

Semioottisesti kyvykkäät koneet sovelluksineen

Timo Honkela

Käsitteellöity, julkaistu Futura-lehdessä 2/2012, ss. 27-41.
Julkaisija Tulevaisuuden tutkimuksen seura ry, Helsinki.

Tässä artikkelissa arvioidaan tekoälyn ja sen lähialueiden tulevaisuuden mahdollisia kehityspolkuja tavoitellen tavanomaisesta poikkeavia näkökulmia. Tavoitteena ei siis ole sen ajankohdan ennustaminen, jolloin robottijoukkue voittaa jalkapallon maailmanmestarijoukkueen, tai jolloin autojen ohjaamiseen on kehitetty automatiikka korvaa ihmiset heitä parempina kuljettajina ja rajaa autolla ajamisen suljettujen ratojen vaaralliseksi harrastukseksi. Tekoälyn tulevaisuutta tarkastellaan lähtökohtaisesti filosofisena, yhteiskunnallisena, kielitieteellisenä ja psykologisena kysymyksenä kuin menetelmällisteknologisena aiheena.

Johdanto

Suomen tekoälyseuran laaja hanke tuotti noin 20 vuotta sitten tuloksenaan Tekoälyn ensyklopedia-nimisen kirjan (Hyvönen et al. 1993). Kirjassa esitellään kattavasti tekoälyn lähitieteitä ja osa-alueita kuten havaitseminen, tietämyksen esittäminen, ongelmanratkenta, ymmärtäminen, oppiminen ja opettaminen. Yksittäisinä aiheina ovat esillä esimerkiksi sääntöjärjestelmät, neuroverkot, puheenkäsittely, konenäkö, päätösanalyysi, luonnollisen kielen käsittely, koneoppiminen ja hypermedia. Sovelluksina tarkastellaan asiantuntijajärjestelmiä, lääketieteen järjestelmiä, teollisuusautomaatiota ja robotiikkaa. Kaikki nämä alueet ovat edelleen aktiivisen tutkimus- ja kehitystyön kohteita. Vaikka pinnallisesti tarkasteltuna tietotekniikan voidaan ajatella kehittyvän huimaa vauhtia, monet nykyaikaisen tietojenkäsittelyn pohjalla olevat ajatukset ovat olleet olemassa jo pitkään. Ideoiden ja käsitteiden tasolla perustavaa laatua olevia kehitysaskeleita otetaan harvakseltaan. Esimerkiksi www:n käytännöllinen merkitys on ollut valtava, mutta jo vuonna 1945 Vannevar Bush muotoili sen perusideat (Bush 1945). Douglas Engelbart ja Ted Nelson jatkoivat näiden ideoiden kehittelyä 1960-luvulla (ks. esim. Heimbürger & Kuhanen 1983). Räjähdyksmäisen kehityksen mahdollisti Tim Berners-Leen vuonna 1990 muotoilema Hypertext Transfer Protocol

(HTTP), joka määritteli tietokoneiden välisen tiedonsiirron standardinomaisesti ja mahdollisti sen, että keskenään yhteensopivista osista rakentui laaja kokonaisuus yhteisesti hyväksytyn kielen puitteissa. HTTP:stä muodostui nopeaan kasvuun lähteneen world wide webin (www) ydin. Tekoälyseuran hallituksessa www:stä keskusteltiin 1990-luvun alussa tuoreeltaan innostuneesti, mutta yritysedustaja toppuutteli intoa todeten, että asiaan ei liity juurikaan kaupallisia intressejä – ja liittyi siten lukuisiin tulevaisuutta epäonnistuneesti arvioineisiin. Vaikka tekoälytutkimuksen käytännön vaikutus on ollut monilta osin erittäin suuri, on siis hyvä muistaa, että monesti tie tutkijan ajatuksista käytännön sovelluksiin on nykyäänkin pitkä. Nopeaa reittiä käytäntöön voivat siirtyä ne ideat, joiden teknologiset puitteet ovat jo olemassa, ja joiden omaksumiseen ja käyttöönottamiseen ihmisillä on riittävät valmiudet (Shove et al. 2012).

Tässä artikkelissa arvioidaan tekoälyn ja sen lähialueiden tulevaisuuden mahdollisia kehityspolkuja tavoitellen tavanomaisesta poikkeavia näkökulmia. Tavoitteena ei siis ole sen ajankohdan ennustaminen, jolloin robottijoukkue voittaa jalkapallon maailmanmestarijoukkueen, tai jolloin autojen ohjaamiseen on kehitetty

automaatiikka korvaa ihmiset heitä parempina kuljettajina ja rajaa autolla ajamisen suljettujen ratojen vaaralliseksi harrastukseksi. Tekoälyn tulevaisuutta tarkastellaan lähtökohtaisesti filosofisena, yhteiskunnallisena, kielitieteellisenä ja psykologisena kysymyksenä kuin menetelmällisteknologisena aiheena.

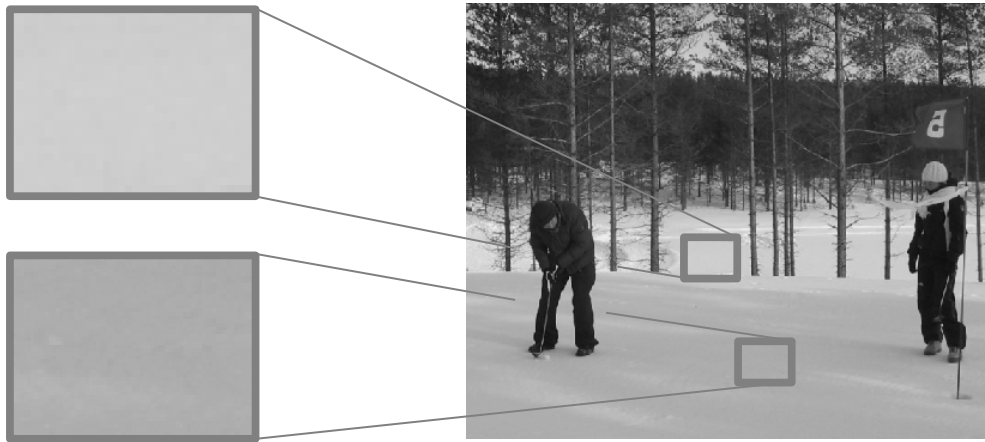
Semioottisesti kyvykkäät koneet

Semioottinen kyvykkyys voidaan määritellä kyvyksi käyttää merkkejä asianmukaisesti yhteisössä muodostuneiden käytäntöjen mukaisesti niin, että järjestelmä tuntee merkkien ja objektien väliset kytkökset ja osaa tarvittaessa ehdottaa uusia käytäntöjä (Honkela 2010). Semioottisesti kyvykäs järjestelmä kykenee esimerkiksi luomaan yhteyden jonkin konkreettista esinettä tai asiaa tarkoittavan sanan ja sen moninaisten visuaalisten ilmentymien välille. Jotta kykyä voitaisiin pitää todellisena, se ei voi perustua ihmisen ohjelmoiman linkin muodostamiseen tietokoneelle talletetun merkkijonon ja yksittäisen tai useampien kohdetta esittävien kuvapistekokoelmien välille. Tällöin kyseessä on ulkoa annettu linkki, jota kohtaan voidaan esittää samanlaisia huomioita kuin mitä John Searle on esittänyt kiinalaisen huoneen ajatuskokeessaan (Searle 1980). Jotta semioottisen linkin muodostaminen olisi mahdollista inhimilliseen tapaan, järjestelmällä täytyy olla hahmontunnistuskkykyä. Keskeinen perinteisen tekoälyn ongelma on se, että symbolit ”roikkuvat ilmassa” eli representaatiot eivät ole linkittyneitä havaintokokemuksen valtavaan hahmovirtaan. Tätä symbolisten esitysmuotojen rajoitusta kutsutaan symbol grounding -ongelmaksi (Harnad 1999). Hahmontunnistusmenetelmillä (Theodoridis & Koutroumbas 2010) voidaan rakentaa järjestelmiä, jotka mahdollistavat symbolien juurruttamiseen hahmomuotoiseen havaintokokemukseen. Itsenäisesti toimivien kieltä oppivien robottien kehittämisessä tällaiset ominaisuudet ovat erityisen tärkeitä (ks. esim. Vogt 2000).

Symbolien juurruttaminen havaintokokemukseen mahdollistaa sellaisen merkitysten käsittelyn koneellisesti, joka ei ole mahdollista perinteisessä symboliseen logiikkaan perustuvassa mallintami-

nessa. Kun loogikot käsittelevät merkitystä, he ovat tyypillisesti jääneet kiinni kielen maailmaan: kielellisiä merkityksiä määritellään kielen sisällä – olkoonkin, että luonnollisen kielen ilmaisujen merkityksiä kuvataan formaalien kielten ilmaisuilla, mutta niiden suhde ympäröivään todellisuuteen on silti ohut ja vaatii aina inhimillisen tulkitsijan. Semioottisesti kyvykkään koneen kehittämisessä perinteinen logiikka on sellaisenaan liian vajavainen väline. Niinkin selväpiirteisteltä vaikuttavien asioiden kuin 'lumi' tai 'valkoisuus' määrittelyssä kielen sisällä pysyminen ei anna mahdollisuutta tavoittaa näiden sanojen merkityssisältöä arkisten tilanteiden näkökulmasta (vrt. kuva 1). Merkitysten tavoittamisessa ei juurikaan auta viittaaminen metakieleen eli loogikon ilmaisu ”lumi on valkoista jos ja vain jos lumi on valkoista” jää köykäiseksi, kuinka suuri rakennelma formaalia tarkastelua sen ympärille rakennetaankaan. Myöskään sanan 'punainen' merkityksen määrittäminen eri konteksteissa ei ole yksinkertaista: ilmaisuissa ”punainen paita”, ”punainen iho” ja ”punaviini” punaisuus viittaa eri kohtiin väriavaruudessa (Gärdenfors 2000).

Jos järjestelmän tekemät havainnot ulkopuolisesta maailmasta esitetään esimerkiksi moniulotteisena jatkuvana avaruutena (Gärdenfors 2000), merkitysten eroja ja samankaltaisuuksia voidaan tarkastella matemaattisina ja tilastollisina suureina (Honkela et al. 2008). Kun tällainen lähestymistapa omaksutaan, symboleita tarkastellaan ulkoisina representaatioina. Kognition sisäisten representaatioiden puolestaan ei tarvitse lainkaan muistuttaa ulkoisia representaatioita (Knuuttila & Honkela 2005). Toisin kuin siis esimerkiksi Fodor (1975) ja monet hänen seuraajansa väittävät, symbolinen tai looginen kieli ei ole mielen sisäinen rakenne vaan se on ulkoinen representaatio, jonka ihmiset yksilöinä ja yhteisönä ovat kehittäneet tiedon jakamista ja jalostamista varten. Tämän näkemyksen mukaan käsitejärjestelmät ovat dynaamisia järjestelmiä, jotka kehittyvät ajan myötä maailman muuttumisen ja erilaisten käyttötarkoitusten syntyminen myötä (Honkela 2010).



Kuva 1. Esimerkki visuaalisesta havainnosta, jossa lumen määrittelemisen sen valkoisuuden avulla ei onnistu.

Ihmisen hahmontunnistuskyyvyt ovat niin kehittyneitä ja monesti luonteeltaan tiedostamattomia, että helposti ei tule ajatelleeksi, kuinka monimutkaisesta prosessista on kyse. Evoluution kuluessa on ollut tärkeää tunnistaa erilaisia hahmoja, jotta voi erottaa ympäristön mahdollisuudet ja uhat. Ihmisellä näyttää olevan joitakin perinnöllisesti viritettyjä hahmontunnistuskyykyjä (esimerkiksi kasvojen tunnistaminen kasvoiksi), mutta suurelta osin yksityiskohdat opitaan. On osoitettu, että edes hyvin perustavaa laatua olevia piirretunnistimia ei tarvitse olla aivoissa valmiiksi ”ohjelmoituina” vaan niitä voidaan oppia näköhavaintokokemuksista (Hyvärinen et al. 2009). Näyttääkin siltä, että hahmontunnistuskyyky perustuu suurelta osin oppimiseen.

Koneoppimisella tarkoitetaan sellaista tietojenkäsittelytieteen ja tekoälytutkimuksen osaluuetta, jossa järjestelmä kehittyy sille annettujen esimerkkien perusteella. Järjestelmän tehtävänä voi esimerkiksi olla se, että se ratkaisee jonkin ongelman ajan myötä aina vain paremmin, tai että se oppimisen kautta muuttaa mallejaan muuttuvasta maailmasta. Järjestelmä oppii kohdealuetta koskevan ongelmanratkaisussa tarvittavan tietämyksen sen sijaan, että tietämys kirjattaisiin järjestelmään käsin esimerkiksi sääntöinä. Ohjaamattomassa oppimisessa järjestelmä muodostaa itsenäisesti mallin sille annetuista esimerkeistä (ks. esim. Oja 2002). Ohjaamatonta oppimista voidaan pitää kognitiivisesti tärkeänä mallina, koska ihmisen

kokemuspiiriin tulevien havaintojen arvioimista ja luokittelusta vastaa ihminen itse. Tämä ei tietysti tarkoita sitä, etteikö havainnossa voi olla mukana muiden ihmisten tarjoamia luokittelevia elementtejä, joista tärkeimpinä voidaan pitää kielellisiä ilmaisuja. Jos aikuinen toistuvasti käyttää samaa sanaa kahden keskenään erilaisen ilmiön tai hahmon yhteydessä, lapsi vähitellen liittää nämä hahmot toisiinsa, vaikka ne eivät muuten muistuttaisikaan toisiaan. Kielelliset konventiot siirtyvät sukupolvelta toiselle mukanaan maailman jäsentämisen lainalaisuudet. Merkitysten tasolla nämä konventiot voivat pahimmillaan sisältää ja ylläpitää epätarkoituksenmukaisia jaotteluita, jotka vaikuttavat ihmisten ajatteluun formatiivisella tasolla. Tässä suhteessa esimerkiksi feministien ja yhteiskuntasuuntautuneiden kielentutkijoiden huolta sukupuolierottelevasta tai rasisisesta kielenkäytöstä voidaan pitää perusteltuna. Tämäntyyppistä kielen vaikutusta ajatteluun voidaan simuloida myös tietokonemalleilla.

Keskeisiä ohjaamatonta oppimista mallintavia tietojenkäsittelymenetelmiä ovat Kohosen itseorganisoiva kartta (Kohonen 1984; 2001) ja riippumattomien komponenttien analyysi (Hyvärinen et al. 2001).

Michael Polanyi on todennut, että yksilöinä tiedämme enemmän kuin voimme kertoa (Polanyi 1967). Tällaisella hiljaisella tiedolla on tärkeä merkitys hyvin monenlaisissa inhimillisissä toiminnoissa polkupyörällä ajamisesta suuryritysten johtamiseen. Asiasta on Suomessakin kirjoitettu paljon (esim. Koivunen

1997), mutta hiljaisen tiedon suhdetta tietokoneisiin on käsitelty varsin vähän. On ehdotettu, että hiljaista tietoa voitaisiin kerätä talteen tekstimuodossa. Hiljaisen tiedon tai intuitioksi kutsutun kyvyn mallintaminen yksityiskohtaisesti kielellisinä esitysmuotoina tai loogis-analyttisillä välineillä on kuitenkin vaikeaa tai jopa periaatteellisesti mahdotonta. Hiljaista tietoa ja intuitiota voidaan kuitenkin hyvin mallintaa neuroverkkojen kaltaisissa koneoppivissa järjestelmissä (Brunak & Lautrup 1990). Eksplisiittisten sääntöjen tai muiden symbolisten representaatioiden sijasta tietoa ja päättelyä mallinnetaan neuroverkoissa suurena joukkona keinotekoisien neuronien välisiä yhteyksiä. Vastaanlaista tietojenkäsittelyä toteutetaan nykyään usein tilastollisissa tai probabilistisissa koneoppimismalleissa (ks. esim. Huopaniemi et al. 2010, Peltonen et al. 2010). Käytännössä edellä mainittu tarkoittaa, että yksittäiseen päätelmään tai luokitteluun voi vaikuttaa jopa tuhansia eri tekijöitä tavalla, jossa yksittäisen tekijän vaikutus ei ole suoraviivaisesti selitettävissä. Tämä ilmiö muistuttaa läheisesti empiirisen psykologian havaintoa, jonka mukaan tiedostamaton (mitä tässä pidetään synonyyminä intuitiivisille) päättely tuottaa monimutkaisissa ongelmanratkaisutilanteissa systemaattisesti paremman tuloksen kuin niin sanottu rationaalinen päättely (Dijksterhuis ym., 2006).

Väliyhteenvetona voidaan todeta, että tulevaisuuden tekoäly saavuttaa lisääntyvässä määrin semioottista kyvykkyyttä, käsitteellistä luovuutta ja intuitiiviseksi luonnehdittavaa päättelykykyä perustana tutkimustulokset tilastollisen koneoppimisen, probabilistisen mallintamisen ja hahmontunnistuksen alueilla. Tutkimus on jo niin pitkällä, että kyse ei ole siitä, tapahtuuko tällaista kehitystä vai ei, vaan siitä, kuinka pitkälle kehitys johtaa ja miten sitä sovelletaan. Kognitiotieteen näkökulmasta mallintaminen voi tarjota uutta tietoa esimerkiksi kielellisen ja muun kognition välisestä suhteesta.

Seuraavassa tarkastellaan muutamia alueita, joilla tekoälyn kehityksellä voi olla voimakas yhteiskunnallinen vaikutus. Esitys koostuu skenaarioista, joiden toteutuminen tekoälyn teknologisista puitteista tarkastellen vaikuttaa mahdolliselta tai jopa todennäköiseltä.

Yhteiskunnan tilanneanalyysi ja simulointi

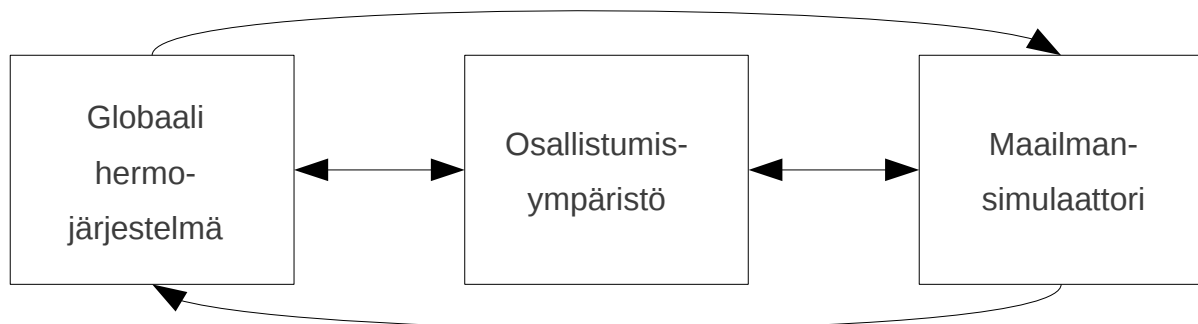
Yhteiskunnat ovat sisäisesti ja toisiinsa eri tavoin kytkeytyneitä valtavia dynaamisia järjestelmiä. Yhteiskunnallisten ja erityisesti talouteen liittyvien ilmiöiden dynamiikkaa mallinnetaan usein matemaattisilla kaavoilla, jotka kuvaavat erilaisten muuttujien kuten kysyntä, hinta ja tuotantokustannukset tai kansantulo ja nimelliskorko välisiä suhteita. Tällaisessa lähestymistavassa esimerkiksi konteksti-ilmiöiden ja eri yksityiskohtaisuustasojen välisten suhteiden mallintaminen on vaikeaa. Lainalaisuudet voivat myös muuttua, kun järjestelmässä tapahtuu olennaisia rakenteellisia muutoksia. Agenttipohjainen mallintaminen on yleistynyt edelläkuvatulle vaihtoehtoisena lähestymistapana. Agenttipohjaisessa mallintamisessa ilmiöitä tarkastellaan toimijajoukon eli simuloidun populaation avulla. Toimijoille on ohjelmoitu käyttäytymis- ja vuorovaikutustavat, joita ne toteuttavat. Koko populaation tarkastelu voi tuoda esille erilaisia emergenttejä eli kehkeytyviä ominaisuuksia. Erinomainen yksittäinen esimerkki simulointilähestymistavan hyödyntämisestä on Linturin (2010) kehittämä malli, joka muun muassa selittää rahan emergenssin vaihdantatalouden sisältä. Joissakin malleissa otetaan koneoppiminen mukaan eli yksittäiset toimijat voivat myös oppia. Vieläkin monimutkaisempia ovat mallit, joissa oppimista tapahtuu sekä evolutiivisesti (yhdestä populaatiosta seuraavaan sukupolveen siirryttäessä) että yksikehityksenä. Evolutiiviset mallit voivat tavoitella biologisen evoluution, kulttuurievoluution tai jopa molempien simulointia.

Tulevaisuuden mahdollisuuksia yhteiskunnan tilan ja sen kehityksen mallintamisessa voidaan tarkastella FuturICT-nimisen hankkeen kautta (www.futurict.eu). Kyseessä on laaja keihäänkärkihankekokonaisuus, joka kilpailee vuoden 2012 aikana viiden muun hakijan kanssa noin miljardin euron rahoituspanoksesta kymmenen vuoden ajalle jaettuna. Hankkeen liikkeellelähdestä suunnittelussa laajuudessa ei siis ole varmuutta. Suunnitelmat antavat kuitenkin hyvin tietoa siitä, minkälaisia välineitä laajamittaisen sosiokognitiivisen mallintamisen

alueella halutaan kehittää yhteiskunnan tilaa tarkasteleviksi sosioskoopeiksi ja sosio-taloudellispoliittisiksi kompasseiksi lainataksemme hankkeen koordinaattoreiden käyttämiä uudissanoja. Kuvassa 2 on esitetty kokonaisjärjestelmän keskeiset osat. Globaalin hermojärjestelmän on tarkoitus kerätä laajasti tietoa yhteiskunnan tilasta. Osallistumis-ympäristössä ihmiset voivat tarkastella erilaisia tietoja, malleja ja ennusteita ja antaa ruohonjuuritason tietoa mallintamisessa hyödynnettäväksi. Maailmansimulaattoria käytetään mallien, ennusteiden ja skenaarioiden tuottamiseen. Yksi tärkeä osa yhteiskunta-analyysia on monimutkaisten verkostojen rakenteen analyysi (ks. esim. Onnela et al. 2007).

tyhjäksi: järjestelmällä ei ole mitään tuntumaa siihen, miltä valtiot, demokraattisuus, ihmiset tai rakastaminen näyttävät, tuntuvat tai kuulostavat.

Ihmiset muodostavat ymmärryksensä sanojen merkityssisällöstä omakohtaisen moniaistisen kokemuksensa kautta. Kukin ihminen kulkee omaa yksilöllistä elämänpolkuaan. Tämän seurauksena jokaisen ihmisen tapa käsitteellistää maailmaa on yksilöllinen. Yksilöllisyys on kuitenkin suhteellista, koska ihmisen biologisen olemuksen, kielen käytön ja maailman säännönmukaisuudet johtavat siihen, että eri ihmisten malli maailmasta sekä malli maailman ja kielen välisestä suhteesta on suurelta osin varsin samankaltainen.



Kuva 2. FuturICT-hankkeessa määritelty suunnitelma järjestelmäksi, jolla voidaan tuottaa laajamittaista ja syvällistä tietoa yhteiskunnan tilasta ja testata sitä koskevia skenaarioita (www.futurict.eu).

Ymmärryksen yksilöllisyys ja merkitysneuvotteleva kone

Perinteisessä tekoälyssä on lähdetty siitä, että kielen ja tiedon perustana olevat käsiterakenteet ovat valtaosin jaettuja. Tietämyksen esittämistä koskevassa tutkimuksessa on sama tietoteoreettinen lähtökohta kuin symbolisessa logiikassa: maailma nähdään erillisinä tapahtumina, olioina, olioiden ominaisuuksina ja olioiden välisinä suhteina. Mallit ovat lähtökohtaisesti kielellisiä. Loogikko määrittelee lauseiden “Suomi on demokraattinen maa” ja “Liisa rakastaa Mattia” merkityksen esimerkiksi predikaattilogiikan ilmaisuina “ $\text{valtio}(x) \wedge \text{demokraattinen}(x) \wedge x = \text{'Suomi'}$ ” ja “ $\text{rakastaa}(a, b) \wedge a = \text{'Liisa'} \wedge b = \text{'Matti'}$ ”. Tällan tietämyksen määrittely jää kuitenkin varsin

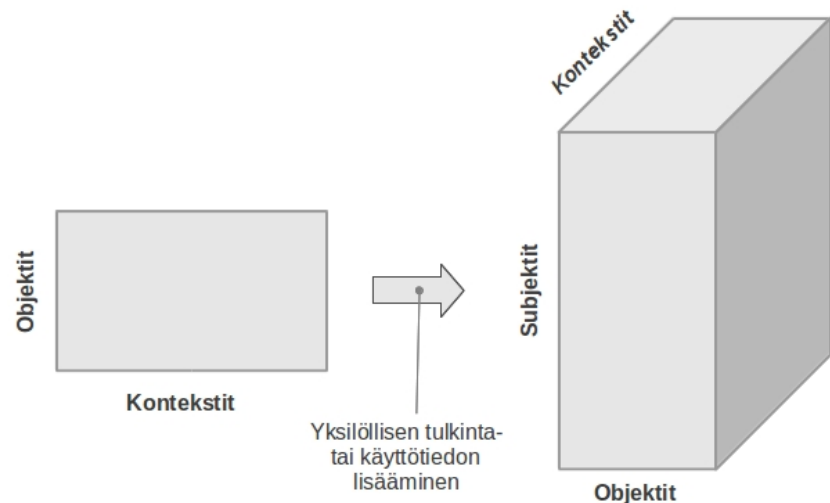
Yksittäisillä elämänalueilla näiden mallien erot voivat kuitenkin olla huomattavia. Klassinen esimerkki eroista maailman hahmottamisessa on nk. Amesin huonetta koskeva koeasetelma, jossa ihmisen pituus näyttää selvästi muuttuvan, kun hän kävelee huoneen nurkasta toiseen. Ihmiset, joiden rakennetussa ympäristössä ei yleensä ole suorakulmaisia rakenteita, näkevät huoneen sellaisena kuin se on vinoine rakenteineen, eivätkä tulkitse sitä suorakulmaiseksi huoneeksi (Segall et al. 1963).

Kielen mallien eroja koskeva triviaali esimerkki on erikoisalojen kieli. Sanat “agenesia”, “affiini” ja “polyploidia” eivät merkitse mitään ihmisille, joille vastaavat lääketieteen, matematiikan ja biologian sisällöt eivät ole tuttuja. Yksilöllisiä merkityseroja voidaan kuitenkin havaita paljon arkisemmissakin tilanteissa. Värien nimeäminen vaihtelee, jos kyse ei ole pääväreistä. Ihmisillä voi olla hyvinkin erilaisia tulkintoja siitä, mitä tarkoittavat sanat “pian”, “reilu” tai “kohtuul-

linen”. Konkreettisiin asioihin viittaavat substantiivit kuten “talo” tai “koira” ymmärretään useimmiten yhteneväisesti, mutta esimerkiksi sanojen “demokratia”, “kannattavuus” tai “oikeudenmukaisuus” sisältö vaihtelee tulkitsijasta toiseen.

Arto Mustajoki (2012) on tutkinut monipuolisesti kielellistä vuorovaikutusta ja siinä tapahtuvia epäonnistumisia. Hänen mukaansa vuorovaikutuksen ongelmia voivat selittää ihmisten erot kielen osaamisessa, kulttuuritaustassa ja kognitiivisissa ominaisuuksissa (Mustajoki 2008). Onnistuneen vuorovaikutuksen perustana on se, että ottaa huomioon viestin vastaanottajan. Tämän prosessin esteenä voivat olla esimerkiksi (a) vaivan välttäminen, (b) erehtyminen luulemaan, että kuulijan ymmärrys ilmiöstä tai tausta on samankaltainen kuin puhujalla, (c) yli-innokkuus käsiteltävänä olevan asian suhteen, jolloin kuulija unohtuu, (d) ongelmallinen fysiologinen tai psyykinen tila (esimerkiksi humala), tai (e) muut viestinnän tavoitteet kuin selkeys (esimerkiksi halu tehdä vaikutus) (Mustajoki 2012). Ihmisten välinen vuorovaikutus ei siis aina onnistu suoraviivaisesti. Yhteistä ymmärrystä voidaan hakea esimerkiksi merkitysneuvotteluiden kautta, mutta yhteisymmärryksen puute voi myös jäädä pysyväksi.

Merkitysten mallintaminen jatkuvina moniulotteisina suureina mahdollistaa jaettujen merkitysten kehkeytymisen ja merkitysneuvotteluiden laskennallisen mallintamisen (Honkela 1993, Steels & Kaplan 2000, Lindh-Knuutila et al. 2006, Honkela et al. 2008). Tulevaisuuden koneet voivat käydä merkitysneuvotteluita sekä keskenään että ihmisten kanssa nykyistä paljon suuremmassa mittakaavassa ja auttaa ihmisiä ymmärtämään kielen käyttöä sekä yksilöllisellä että yhteisöllisellä tasolla. Askel tähän suuntaan on GICA (Grounded Intersubjective Concept Analysis) -niminen menetelmä (Honkela et al. 2012). Menetelmän ytimessä ovat subjekti-objekti-konteksti -tietorakenteet, jotka kuvaavat kuinka eri ihmiset tulkitsevat ilmaisuja eri konteksteissa.



Kuva 3. GICA-menetelmässä käytetyt subjekti-objekti-konteksti -tietorakenteet perinteisten sanadokumenttimatriisien ja niiden yleistyksien laajennuksena (Honkela et al. 2012).

Sanojen merkityksiä on pystytty jo yli kahden vuosikymmenen ajan mallintamaan automaattisesti keräämällä tietoa siitä, millaisissa konteksteissa sanoja käytetään ja sitten analysoimalla tätä tietoa tilastollisesti (Ritter ja Kohonen 1989, Deerwester et al. 1990). Sanojen merkitysten samankaltaisuus paljastuu sitä kautta, että samaan merkityserheeseen kuuluvia sanoja käytetään suhteellisesti useammin samassa tekstidokumentissa kuin satunnaisia sanapareja. Kun dokumentteja analysoidaan suuri määrä, sanojen väliset suhteet alkavat hahmottua luotettavasti lukuunottamatta hyvin harvoin esiintyviä sanoja.

GICA-menetelmä laajentaa edellä kuvattua perusasetelmaa lisäämällä siihen subjektiuden: sanojen käyttöä ja tulkintaa tarkastellaan yksilökohtaisesti (kuva 3) (Honkela et al. 2012). Yksinkertaisia esimerkkejä kielen tulkinnan vaihtelusta ovat alueelliset erot. Esimerkiksi sana tuima tarkoittaa eri asioita eri puolilla Suomea. Tällöin yksilöllistä kielenkäyttöä selittävä tekijä on maantieteellinen alue. Toinen esimerkki voidaan poimia naapurimaastamme Ruotsista. Siellä käydään keskustelua lainsäädännön muutoksesta koskien sairauseläkkeelle jäämisen ehtoja. Debatin kohteena ovat fraasit ”reguljära arbetsmarknaden” ja ”normalt förekommande arbete”. Eri poliittisten puolueiden edustajilla,

oikeuslaitoksen edustajilla, sosiaalityöntekijöillä ja maallikoilla on vaihtelevia käsityksiä siitä, mitä näillä fraaseilla tarkoitetaan ja miten niiden varaan rakennettu päätöksenteko toimii erilaisissa tilanteissa. Tällainen lainsäädäntöön ja sen soveltamiseen liittyvä tapaus on hyvä esimerkki tilanteesta, jossa merkityssisällön ja -variaation huolellinen analyysi olisi erittäin tärkeää. Esimerkki nostaa hyvin myös GICA-menetelmän peruselementit: objekteja ovat tässä kaksi edellämainittua ruotsinkielistä fraasia, subjekteja ovat henkilöt eri intressiryhmistä, ja konteksteja ovat tilanteet, joihin fraasien katsotaan tai ei katsota viittaavan. Prosessin toimivuuden varmistamiseksi GICA-analyysi olisi tällaisessa tilanteessa hyvä tehdä niin, että mukana olisi vähintään kymmeniä poliitikkoja ja lainoppineita sekä mieluummin satoja tai tuhansia maallikoita. Tulevaisuuden yhteiskunta voisi olla sisällyttää kaikkeen yhteiskunnalliseen keskusteluun, päätöksentekoon ja lakien soveltamiseen tämentyyppisiä tietokoneavusteisia kieli- ja käsiteanalyyssejä. Jopa paradoksaaliselta vaikuttavalla tavalla tietokoneiden käyttö tähän tapaan auttaisi siis välttämään nyky-yhteiskunnan kritisoitua asiantuntijavaltaa tekemällä yhteisistä asioista sopimisesta kielellisesti läpinäkyvämpää ja tilannekohtaisesti hienovaraisempaa. Samalla ihmisten oikeustaju saisi kanavan tulla kuulluksi nykyistä paljon paremmin, kun kone auttaisi eri taustoista tulevia ihmisiä toimimalla heidän merkitysneuvotteluidensa apuvälineenä.

Koneiden käyttö edellä kuvatulla tavalla oikeuden välineenä on kaukana niistä visioista, joita esimerkiksi 1980-luvulla tarjottiin. Ajateltiin, että lakeja voitaisiin kirjata sääntöjärjestelminä, joita voidaan sitten soveltaa kuin logiikan lakeja seuraten. Tuollaista tavoitetta voidaan pitää epärealistisena, kun otetaan huomioon kielen ja maailman monitahoisuus ja -mutkaisuus, ja mitä ollaan opittu kielen rajoista tiedon esittämisessä.

Koulutuksen ja tieteen käytännöt ja infrastruktuurit

Nykyaikainen tutkija käyttää työssään hakukoneita, kirjoittaa artikkelinsa tietokoneella, ne leviävät internetin kautta ja tutkijat

keskustelevat työnsä sisällöstä erilaisten tietoverkkojen varaan rakennettujen palvelujen kuten sosiaalisen median kautta. Useilla aloilla tietokone on myös osa tutkimusmenetelmistöä, esimerkkeinä tilastolliset tietojenkäsittelyohjelmat ja laadullista tutkimusta tukevat ohjelmistot. Kouluissa samantyyppiset prosessit toistuvat erilaisilla painotuksilla. Voidaan esittää kysymys, missä määrin tekoälytutkimus on ollut mukana tässä kehityksessä. Ainakin koneoppimista ja tietämyksen esittämistä koskevalla tekoälytutkimuksella on selkeä yhteys tutkimuksen ja koulutuksen alueeseen. Konkreettista vaikutusta voidaan toistaiseksi kuitenkin pitää varsin vähäisenä. Tietämyksen esittämisen alueella on kaksi selkeää poikkeusta. Käsitekarttoja voidaan pitää tekoälytutkimuksen sivutuotteena, jota käytetään laajalti oppimisen apuvälineenä (ks. esim. Raike 2005, Koponen & Pehkonen 2008). Semanttisen webin tutkimuksessa tuotettavilla ontologioilla puolestaan pyritään tietämyksen esittämiseen systemaattisella tavalla (Mäkelä et al. 2012). Ontologioita voidaan pitää koneen tulkittavissa olevina tesauksina, jotka kuvaavat – käsitekarttojen tapaan mutta yksityiskohtaisemmin ja systemaattisemmin – valitun kohdealueen käsite-rakenteet. Ontologiat eivät kuitenkaan sellaisenaan ratkaise semioottisesti kyvykkäiden koneiden rakentamisen haastetta. Seuraavaksi voidaankin tarkastella kehitysaskeleita, jotka tullevat mahdollisiksi tulevaisuudessa, jos aiemmin kuvatun kaltaiset semioottiset koneet tulevat mahdolliseksi.

Semioottisesti kyvykäs kone voisi auttaa ketä tahansa ihmistä kulkemaan eteenpäin lähikehityksen vyöhykkeellään (Vygotski 1982). Lähikehityksen vyöhyke on se osaamisen alue, joka on ihmisen seuraavaksi saavutettavissa, jos hän saa siihen tukea. Jo nykyaikana on mahdollista rakentaa järjestelmiä, jotka analysoivat tekstien vaikeustasoa suhteessa lukijan tietämyksen tasoon tekstin alueella (Paukkeri et al. 2012). Lukijan tietämystä arvioidaan analysoimalla hänen kirjoittamiaan tekstejä, joita täytyy olla riittävästi saatavilla kattavan mallin aikaansaamiseksi. Askel tästä eteenpäin on laittaa kone suunnittelemaan, mitä kannattaa seuraavaksi lukea, jotta oma osaaminen edistyy haluttuun suuntaan. Tällaisessa skenaar-

riossa tietämistä tarkastellaan käsitteellisten kokonaisuuksien tasolla eikä esimerkiksi yksittäisten faktojen muistamisena. Tämä on lähtökohta myös konstruktivistisesta oppimisesta (Bruner 1961, Glasersfeld 1989, von Foerster 2002, Hakkarainen et al. 2004) ja käsitteellistä muutosta (Kuhn 1962, Vosniadou & Brewer 1992, Hakkarainen et al. 2004) koskevassa tutkimuksessa. Käsitteellisten kokonaisuuksien prosessoinnissa on olennaista mallintaa kielellisen tason lisäksi hahmoluonteinen moniulotteinen tieto maailmasta (Gärdenfors 2000, Honkela et al. 2008). Tämä hahmoluonteinen kokemuksellinen tieto maailmasta on tyypillisesti paljon yksityiskohtaisempaa kuin mitä kielen tasolla kuvataan. Kieltä ja erilaisia luokittelujärjestelmiä onkin pidettävä pikemminkin tapana pakata tietoa ja mahdollistaa vuorovaikutusta kuin tapana kuvata maailmaa jotenkin ehdottoman totuudenmukaisesti ja jäännöksittä. Koulutukseen liittyvä esimerkki tästä ovat todistukset, joita saamme eri vaiheissa. Koulutodistus sisältää pakatussa muodossa tietoa siitä, kuinka paljon oppilaan katsotaan oppineen kultakin alueelta. Lukion päättötodistuksessa on annettu arvosana äidinkielestä ja kirjallisuudesta, toisesta kotimaisesta ja vieraista kielistä, matematiikasta, fysiikasta, kemiasta, jne. – yhteensä vajaasta kahdestakymmenestä oppiaineesta. Matemaattisesti on kyse noin 20-ulotteisesta vektorista, joka pakatussa muodossa kuvaa oppilaan saavuttamaa tiedon ja osaamisen tasoa. Tätä kuvausta voidaan pitää hyvin ylimalkaisena ja syy siihen, että kuvaus ei ole tarkempi liittyy tarpeeseen pakata tietoa toisten ihmisten luettavissa olevaan muotoon. Semioottisten kyvykkäiden koneiden aikana tämä ei ole tarpeen vaan minimiin pakatuista osaamisen kuvaamismuodoista voidaan luopua ja siirtyä tietokoneiden ylläpitämiin tuhansia tai jopa miljoonia yksityiskohtia sisältäviin tapoihin kuvata ihmisen tietoja, taitoja ja osaamista. Vaikka tähän skenaarioon liittyy myös mahdollisesti huolestuttavia tekijöitä koskien yksityisyyden suojaa ja koneiden roolia yhteiskunnassa, voidaan sitä myös pitää deliberatiivisena: ihmistä ei määriteltäisi pienen muuttujamäärän kautta vaan ihmisen tiedollinen ja osaamisen taso tuli myös virallisesti määritellyksi nykyistä paljon paremmin sellaisena huomattavan monimutkaisena ilmiönä kuin mitä

se todellisuudessa on.

Edellä kuvattuun skenaarioon liittyvä radikaali tulevaisuudenkuva on, että oppimisen ja koulutuksen määrämuotoisuudesta luovutaan. Kunkin ihmisen luontaiset kyvyt ja motivaatio muodostavat lähtökohdan, jonka voidaan antaa ohjata oppimista. Toisaalta ihmiset opiskelevat voidakseen ratkaista erilaisia käytännön ongelmia ja toimia tehtävissä, joihin heidät on valittu. Kumpi tahansa motivaatio oppimiseen ja osaamisen kehittämiseen on joka tapauksessa parempi kuin geneerinen opinto-ohjelman seuraaminen. Lönnrot lähti 26-vuotiaana ensimmäiselle runojenkeruumatkalleen. Miksi nykyäänkään opiskelijoiden pitäisi tyytyä suorittamaan määrämittäisiä kursseja ja täyttämään opettajien kansioita tenttivastauksilla ja opinnäytetöillä, joiden ainoa tarkoitus on osoittaa osaamista. Tutkivan oppivan periaatteen soveltamista voidaankin pitää kehityksenä oikeaan suuntaan (Hakkarainen et al. 2004). On perustaitoja ja -tietoja, joita tarvitaan soveltavassa ongelmanratkaisussa, mutta on hyvä, jos näiden oppimiseen on itse koetun kautta saatu motivationaalinen perusta.

Osaamisen kysynnän ja tarjonnan välistä tasapainoa voidaan tarkastella semioottisesti kyvykkäiden koneiden avulla, ja ohjata sitä nykyistä paljon hienovaraisemmin ja jopa ilman muodollista järjestelmää määrämittäisine koulutuskokonaisuuksineen ja todistuksineen. Nykyään poliitikot ovat esimerkiksi huolissaan opiskeluajoista ja muista karkeista mittareista ilman, että olisi todellista tietoa siitä, mitä opiskelijoiden elämässä todella tapahtuu. Nykyisillä mittaustavoilla opiskelija, joka perustaa oman yrityksen kolmannen opintovuotensa jälkeen ja työllistää mahdollisesti satoja ihmisiä, on tilastoissa samassa asemassa kuin opiskelija, joka on muutaman opintovuotensa jälkeen huomannut olevansa väärällä alalla, mutta joidenkin säännösten aiheuttamien ongelmien takia luopunut urasuunnitelmistaan ja suistunut alkoholismiin. Nykyisin vain määrämuotoiset suoriutumiset yhteiskunnan määrittelemissä järjestelmissä lasketaan, mitä voidaan itse asiassa pitää rakenteellisena väkivaltana yksilöä kohtaan. Syynä tähän voidaan yhtäältä pitää kielen ja

luokittelujärjestelmien sisäänrakennettuja ominaisuuksia (Bowker & Star 2005) ja toisaalta halua kontrolloida yhteiskuntaa tavalla, joka rajoittaa vastuullisten ja itsenäisten kansalaisten yhteiskunnan kehittymistä.

Elinikäinen oppiminen voisi siis olla paljon konkreettisemmin todellisuutta ilman selvää jakoa kouluun, opiskeluun ja työelämään ja ilman erilaisia portaita tutkintoineen. Jokaisella ihmisellä on oma osaamisprofiilinsa, jota hänellä on mahdollisuus kehittää haluamaansa suuntaan hyvin monenlaisilla tavoilla. Automaation jatkaessa tunkeutumistaan erilaisille inhimillisen toiminnan osa-alueille on epätodennäköistä, että tulevaisuuden osaamistarpeet olisivat kovin määrämuotoisia ja -mittaisia. Näin ollen ihmisen ”rehevä kasvu” sisäisen motivaation pohjalta lienee järkevämpi suunta kuin ulkoa ohjattu oppimisen sääntely, joka väkisinkin perustuu hyvin puutteelliseen tietoon tulevaisuudesta. Lienee kestävä kehitystä edistävää auttaa ihmisiä nuoresta pitäen itsenäiseen oppimiseen ja tekemiseen, joka mahdollistaa myöhemmin omaehtoisen selviämisen tilanteissa, jotka voivat poiketa paljonkin aiemmista. Tällainen kasvatus ja koulutus vaatii rohkeutta antaa vastuuta ja valtaa sekä toimia kaavamaisten mallien tarjoaman puitteen ulkopuolella (Saarinen 2008).

Määrämittaisten vaatimusten asettamisen sijasta ohjaamisen ja oppimisen arvioinnin kohteena voisi tulevaisuudessa olla motivaatiotaso ja edistyminen omaehtoisemmin valituilla poluilla. Määrätietoista tavoitteenasettelua voidaan tukea jo varhain ja tällöin myös erilaiset välineelliset sisällöt kuten matematiikka ja kielet tulevat luonnostaan mukaan oppimisen kohteiksi. Osaamisen ja talouden periaatteita voitaneen tarkastella analogisina ilmiöinä. On selvää, että kaikki eivät voi olla yhtä varakkaita, mutta toisaalta liian epätasaista varakkuuden jakaumaa voidaan pitää ongelmallisena. Samantyyppinen tarkastelu koskee osaamista, jossa kaikkein eniten osaavia ei tarvitse ”verottaa” estämällä heitä oppimasta kuinka paljon tahansa vaan tärkeitä on pitää huolta siitä, että osaamista löytyy riittävästi koko populaatiosta, mihin Suomessa on Pisa-tutkimusten mukaan päästy erittäin hyvin.

Semioottisesti kyvykkäiden koneiden aika tulee

olemaan myös emotionaalisesti ja sosiaalisesti kyvykkäiden koneiden aikaa. Tällöin kone voi olla opettajan yhteistyökumppani esimerkiksi näyttämässä, minkälaiset tekijät ovat estämässä oppimista ja kehittymistä kohti vahvenevaa uskoa omiin kykyihinsä. Tällaisessa skenaariossa arvioinnin kohteena eivät olisi ensijaisesti oppimistulokset vaan tapa, jolla oppimista tukevaa puitetta ylläpidetään. Sosiaalisesti ja semioottisesti kyvykäs kone voisi kiinnittää huomiota oppilaiden väliseen kiusaamiseen tai opettajan kohdistamaan lannistavaan vuorovaikutukseen jotakin oppilasta kohtaan. Tällaisessa utopiassa kone ei olisi määräysvaltaa käyttävä ”big brother” vaan arvovaltaa omaava ”viisas kone”, joka voisi interventioillaan rakentaa toimivaa yhteisöä paikoissa ja tilanteissa, joihin ei riitä vastaavia taitoja omaavia ihmisiä. Samalla kun edellä kuvattu skenaario oppimisen yksilöllistymisestä painottaa yksilökeskeisyyttä, on ilmeistä, että yhteisöllisyyttä edistäviä voimia on syytä olla mukana tasapainottamassa kehitystä. Tavoitteena voi siis myös olla sosiaalisten taitojen oppiminen – ei hyväksikäyttävässä mielessä vaan kehittävänä empaattisuutena ja vastuullisuutena. Se, mitä koulussa arvioidaan yhden arvosanan – käyttäytyminen – puitteissa, voidaan jakaa kymmeniin osa-alueisiin, joita koskevaa tietoa voidaan kerätä tuhansina yksityiskohtina.

Tieteen käytäntöjen kehittämisessä semioottisesti kyvykkäillä voi olla monia rooleja. Esimerkkejä ovat tutkimussuunnitelmien ja -artikkeleiden laadun arviointi sekä tieteellisen tiedon esittäminen ja jatkohyödyntäminen. Tutkimuksessa käytetään paljon vertaisarviointia, johon liittyvä ongelma on arvioijien valinta. Arvioijien valinnassa semioottisesti kyvykkäillä koneilla voisi olla paljon annettavaa. Artikkeleiden ja suunnitelmien sisältöjen sekä arvioijapoolissa mukana olevien asiantuntijoiden osaamisprofiilien välistä suhdetta voidaan tarkastella hyvin yksityiskohtaisella tasolla ja tehdä siten nykyistä huolellisempia valintoja. Myös itse arviointiprosessissa koneet voivat olla lisääntyvässä määrin mukana. Nykyisin koneellista analyysia käytetään lähinnä plagioinnin tunnistamiseksi.

Tulevaisuudessa koneen rooli tiedon tuottamisen

prosesseissa voi olla nykyistä paljon laajempi. Tiedon esittämisessä on nähtävissä kaksi tulevaisuuden mahdollisuutta. Ensinnäkin semioottisesti kyvykkäiden koneiden ansiosta tarve yrittää harmonisoida tiedon esitysmuotoja vähenee. Se lisäarvo, jota nykyään tavoitellaan avoimen yhdistetyn tiedon (open linked data) avulla, saavutetaan nykyistä paljon helpommin. Kone löytää asioiden välisiä yhteyksiä siitä huolimatta, miten ne on esitetty. Tämä näkökohta on syytä ottaa huomioon jo nykyään, kun suunnitellaan tieteen infrastruktuureja. Toinen tärkeä kehitysskaskel on tiedon esittäminen laajenevassa määrin kontekstuaalisena. Matemaattiset ja luonnontieteet ovat eräässä mielessä helppoja tieteitä, koska niissä kertyvä tieto ei yleensä riipu ajasta eikä paikasta. Humanistisissa ja sosiaalitieteissä tieto on usein paikkaan ja aikaan sidottua, mikä tekee tutkimuksesta haasteellista (von Foerster 2002, Honkela 2011). Tilannetta helpottanee se, että kontekstisidonnaisen tiedon esittämiseen on kehittymässä aiempaa parempia tapoja. Kvantitatiivinen esimerkki tästä on ehdollisten todennäköisyyksien hyödyntäminen. Sen sijaan, että väitettäisiin jotakin maailmasta yksinkertaisesti propositiona $A, B, C, \text{ jne.},$ tietoa maailmasta kirjataan ehdollisina todennäköisyyksinä muodossa A ehdolla B eli $P(A | B)$. Esimerkiksi paras tämän hetkinen konekäännös-järjestelmä, Google Translate, perustuu satoihin miljardeihin tämäntyyppisiä tilastollisesti kerättyjä ja todennäköisyysslaskennan puitteissa kuvattuja riippuvuussuhteita (Brants et al. 2007). Ehdolliset todennäköisyydet ovat yksi tapa esittää kontekstuaalista tietoa, jonka täydennyksenä tarvitaan tapoja esittää tietoa hahmoluonteisesti (Gärdenfors 2000, Honkela et al. 2008). Tällaista lähestymistapaa tarvitaan, kun halutaan tehdä nykyistä tarkemmin päätelmiä esimerkiksi historia- ja kontekstisidonnaisista psykologisista tai yhteiskunnallisista ilmiöistä. Kvantitatiivisilla menetelmillä yksinään on vaikeaa tavoittaa monimutkaisimpia ilmiötä, joten kvantitatiivisia ja kvalitatiivisia menetelmien yhdistäminen saavuttaa suosiota (ks. esim. Janasik et al. 2009).

Kontekstualisoivaa tieteen tulosten tekstilouhintaa tehdään jo nykyisellään aivotutkimuksen alueella. Käymällä läpi suuri joukko raportteja aivokuvantamistuloksista,

voidaan mallintaa miten kognitiiviset toiminnot sijoittuvat aivojen eri osiin (ks. esim. Wilkowski et al. 2010). Pitkälle menevien johtopäätösten tekemisessä siitä, mitä aivojen eri osien havaitaan tekevän, täytyy tietysti olla varovainen, koska aivot eivät toimi erillään muusta maailmasta vaan vuorovaikutuksessa muun kehon ja maailman kanssa. Kohosen itseorganisoiva kartta selittää abstraktilla tasolla, miten toimintojen paikallisuus ja kokonaisvaltaisuus voidaan ymmärtää (Kohonen 2001). Mielenkiintoisena voidaan pitää aivotutkimuksen tuloksia, joissa on voitu varsin hyvällä tarkkuudella päätellä aivokuvantamisen avulla, minkä sanaluokan sanaa koehenkilölle kulloinkin näytetään (Fyshe et al. 2012). Maguire et al. (2000) osoittivat, minkälaisia aivokuvantamisessa havaittavia spatiaalista hahmottamiskykyä koskevia muutoksia taksikuskiensa aivoissa on verrattuna muuhun väestöön. Näistä tuloksista voidaan ennakoida tulevaisuuden tilanne, jossa aivokuvantamisen perusteella voidaan tehdä varsin tarkkojakin päätelmiä ihmisen osaamisalueista, vaikkei yksittäisten ajatuksien lukeminen tulisikaan mahdolliseksi.

Tietokone ihmisten ja ympäristön hyvinvoinnin edistäjänä

Edellä on kuvattu sitä, kuinka ihmisen tietämisestä ja osaamisesta on mahdollista kerätä nykyistä paljon yksityiskohtaisempaa tietoa. Sama koskee myös ihmisten arjen tekemisiä ja olotilaa kulloisenakin ajanhetkenä. Kuumemittari on pitkään käytössä ollut teknologiaa oman tilan mittaamiseen poikkeustilanteissa. Nykyaikaisia välineitä ovat esimerkiksi sykemittarit monine toimintoineen sekä sähköiset päiväkirjat esimerkiksi urheilusuoritusten, syötyjen ruokien ja painon kirjaamiseen. Tällaisten ilmiöiden tarkasteluun käytäntöjen teoria on erityisen hedelmällinen. Se tekee erottelun käytäntöjen idean ja niiden liittyvän osaamisen ja materiaalien puitteiden välillä (Shove et al. 2012).

Äärimmilleen arjen mittaaminen on viety quantified self -kulttuurissa, joka on nimettyä ilmiönä niin tuore, että sille ei ole vakiintunut suomenkielistä nimitystä eikä siinä ole juuri kirjoitettu tieteellisiä artikkeleita. Yksi ehdokas suomennokseksi on "mitattu minä", jolla viitataan

siihen, että omasta elämästä kerätään laajasti määrämutoista tietoa, jota voidaan analysoida eri tavoilla. Mitattu minä on alakulttuuri, mutta piirteitä siitä siirtyy jokaisen arkeen esimerkiksi mobiiliteknologian kautta. Mobiiliteknologialla toteutettavat palvelut ovat joissakin tapauksissa siinä määrin kiinnostavia tai hyödyllisiä, että ihmiset päättävät antaa itsestään tietoa palveluntarjoajalle yksityisyydensuojan kustannuksellakin.

Arjen seurantaan ja mittaamiseen liittyvillä sovelluksilla on monia mahdollisia yhteyksiä semioottisesti kyvykkäiden koneiden tulevaisuuteen. Yksi niistä on kontekstitiedon keruu. Jos mukana kulkeva laite tallentaa sekä kuvallista että ääni-informaatiota, semioottisesti kyvykäs kone voi kehittyä ymmärtämään omistajaansa kontekstuaalisesti. Tällöin käytettävissä olisi henkilökohtainen assistentti, joka pystyisi esimerkiksi toteuttamaan komentojen mukaisia toimintoja asiayhteyden edellyttämällä tavalla. Fyysisen ja henkisen hyvinvoinnin edistäjänä tällaisilla järjestelmillä voi olla – niin halutessamme – valtava merkitys. Nykyistä terveyttä ja hyvinvointia koskevaa tutkimusta rajoittaa se, että tietoa on suhteellisesti ottaen varsin rajatusti. Jos niitä tekijöitä, jotka vaikuttavat hyvinvointiin voitaisiin seurata kattavasti, johtopäätöksiä kussakin tilanteessa ja kullekin ihmiselle sopivista interventioista voitaisiin tehdä paljon nykyistä tarkemmin, kontekstiherkemmin ja yksilöllisemmin. Tähänkin kehitykseen vaikuttavat päätökset yksityisyydensuojasta, mutta toisaalta jo nykyisellään monet ihmiset ovat päättäneet antaa terveystietojaan laajasti esimerkiksi Cure Together -nimisen palvelun ylläpitäjille (curetogether.com). Palvelussa käyttäjät voivat kirjata, mitä hyötyä heille on ollut erilaisista interventioista heidän sairauksissaan. Yksi palvelun etu on se, että interventioiden joukko on avoin eli mukana voi olla tavanomaisten lääketieteellisten hoitotoimenpiteiden lisäksi esimerkiksi elämäntapamuutoksia tai vaihtoehtoisia hoitomuotoja. Tällöin näyttöön perustuva lääketiede kattaa sellaisetkin lähestymistavat, jotka eivät ehkä pääse läpi lääketieteen tutkijoiden seulasta. Mainitussa Cure Together -palvelussa on muun muassa todettu, että liikunta on jopa paras hoitokeino lievässä tai

keskivaikeassa masennuksessa. Ihmisten hyvinvointi ja toimintakyky on sekä yksilön että yhteiskunnan kannalta tärkeä asia, mutta yhden liiketoiminta-alueen kannalta ei välttämättä ole suotuisaa, että jotakin ongelmaa hoidetaan lääkevalmisteiden sijasta ottamalla yhteyttä vaikkapa Vierumäen urheiluopistoon liikunnallisen elämänvalmennuksen saamiseksi. Kokonaisuuden kannalta on kuitenkin tärkeää, että tällaiset toimialojen väliset painopisteiden muutokset voivat tapahtua silloin, kun ne ovat kokonaisuuden kannalta hyödyllisiä. Vastaavalla tavalla tarkastelulla voidaan ihmetellä, onko esimerkiksi tupakanviljelijöille annettavissa tukiaisissa järkeä vain työllisyyden ylläpitämisen takia.

Jos ihminen on kerännyt laajasti tietoja itsestään mitattu minä -periaatteen mukaisesti, on hyvinkin mahdollista, että erilaisiin tilanteisiin löytyy ratkaisukeinoja paljon tehokkaammin kuin, jos tällaisia tietoja ei ole kerätty. Tulevaisuuden skenaariona voi myös ajatella, että mukana kulkevan laitteen keräämä tieto antaa semioottisesti kyvykkäälle koneelle mahdollisuuden toimia avustajana vuorovaikutuksessa lääkäreiden ja lääkäreiden kanssa yhteistyössä toimivien, lääke- ja hoitotieteisiin keskittyneiden koneiden kanssa. Tässä yhteydessä on tärkeää huomata yksi seikka semioottisesti kyvykkäistä koneista. Eri ihmisten aivot eivät muodosta yhdessä yhtä keskitettyä suuraivoa. Samaan tapaan joukko semioottisesti kyvykkäitä koneita ei muodosta yhtä suurta älykästä konetta vaan niiden vuorovaikutusta koskevat lähes samantyyppiset lainalaisuudet kuin ihmisiäkin. Periaatteessa on toki mahdollista, että jotakin henkilöä edustava laite antaa vaikkapa hoitokoneelle tiedoksi kaiken sen kokemustiedon, jota henkilö on kerännyt elämänsä aikana ja näin mahdollistaa tuosta kokemuksesta oppimisen. Lisäksi on oletettava, että laskentakapasiteettia on niin paljon käytettävissä, että kokemuksesta oppiminen voidaan puristaa lyhyeen aikaan, tai että kokemusvirrasta voidaan valita relevantti osajoukko. Jos näin kyvykkäiden koneiden aika koittaa, jää tietysti monenlaisia kysymyksiä vastattavaksi. Miten ihminen toimii, jos mukana on kone, joka osaa kertoa, mitä olisi syötävä, milloin nukuttava, keiden kanssa vietettävä aikaa, ja vastauksia moniin muihinkin kysymyksiin.

Ihminen voisi tällaisessa tilanteessa taantua lapsen asemaan tottelemaan jatkuvasti viisaan vanhemman antamia neuvoja, tai ehkä kieltäytyisi tällaisesta palvelusta kokonaan. Toisaalta jo nyt tiedämme monista tavoista niiden haitat ja riskit, ja yhteiskuntamme säättävät lakeja ja muodostavat normeja, joissa näitä tapoja säädellään, eli tätä rajankäyntiä yksilön vapauden ja yhteiskunnan ohjauksen välillä käydään jatkuvasti.

Jos viisas kone on henkilökohtainen, suora yhteiskunnallinen ohjaus ja valvonta tuon koneen kautta ei toteudu, ja ihminen voi periaatteessa päättää nykyiseen tapaan tekemisistään. Ihmisyhteisön hajautetun luonteen vuoksi ja tietäen evolutiivisten ja itseorganisoituvien järjestelmien voiman, on todennäköistä, että hajautettuun päätöksentekoon perustuva malli on se, mihin edelläkuvattujen teknologisten mahdollisuuksien mahdollistama kehityskulku johtaisi. Semioottisesti älykkäät koneet voisivat sitä paitsi auttaa ihmisiä muodostamaan nykyistä voimakkaampia luottamusverkostoja johtuen merkitysneuvottelukyvyyistään, jotka tukevat yhteisymmärryksen syntymistä erilaisissa tilanteissa.

Ihmisten hyvinvoinnin tarkastelu on riittämätöntä, jos huomioon ei samalla oteta ympäristön hyvinvointia. Ihminen on suoraan riippuvainen fyysisestä ja biologisesta ympäristöstään, joita koskeva kestävä kehitys on ihmisen tulevaisuuden kannalta keskeistä. Kestävää kehitystä estäviä tekijöitä ovat esimerkiksi saastuminen ja muut tekijät, jotka muuttavat tai ovat vaarassa muuttaa ekosysteemejä radikaaleilla tavoilla. Semioottisesti kyvykkäät koneet voivat edistää kestävä kehitystä eri tavoilla. Niiden avulla voidaan kerätä ympäristön tilaa koskevaa tietoa arkisten havaintojen kautta (ks. esim. Wynne 2004). Ne voivat osallistua ympäristön tilaa koskevien tilannearvioiden ja skenaarioiden tekemiseen aiemmin tässä artikkelissa kuvatun FuturICT-hankkeessa suunnitellun järjestelmän puitteissa. Ne voivat myös toimia tulkkeina arjessa ja korkeammankin tason päätöksentekotilanteissa ja auttaa ihmisiä tekemään viisaita valintoja. Ympäristötutkimuksessa ei riitä se, että tarkastellaan ympäristöä esimerkiksi biologian, kemian ja fysiikan näkökulmasta, vaan kyse on pitkälti siitä,

miten ihminen vaikuttaa ympäristöön ja silloin tarkastelun kohteena täytyy olla esimerkiksi myös käsitteellisiä ja sosiokognitiivisia kysymyksiä (Hukkinen 2008).

Johtopäätöksenä voidaan todeta, että kehittynyttä tekoälyä on mahdollista hyödyntää monenlaisissa yksittäisiä ihmisiä ja kokonaisia yhteiskuntia palvelevissa tarkoituksissa, joista monet ovat nykyisen teknologioiden kehittymistä koskevien keskusteluiden ulkopuolella tai joiden suhteen päähuomio on keskittynyt vaaroihin mahdollisuuksien sijasta. Keskustelua siitä, minkälainen rooli tässä kirjoituksessa ennakoituille semioottisesti kyvykkäille koneille annetaan, on syytä jatkaa. Kuvilteltavissa on kuitenkin tilanne, jossa tämä kehitys johtaa hyvinkin suotuisiin lopputuloksiin ihmiskunnan ja koko biosfäärin kannalta – aikana, jolloin palveluksessamme ovat viisaat koneet.

Lähdeviitteet

Bishop, Steven, Dirk Helbing, Paul Lukowicz & Rosaria Conte (2011): FuturICT: FET Flagship Pilot Project. *Procedia Computer Science*. <http://ssrn.com/abstract=1895523>

Bowker, Geoffrey & Susan Leigh Star (2000): *Sorting Things Out: Classification and Its Consequences*. MIT Press.

Brants, Thorsten, Ashok C. Papat, Peng Xu, Franz J. Och & Jeffrey Dean (2007). *Large Language Models in Machine Translation*. *Proceedings of EMNLP-CoNLL 2007*, pp. 858-867.

Brunak, S. & B. Lautrup (1990): *Neural networks: computers with intuition*. World Scientific Publishing Co., Inc., River Edge, NJ.

Bruner, J. S. (1961): *The act of discovery*. *Harvard Educational Review*, 31(1), s. 21–32.

Bush, Vannevar (1945): *As we may think*. *Atlantic Monthly*, July.

Deerwester, Scott, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, Richard Harshman (1990): *Indexing by Latent Semantic Analysis*. *Journal of the American Society for Information Science* 41(6), s. 391–407.

Dijksterhuis, A., M.W. Bos, L.F. Nordgren & R.B. van Baaren (2006): Complex choices better made unconsciously? *Science* 313, s. 760-761.

Fodor, J.A. (1975): *The Language of Thought*. Harvard University Press, Cambridge, Massachusetts.

Fyshe, A., E.B. Fox, D.B. Dunson & T.M. Mitchell (2012): Hierarchical Latent Dictionaries for Models of Brain Activation. *Proceedings of AISTATS 2012*, April.

Foerster, Heinz von (2002): *Understanding understanding: conversations on epistemology and ethics*. Springer.

Glaserfeld, E. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, 80(1), s. 121-140.

Gärdenfors, Peter (2000): *Conceptual Spaces*. MIT Press.

Hakkarainen, Kai, Kirsti Lonka & Lasse Lipponen (2004): *Tutkiva oppiminen - Järki, tunteet ja kulttuuri oppimisen sytyttäjinä*. WS Bookwell, Porvoo.

Hakkarainen, Kai, Ritva Engeström, Sami Paavola, Pasi Pohjola & Timo Honkela (2009): Knowledge practices, epistemic technologies, and pragmatic web. Teoksessa *Proceedings of I-KNOW'09 and I-SEMANTICS'09: the 4th AIS SigPrag International Pragmatic Web Conference Track (ICPW 2009)*, 683–694. Verlag der Technischen Universität Graz, Graz, Itävalta.

Harnad, Stevan (1990): The Symbol Grounding Problem. *Physica*, D42: 335-346.

Heimbürger, Anneli & Taru Kuhanen (1983): *Hypermedia. Tekoälyn ensyklopedia. Suomen tekoälyseura r.y.* s. 327 - 333.

Honkela, Timo (1993): Neural nets that discuss: a general model of communication based on self-organizing maps. *Proceedings of ICANN'93, International Conference on Artificial Neural Networks*, Springer-Verlag, London, s. 408–411.

Honkela, Timo, Ville Könönen, Tiina Lindh-Knuutila, and Mari-Sanna Paukkeri (2008): Simulating processes of concept formation and communication. *Journal of Economic*

Methodology 15(3), s. 245-259.

Honkela, Timo (2010): Tietokone tallentajasta tulkitsijaksi: Representaatio tietojenkäsittelytieteessä. Teoksessa Tarja Knuutila & Aki Petteri Lehtinen (toim.): *Representaatio - Tiedon kivijalasta tieteiden työkaluksi*, 305-330. Gaudeamus, Helsinki.

Honkela, Timo (2011): Kunnioitusta erilaisuutta kohtaan. *Tieteessä tapahtuu*, 29(4-5): s. 38-39. <http://ojs.tsv.fi/index.php/tt/article/view/4260/3974>

Honkela, Timo, Juha Raitio, Krista Lagus, Ilari T. Nieminen, Nina Honkela and Mika Pantzar (2012): Subjects on Objects in Contexts: Using GICA Method to Quantify Epistemological Subjectivity. *Proceedings of IJCNN 2012 Conference*, painossa.

Hukkinen, J. (2008): *Sustainability Networks: Cognitive Tools for Expert Collaboration in Social-Ecological Systems*. Routledge, London.

Huopaniemi, Ilkka, Tommi Suvitaival, Janne Nikkilä, Matej Orešic, and Samuel Kaski (2010): Multivariate multi-way analysis of multi-source data. *Bioinformatics*, 26, s. i391–i398.

Hyvärinen, Aapo, Juha Karhunen & Erkki Oja (2001): *Independent Component Analysis*. John Wiley & Sons.

Hyvärinen, Aapo, Jarmo Hurri & Patrik O. Hoyer (2009): *Natural Image Statistics - A probabilistic approach to early computational vision*. Springer.

Hyvönen, Eero, Ilkka Karanta & Markku Syrjänen (toim.) (1983): *Tekoälyn ensyklopedia*. Gaudeamus, Helsinki.

Janasik, Nina, Timo Honkela & Henrik Bruun (2009): Text mining in qualitative research: Application of an unsupervised learning method. *Organizational Research Methods* 12(3), s. 436-460.

Knuutila, Tarja & Timo Honkela (2005): *Questioning External and Internal Representation: The Case of Scientific Models*. Teoksessa Magnani, L. (toim.) *Computing, Philosophy, and Cognition*. King's College Publishing, Lontoo.

Kohonen, Teuvo (1984): *Self-Organization and Associative Memory*. Springer, Berlin.

- Kohonen, Teuvo (2001): *Self-Organizing Maps*. Third, extended edition. Springer.
- Koivunen, Hannele (1997): *Hiljainen tieto*. Otava, Helsinki.
- Koponen, Ismo & Maija. Pehkonen (2008): Physics concepts and laws as network-structures: Comparisons of Structural features in Experts' and Novices' concept maps. Teoksessa *Proceedings of 3rd Conference on Concept Mapping*. Julkaisut ladattavissa osoitteessa <http://cmc.ihmc.us/cmc2008/cmc2008Program.html>
- Kuhn, Thomas S. (1962): *The structure of scientific revolutions*. University of Chicago Press, Chicago, Illinois.
- Lagus, Krista & Timo Honkela (2006): Kognitiivinen mallintaminen. Teoksessa H. Hämäläinen, M. Laine, O. Aaltonen & A. Revonsuo (toim.): *Mieli ja aivot*. Kognitiivisen neurotieteen oppikirja, 61-70. Kognitiivisen neurotieteen tutkimuskeskus, Turun yliopisto, Turku.
- Lindh-Knuutila, Tiina, Timo Honkela & Krista Lagus (2006): Simulating meaning negotiation using observational language games. Teoksessa P. Vogt et al. (toim.) *Symbol Grounding and Beyond: Proceedings of the Third International Workshop on the Emergence and Evolution of Linguistic Communication*, 168-179, Springer, Berlin/Heidelberg.
- Linturi, Risto (2010): Social simulation of networked barter economy with emergent money. Technical Report TKK-ICS-R40, Aalto University School of Science and Technology, Department of Information and Computer Science, Espoo.
- Maguire, E. A., D. G. Gadian, I. S. Johnsrude, C. D. Good, J. Ashburner, R. S. J. Frackowiak & C. D. Frith, (2000): Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, 97(8), s. 4398-4403. The National Academy of Sciences.
- Mehmood, Yasir, Mudassar Abbas, Xi Chen & Timo Honkela (2011): Self-organizing maps of nutrition, lifestyle and health situation in the world. Teoksessa *Advances in Self-Organizing Maps - Proceedings of WSOM 2011*, 8th International Workshop, 160–167. Springer.
- Mitchell T. M., S.V. Shinkareva, A. Carlson, K. Chang, V.L. Malave, R.A. Mason & M.A. Just (2008): Predicting human brain activity associated with the meanings of nouns. *Science* 320, 1191-1195
- Mustajoki, Arto (2008): Modelling of (mis)communication. *Prikladna lingvistika ta ligvistitshni tehnologii: Megaling-2007*, s. 250-267. Dovira, Kiev.
- Mustajoki, Arto (2012): A multidisciplinary and multidimensional approach to risks and causes of miscommunication. *Language and Dialogue*, hyväksytty julkaistavaksi.
- Mäkelä, Eetu, Eero Hyvönen & Tuukka Ruotsalo (2012): How to deal with massively heterogeneous cultural heritage data – lessons learned in CultureSampo. *Semantic Web – Interoperability, Usability, Applicability*, 3(1), s. 85-109
- Oja, Erkki (2002): Unsupervised learning in neural computation. *Theoretical Computer Science* 287(1), s. 187-207.
- Onnela, J.-P., J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész & A.-L. Barabási (2007): Structure and tie strengths in mobile communication networks. *PNAS* 104(18), s. 7332-7336.
- Paukkeri, Mari-Sanna, Marja Ollikainen & Timo Honkela (2012): Assessing user-specific difficulty of documents. *Information Processing & Management*, hyväksytty julkaistavaksi.
- Peltonen, Jaakko, Yusuf Yaslan & Samuel Kaski (2010): Relevant subtask learning by constrained mixture models. *Intelligent Data Analysis* 14(6), s. 641-662.
- Pereira, F., Detre, G. & Botvinick, M. (2011): Generating text from functional brain images. *Frontiers in Human Neuroscience*, 5, 72.
- Polanyi, M. (1967): *The Tacit Dimension*. Routledge & Kegan Paul, London.
- Raike, Antti (2005): *Löytäjät - Elokuvantajua rakentamassa*. Väitöskirja, Taideteollinen korkeakoulu. <http://www.uih.fi/ISBN/951-558->

- Ritter, Helge and Teuvo Kohonen (1989): Self-Organizing Semantic Maps. *Biological Cybernetics*, 61, s. 241-254.
- Saarinén, Esa (2008): Philosophy for Managers: Reflections of a Practitioner. *Philosophy of Management*, Vol. 7, Supplement.
- Searle, John R. (1980): Minds, brains, and programs. *Behavioral and Brain Sciences* 3(3), s. 417-457.
- Segall, Marshall H., Donald T. Campbell & Melville J. Herskovits (1963): Cultural Differences in the Perception of Geometric Illusions. *Science* 139(3556), s. 769-771.
- Shove, Elizabeth, Mika Pantzar & Matt Watson (2012): *The Dynamics of Social Practice – Everyday Life and how it Changes*. Sage Publications, New York.
- Steels, L. & F. Kaplan (2000): AIBO's first words: The social learning of language and meaning. *Evolution of Communication* 4(1), s. 3-32.
- Theodoridis, S. & K. Koutroumbas (2009): *Pattern Recognition* (4th edition), Elsevier.
- Vosniadou, S. & W.F. Brewer (1992): Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, s. 535-585.
- Vygotski, L.S. (1982; alkuperäinen 1962): *Ajattelu ja kieli*. Weilin&Göös, Espoo.
- Wilkowski, Bartłomiej, Marcin Szewczyk, Peter Mondrup Rasmussen, Lars Kai Hansen and Finn Årup Nielsen (2010): *BredeQuery: Coordinate-Based Meta-analytic Search of Neuroscientific Literature from the SPM Environment*. *Biomedical Engineering Systems and Technologies*, 52, Springer-Verlag, Berlin Heidelberg, s. 314-324.
- Wynne, B. E. (2004): *May the sheep safely graze? A reflexive view of the expert-lay knowledge divide. Risk, environment and modernity: towards a new ecology*. Sage.

Semiotically competent machines and their applications

Timo Honkela

Futura, 2/2012

Abstract

In this article, the future of artificial intelligence is considered from an interdisciplinary point of view. A scenario in which machines that can learn to understand in a contextually appropriate manner is presented. These kinds of machines are called to be semiotically competent. This includes an ability to learn and refine multimodally grounded and contextually relevant conceptual systems and to conduct meaning negotiations. Statistical machine learning, pattern recognition and probabilistic modeling form a methodological basis for such developments. Grounded Intersubjective Concept Analysis (GICA) is presented as an example of methodology that paves way to the presented scenario. The GICA method can be used to analyze similarities and differences in understanding regarding different individuals and contexts. Regarding these kinds of developments in artificial intelligence, different potential applications in and implications for education, science and society in general as well as developments in the health and wellbeing area and environmental domain are discussed. The main conclusion is that there are many possibilities for positive outcomes and deliberating consequences, opposite to the dark views often presented in association with intelligent technologies.