täydellistä elektronista elämää. Se tarkoittaa myös sitä, että voimme rajoittaa kehitystä, kuten estää tekoälyä saavuttamasta tietoisuutta.

Keväällä 2023 keskusteluissa nousikin esille mielipiteitä, joiden mukaan meidän tulisi pidättäytyä kehittämästä itseämme suurempaa älykkyyttä – pysäyttää

59. Kriittisesti ihmisoikeuksien inflaatiosta erityisesti Sophia-robotille vuonna 2017 myönnetyn Saudi-Arabian kansalaisuuden yhteydessä ks. Jaana Parviainen – Mark Coeckelbergh, The political choreography of the Sophia robot: beyond robot rights and citizenship to political performances for the social robotics market. *AI & Society* 36(3) 2021, s. 715–724. Silloin kun tekoälylle myönnetään oikeuksia esimerkiksi joidenkin tehtävien suorittamiseen, ne rinnastuvat yrityksiin ja muihin oikeushenkilöihin, jolloin merkitystä ei ole sillä, onko niillä tietoisuutta tai ei. Ks. esim. Hildt 2019, s. 2–3. Myös oikeuksien rajoittuminen siihen, mikä on tekoälyn tehtävän kannalta tarkoituksenmukaista (*a posteriori*), tarkoittaa, ettei tästä seuraa ihmisoikeuksien inflaatiota. Mutta tietoisuus luo velvollisuuden ihmiselle joko suojella tekoälyä tai nostaa se tasavertaiseksi oikeushenkilöksi *a priori* oikeuksineen ja velvollisuuksineen.

60. Robinson 2013, s. 211, 213, 216.

61. Freeland 2013, s. 92.

62. Ks. esim. Lovelock & Appleyard 2019, s. 84 ja Robinson 2013, s. 209, 216.

tekoälyn kehitys kokonaan ja jopa ottaa pari askelta taaksepäin.<sup>63</sup> Havaitsemme jälleen paljon yhtymäkohtia keskusteluun kohtaamisesta maan ulkopuolisen elämän kanssa. Monet tunnetut tiede-elämän edustajat, kuten *Stephen Hawking* ja *Neil DeGrasse Tyson*, ovat ilmaisseet huolensa siitä, mikä on ihmisen asema tällaisissa tapaamisissa.<sup>64</sup> *John Gertzin* mukaan meidän ei kannattaisi pitää meteliä itsestämme.<sup>65</sup> Hän onkin esittänyt Asilomarin tekoälyä koskevien periaatteiden kaltaisten rajoitteiden asettamista myös avaruustutkimukselle.<sup>66</sup>

Pyrkimykset rajoittaa tekoälyn kehitystä ovat kuitenkin olleet asynkronisia ja siten tehottomia. Avoimista kirjeistä ja julkilausumista huolimatta teknologinen kehitys ei ole pysähtynyt hetkeksikään. Sama toistuu myös ihmisen ominaisuuksien geneettisessä ja teknologisessa parantelussa. Pyrimme estämään ihmisten liian pitkälle vietyä manipulointia siihen liittyvien eettisten syiden vuoksi. Reaalipoliittisista syistä emme kuitenkaan pysty tehokkaasti kontrolloimaan teknologian ja ihmisen evoluutiota, vaikka haluaisimme.<sup>67</sup> Mitä siis teemme, jos vastaamme käveleekin tuskin ihmiseksi tunnistettava kyborgi: Onko hänellä silti ihmisoikeudet? Jos on, niin millaiset? Pitääkö meidän kirjoittaa uudelleen oman yhteiskuntamme metalait?

Reaalisista ja spekulatiivisista uhista huolimatta tekoäly ja sen sovellukset voivat osoittautua ihmiskunnan selviytymisen kannalta välttämättömäksi: me tarvitsemme tekoälyä enemmän kuin tekoäly meitä.<sup>68</sup> Metalakeihin liittyvä keskeinen kysymys onkin se, miten me käyttäydymme.<sup>69</sup> Tietoinen tekoäly on todennäköisesti vielä kaukana, mutta koskaan ei ole liian aikaista tai antroposentristä pohtia, millaisten metalakien varaan ihmiskunnan tulevaisuus rakennetaan.

63. Ks. esim. Eliezer Yudkowsky, Pausing AI Developments Isn't Enough. We Need to Shut it All Down. TIME ideas 29.3.2023 osoitteessa <https://time.com/6266923/ai-eliezer-yudkowsky-open-letter-not-enough/> (vierailtu 30.8.2023); Muzyka 2020, s. 106–107 ja Robin Murphy – David D. Woods, Beyond Asimov: The Three Laws of Responsible Robotics. IEEE Intelligent Systems 24(4) 2009, s. 14–20.

64. Ks. John Gertz, Reviewing METI: A Critical Analysis of the Arguments. arXiv.org 2021. Myös Mars-siirtokuntaa suunnitteleva Elon Musk on allekirjoittanut anti-METI-vetoomuksen.

65. Ks. Gertz 2016, s. 264.

66. Gertz 2016, s. 268.

67. Ks. esim. NATO Science & Technology Organization, Science & Technology Trends 2020–2040. NATO Science & Technology Organization 2020 ja Doris Vogl, Human Enhancement with Chinese Characteristics: Public Discourse Patterns on Human-machine Teaming in the People's Republic of China. Maanpuolustuskorkeakoulu 2023.

68. Ks. Lovelock & Appleyard 2019, s. 106–113.

69. Bohlander 2021, s. 402–403.