

Epäluottamuksen jatkumot ihmisen ja tekoälyn välisissä sovelluskeskusteluissa

Tarkastelemme tässä esitelmässä ihmisen ja generatiivisen tekoälyn välistä vuorovaikutusta poliittisessa keskustelussa. Demokraattisen yhteiskunnan voi katsoa nojaavan yksilön vapauteen valita edustajansa yhteiskunnalliseen päätöksentekoon oikean tiedon perusteella. Kansalaisen ja poliitikon kohtaamista värittää näin lähtökohtaisesti poliitikon moraalinen selontekovelvollisuus äänestäjilleen, joilta mandaatti yhteiskunnalliseen vallankäyttöön on saatu (ks. Berg 2003).

Tekoälyteknologian nopea kehitys on kuitenkin muuttanut poliittista toimintaa muiden instituutioiden tapaan jopa siinä määrin, että poliittisia ihmistoimijoita voivat julkisuudessa edustaa heidän – joko legitiimisti tai tahallisenä väärennöksenä – luodut tekoälyklooninsa.

Esitelmässä pyrimme valottamaan tällaisten uudenlaisten, ei-inhimillisiä olioita käsittävien poliittisten kohtaamisten vuorovaikutuksellista kompleksisuutta. Aineistomme tulee tutkimuskokeilusta, jossa yleisölle avattiin mahdollisuus keskustella vuoden 2024 Suomen presidentinvaaliehtokkaille luotujen tekoälyversioiden kanssa kokeilua varten kehitetyssä digitaalisessa sovelluksessa. Tekoälyjen koulutusmateriaalina käytettiin erityisesti verkossa ollutta vaalikonetta, ja tekoälyehtokkaita ohjeistettiin keskustelemaan tavoilla, jotka mahdollistavat äänestäjien kannatuksen ja vaalivoiton. Sovellus oli yleisölle auki reilun viikon, minä aikana siellä saattoi käydä keskustelemassa valitsemiensa chatbot-poliitikkojen kanssa. Sovelluskeskustelut muodostivat näin omanlaisensa digitaalisen vuorovaikutustilan, jonka teknologisiin resursseihin lukeutui myös kohtaamisen toinen osapuoli. Samalla siinä materialisoitui kuitenkin myös muita sovelluskeskustelun ulkopuolisia tiloja ja aikoja, jotka liittyivät paitsi politiikan myös ohjelmistokehityksen ja tutkimuksen konteksteihin.

Tutkimme digitaalista keskusteluanalyysia soveltaen (ks. esim. Vepsäläinen 2024) aineistollemme tyypillistä kysymys-vastaus-sekvenssiä, jossa sovelluksen käyttäjä esittää ensin tekoälyehtokkaalle kysymyksen ja korjaa tai haastaa sitten saamansa vastauksen. Analysoimme vastausta korjaavien ja haastavien viestien kielellisiä keinoja, kuten uudelleenmuotoiluja (*eli sinä haluat --*), vaihtoehtokysymyksiä (*kyllä vai ei*), kontrastoivia referointeja (*mutta eilen sanoit, että --*) tai evaluointeja (*huh huh mitä paskapuhetta*) sekä tapaa, jolla erilaiset muotoilut nostavat käsittelyyn luottamukseen liittyviä ongelmia (ks. Lindholm 2016). Tuloksemme osoittavat, että

käyttäjän korjaavat ja haastavat viestit voivat paikantaa tekoälyehdokkaaseen kohdistuvan epäluottamuksen erilaisille jatkumoille, minkä johdosta myös tekoälylle rakentuu sovelluskeskusteluissa erilaisia identiteettikategorioita keskustelukyvyyltään rajallisena ja etukäteen koodattuna teknologiana, kiertelyyn taipuvan poliittisen diskurssin edustajana tai kehollisen muodon ja historian omaavana poliitikkona tunnistettavine poliittisine toimineen ja pyrkimyksineen. Päättänessä pohdimme vielä tulostemme merkitystä suhteessa aiempaan ihmisen ja tekoälyn välistä vuorovaikutusta koskevaan keskusteluntutkimukseen (ks. Mlynář ym.) sekä suhteutamme niitä muilla aloilla käytyihin keskusteluihin tekoälyn toimijuudesta ja sen kerroksellisuudesta (ks. Laapotti ym. 2024).

Berg, M. 2003. *Syytöksiä ja epäilyksiä: Toimittajan ja poliitikon vuorovaikutuksesta televisiokeskustelussa*. SKS.

Laapotti, T., Eronen-Valli, M. & Raappana, M. 2024. Viestinnällinen tutkimusmatka toimijuuteen ja tekoälyyn. *Media & Viestintä*, 47(4): 142–153.

Lindholm, C. 2016. Luottamuksen haasteet. – M. Stevanovic & C. Lindholm (toim.), *Keskustelunanalyysi: Kuinka tutkia sosiaalista toimintaa ja vuorovaikutusta*, 295–309. Vastapaino.

Mlynář, J., de Rijk, L., Liesenfeld, A., Stommel, W. & Albert, S. 2024. AI in situated action: A scoping review of ethnomethodological and conversation analytic studies. *AI & Society* 40: 1497–1527.

Vepsäläinen, H. 2024. Katsaus. Keskustelukumppanina chatbotti: Keskustelun simuloiminen chatbotti-avusteisessa valmennuksessa. *Virittäjä* 128: 398–421.