

Veikka Liukkonen

TEKOÄLYN JA MULTIMODAALISUUDEN HYÖDYNTÄMINEN HAKUKONEISSA

Informaatioteknologian ja viestinnän tiedekunta
Kandidaattitutkielma
Toukokuu 2024

TIIVISTELMÄ

Veikka Liukkonen: Tekoälyn ja multimodaalisuuden hyödyntäminen hakukoneissa
Kandidaattitutkielma
Tampereen yliopisto
Tietotekniikan tutkinto-ohjelma
Toukokuu 2024

Tekoäly on kehittynyt viime vuosina nopeasti ja suuriin kielimalleihin (engl. Large Language Model, LLM) perustuvia keskustelubotteja käytetään monipuolisesti. Suosituimmat hakukoneet eivät ole kuitenkaan juuri kehittyneet ulkoisesti viime vuosikymmenen aikana. Multimodaalisuudella tarkoitetaan eri moodien yhdistämistä. Erilaisia moodeja ovat esimerkiksi ääni, kuva ja teksti. Tämän työn tavoitteena on tutkia, miten tekoälyä ja multimodaalisuutta voidaan hyödyntää hakukoneissa. Multimodaalisuutta tarkastellaan tässä suhteessa tekoälyyn. Tarkastelun kohteena on siis se, miten hakukoneiden multimodaalisia ominaisuuksia voidaan kehittää tekoälyn avulla.

Suuret kielimallit ovat edistyneitä tekoälymalleja, jotka kykenevät käsittelemään luonnollista kieltä. Niiden luotettavuudessa on kuitenkin merkittäviä ongelmia. Kielimallit voivat viitata olemattomiin lähteisiin tai antaa epäolennaisia tuloksia. Suurten kielimallien toiminta ei ole myöskään läpinäkyvää, joten mahdollisia vinoumia niiden päätöksenteossa on vaikea havaita. Suurten kielimallien ongelmat johtuvat muun muassa koulutusdatan vinoumista ja aihepiiristä riippuen ylisovittumisesta, joka johtuu parametrien liiallisesta määrästä suhteessa koulutusdataan. Voidaan siis todeta, että yleisimmät, ChatGPT:n kaltaiset, suuriin kielimalleihin perustuvat keskustelubotit eivät ole luotettavia lähteitä tiedonhaussa.

Erilaisia tekoälysovelluksia käytetään yleisesti hakukoneissa. Erityisesti akateemisissa hakukoneissa on pyritty hyödyntämään suuria kielimalleja tiedonhaun helpottamiseen. Suurten kielimallien ongelmat ilmenevät siten myös tekoälyä hyödyntävissä hakukoneissa. Multimodaalisuuden mahdollisuuksia ei ole vielä hyödynnetty laajasti hakukoneissa. Hakutuloksissa eli hakukoneiden tulosteissa esiintyy useita eri moodeja, mutta syötteissä moodien yhdistäminen on vielä rajallista. Suuret kielimallit pystyvät käsittelemään tekstin lisäksi myös visuaalisia syötteitä, joten tekoäly voisi mahdollistaa edistyneempien multimodaalisten ominaisuuksien käytön hakukoneissa.

Tekoäly luo uusia mahdollisuuksia sille, mitä hakukoneet voivat tarjota, mutta sen ongelmia ei voi jättää huomiotta. Oikein toteutettuna tekoäly mahdollistaa hakukoneiden nykyistä monipuolisemmat syötteet ja osuvammat tulokset. Erityisesti suurilla kielimalleilla voitaisiin kehittää hakukoneita nykyisestään helpottamalla käyttäjän ja hakukoneen välistä kommunikaatiota esimerkiksi mahdollistamalla luonnollisen kielen käytön syötteessä.

Avainsanat: Hakukone, multimodaalisuus, tekoäly, suuret kielimallit, LLM

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

SISÄLLYSLUETTELO

1. JOHDANTO	1
2. HAKUKONEET	3
3. TEKOÄLY	4
3.1 Tekoälyyn liittyviä käsitteitä	4
3.2 Tekoälyn ongelmia	5
4. TEKOÄLY NYKYISISSÄ HAKUKONEISSA	8
4.1 Akateemiset hakukoneet	8
4.2 Verkon hakukoneet	10
5. MULTIMODAAALISUUS	11
6. YHTEENVETO	13
LÄHTEET	15

1. Johdanto

Tekoäly on kehittynyt viime vuosien aikana hyvin nopeasti ja erilaiset tekoälytyökalut ovat myös laajalti kuluttajien saatavilla ilmaiseksi tai kohtalaisen maltillisilla hinnoilla. Erityisesti suosioon ovat nousseet *suuria kielimalleja* (engl. Large Language Model, LLM) hyödyntävät tekoälysovellukset. Näistä tunnetuimpia ovat OpenAI:n kehittämä ChatGPT, Microsoftin Copilot ja Googlen Gemini (entinen Bard). Suuret kielimallit hyödyntävät suuria datamääriä luodakseen ihmismäistä tekstiä (Tikhonova & Raitskaya, 2023).

Multimodaalisuus käsitteenä tarkoittaa useiden *moodien* hyödyntämistä yhdessä. Jyväskylän yliopiston kielikellon mukaan erilaisia moodeja ovat esimerkiksi ääni, kuva ja teksti. Tekoälyn ja suurten kielimallien kyky yhdistellä ja käsitellä suuria määriä erilaista dataa luo uusia mahdollisuuksia multimodaalisuudelle hakukoneissa, esimerkiksi hyödyntämällä useita moodeja niin syötteissä kuin tulosteissa.

Hakukoneet ovat järjestelmiä, jotka mahdollistavat tiedon hakemisen laajoista tietokannoista esimerkiksi hakusanojen avulla. Toisin kuin tekoäly, laajimmassa käytössä olevat verkon hakukoneet eivät ole juurikaan kehittyneet tai muuttuneet ulkoisesti käyttäjän näkökulmasta viimeisen vuosikymmenen aikana. Tämän vuoksi on syytä tutkia, miten hakukoneet hyödyntävät tekoälyä ja multimodaalisuutta, sekä millaisia uusia ominaisuuksia ne voisivat mahdollistaa hakemiselle. Tämän työn tarkoituksena on siis vastata tutkimuskysymykseen: *Miten tekoälyä ja multimodaalisuutta voidaan hyödyntää hakukoneissa?*

Tarkemmin ottaen multimodaalisuutta tutkitaan erityisesti tekoälyn näkökulmasta eli kuinka tekoäly mahdollistaa multimodaalisia ominaisuuksia hakukoneissa. Tässä työssä ei siis käsitellä esimerkiksi hakukoneiden verkkosivujen asetteluun liittyvää multimodaalisuutta, ellei tekoäly liity siihen suoranaisesti.

Suurten kielimallien ja tekoälyä hyödyntävien teknologioiden ongelmat voivat olla kriittisiä tiedonhaun kannalta. Ongelmien tiedostaminen on olennaista, jotta uusien tekoälyteknologioiden mahdollisuudet voidaan toteuttaa käytännössä hakukoneissa. Kyseiset

ongelmat voidaan havaita esimerkiksi suuriin kielimalleihin perustuvien keskustelubotien epäjohdonmukaisissa vastauksissa. Näihin ongelmiin lukeutuvat muun muassa niin sanottu hallusinointi ja mustan laatikon ongelma.

Työssä käsitellään aiheita seuraavassa järjestyksessä: Luvun 2 tarkoituksena on kertoa hakukoneista ja niiden toiminnasta yleisesti. Hakukoneiden kohtalainen ymmärrys on olennaista, jotta niitä voidaan käsitellä tekoälyominaisuuksien näkökulmasta. Tämän työn tarkoituksena ei siis ole tarkastella yksityiskohtaisesti hakukoneiden toimintaperiaatetta ja taustalla olevia teknologioita. Luvussa 3 käsitellään tekoälyyn liittyviä käsitteitä ja ongelmia. Tekoäly on itsessään hyvin laaja käsite, joten selitettäväksi on valittu hakukoneiden kannalta olennaisimmat termit. Luvussa 4 tarkastellaan jo olemassa olevia tekoälyratkaisuja hakukoneissa ja ylipäätään tiedonhaussa. On syytä selvittää, millaisia tekoälytyökaluja on luotu helpottamaan tiedonhakua. Tekoälyn ja suurten kielimallien käyttö on kasvanut nopeasti viime vuosien aikana, joten monet työkalut ovat suhteellisen uusia. Luvussa 5 käsitellään multimodaalisuutta ja sen mahdollisuuksia tekoälyn suhteen. Perinteisissä hakukoneissa haun syötteenä käytetään hakusanaa ja tulosteena saadaan yleensä lista linkkejä, kuvia tai videoita. Muiden syötetapojen kuin tekstin mahdollisuuksia on siis syytä pohtia. Luvun 6 tarkoituksena on toimia keskusteluna ja yhteenvetona työn tuloksista.

Työn tutkimusmenetelmä on kirjallisuuskatsaus. Sopivia lähteitä on etsitty lähinnä Andorista ja ACM Digital librarystä. Haussa usein käytettyjä avain- ja hakusanoja ovat muun muassa: artificial intelligence, search engine, large language model, sekä näiden lyhenneet. Löydetyistä aineistoista on pyritty valitsemaan mahdollisimman tuoreita vertaisarvioituja lähteitä. Suurten kielimallien suosio ja kehitys on kasvanut varsin nopeasti parin viime vuoden aikana, joten sitä vanhemmat lähteet saattavat olla jokseenkin vanhentuneita.

2. Hakukoneet

Tämän luvun tarkoituksena on tarkastella hakukoneiden perusteita. Yksinkertaisimmillaan hakukoneelle annetaan *syöte* (engl. input), jonka jälkeen hakukone antaa vastauksena *tulosteen* (engl. output). Syöte on usein hakusana ja tulosteena usein lista linkkejä verkkosivuille. Luku perustuu pitkälti Dirk Lewandowskin kirjaan *Understanding Search Engines* (2023).

Hakukoneella yleisesti tarkoitetaan yleensä verkon hakukoneita, jotka pyrkivät kattamaan mahdollisimman laajasti verkossa olevia sivustoja. Koko verkon kattavia hakukoneita kutsutaan myös *horisontaalisiksi* hakukoneiksi. *Vertikaaliset* hakukoneet (engl. vertical search engines) pyrkivät sen sijaan keskittymään tiettyyn aihepiiriin tai sivustoon, jotta hakukone voi antaa olennaisempaa tietoa kuin horisontaaliset koko verkon kattavat hakukoneet. Vertikaalisten hakukoneiden tapauksessa hakukoneen kattamat sivustot valitaan yleensä manuaalisesti. Paras esimerkki horisontaalisesta hakukoneesta lienee Googlen hakukone, koska se on ylivoimaisesti tunnetuin ja käytetyin hakukone. Muita horisontaalisia hakukoneita ovat esimerkiksi Bing, Yahoo! ja DuckDuckGo.

Lewandowski listaa seuraavat syyt miksi vertikaalisia hakukoneita ei voida korvata horisontaalisilla:

1. Teknisten rajoitteiden vuoksi horisontaaliset hakukoneet eivät voi kattaa koko verkkoa.
2. Taloudelliset esteet rajoittavat sisällön keruuta ja indeksointia.
3. Horisontaaliset hakukoneet ovat suunnattu keskivertokäyttäjille.
4. Horisontaaliset hakukoneet joutuvat jatkuvasti indeksoimaan kaiken sisällön, jotta kaikki tieto on haettavissa yhdessä.

Tekoäly tuskin tekee vertikaalisten hakukoneiden syrjäyttämistä mahdollista, mutta mahdollisesti tekoälyä voitaisiin käyttää vertikaalisten hakukoneiden parantamiseen. Indeksoitavien sivujen manuaalista valitsemista voitaisiin korvata tai täydentää tekoälyn avulla.

3. Tekoäly

Tekoäly on laaja käsite ja sillä on useita eri osa-alueita. Kielitoimiston sanakirjan määritelmän mukaan tekoäly tarkoittaa tietokoneen toimintoja, jotka jäljittelevät ihmiselle tyypillistä älykkyyttä vaativia toimintoja. Tässä luvussa käsitellään tekoälyä yleisesti, erityisesti suhteessa tekoälyyn perustuviin hakukoneisiin. Luku on jaettu alilukuihin 3.1. ja 3.2. Ensimmäisessä aliluvussa käydään läpi hakukoneiden ja multimodaalisuuden kannalta olennaisia käsitteitä, kun taas jälkimmäisessä tarkastellaan tekoälyyn liittyviä ongelmia ja riskejä.

3.1. Tekoälyyn liittyviä käsitteitä

Luonnollisen kielen käsittely (engl. natural language processing, NLP) on tekoälyn osa-alue, joka mahdollistaa tietokoneille ihmisten tuottaman luonnollisen kielen analysoinnin ja käsittelyn (Karanikolas et al., 2024). Luonnollisen kielen käsittely on hyvin keskeinen osa tekoälyssä, koska ihmismäiseen toimintaan kuuluu olennaisesti ihmisten tuottaman puheen ja tekstin ymmärtäminen. Hakukoneet ovat tehty ihmisille, joten ideaalisesti hakukoneet ymmärtäisivät syötteenä luonnollista kieltä.

Suuret kielimallit ovat pitkälle kehittyneitä tekoälymalleja, jotka kykenevät suorittamaan luonnollisen kielen käsittelyä. Luonnollisen kielen käsittely painottuu yhä enemmän suuriin kielimalleihin verrattuna aiempiin luonnollisen kieleen käsittelyssä käytettyihin sääntöpohjaisiin järjestelmiin. (Karanikolas et al., 2024)

Suurista kielimalleista tunnetuimpia lienevät OpenAI:n GPT-kielimallit (Generative Pre-trained Transformer), joihin ChatGPT-tekoälykeskustelubotti ja monet muut tekoälysovellukset pohjautuvat. Suuret kielimallit lienevät pääsyy tekoälyn käytön yleistymiselle maallikoiden keskuudessa esimerkiksi keskustelubottien muodossa.

Olellainen teknologia suurien kielimallien kannalta on muuntaja (engl. transformer), joka on yksi neuroverkkoarkkitehtuuryyppi (Amaratunga, 2023). Enemmistö suurista kielimalleista Amaratungan (2023) mukaan perustuu muuntajiin, kuten esimerkiksi GPT-kielimalliperhe, jossa T-kirjain merkitsee transformer:ia, eli muuntajaa.

Kielimalleista suuria tekee suuri määrä parametrejä, sekä massiivinen määrä koulutusdataa. Moderneissa suurissa kielimalleissa parametrejä voi olla jopa satoja miljardeja. Yleisesti voidaan sanoa, että mitä enemmän kielimallissa on parametrejä, sitä enemmän se voi oppia. Kielimallin tarkoituksena on saavuttaa yleistäminen (engl. Generalization). Yleistämisen ideana on se, että suuren parametri- sekä koulutusdatamäärän avulla kielimalli kykenee käsittelemään oikein koulutusdatasta erillistä dataa. Varjopuolena tästä syntyy ylisovittumisen (engl. overfitting) ongelma. Jos kielimallissa on liian paljon parametrejä koulutusdataan verrattuna, niin kielimalli yleistämisen sijaan "opettelee ulkoa" koulutusdatan piirteet. Toisin sanoen suurempi määrä parametrejä ei aina ole hyödyksi. Ylimoituksessa kielimalli suoriutuu hyvin koulutusdatalla, mutta tekee paljon virheitä muulla datalla. (Amaratunga, 2023)

Semanttinen samankaltaisuus (engl. semantic similarity) on tapa laskea, kuinka samankaltaisia esimerkiksi luonnollisen kielen sanat tai lauseet ovat. Semanttisen samankaltaisuuden avulla tekoäly pystyy esimerkiksi tunnistamaan synonyymeja, joka on hyvin tärkeää ihmisten kanssa kommunikoinnin mahdollistamiseksi. (Kulmanov et al., 2024)

3.2. Tekoälyn ongelmia

Tekoäly on oiva työkalu, mutta sen hyödyntämisessä tulee ottaa huomioon myös haittapuolet. Suurten kielimallien tunnetuimpiin ongelmiin lukeutuvat Gusenbauerin (2023) mukaan mustan laatikon ongelma, hallusinointi ja mahdolliset vinoutumat (engl. bias) päätöksenteossa. Semanttisen samankaltaisuuden tapauksessa samankaltaisuuden mittauksen laatu riippuu valitusta mittarista ja asiayhteydestä (Kulmanov et al., 2024). Myös semanttinen samankaltaisuus kärsii datan vinoumista. Merkittävän ongelman tiedonhauulle luo myös tekoälyhakukoneiden eriävät hakutulokset (Gusenbauer, 2023).

Tällä tarkoitetaan sitä, että samalla hakusanalla saa eriäviä tuloksia, vaikka haut tois-
taisi välittömästi peräkkäin. Erityisesti tieteellisessä tutkimuksessa toistettavuus on tär-
keää, joten hakutulosten jatkuva vaihtuminen samalla hakusanalla on kriittinen vika.

Suurten kielimallien hallusinoinnissa on kyse siitä, että tuotetussa tekstissä oleva tieto
on virheellistä tai aiheeseen liittymätöntä (Ahmad et al., 2023). Hallusinointi voi johtua
Ahmadin et al. (2023) mukaan muun muassa siitä, että kielimallia ei ole koulutettu aihee-
seen liittyvällä tiedolla, jolloin se keksii tyhjästä, eli hallusinoi, perätöntä tietoa. Toinen
hallusinoinnin muoto on olemattomien lähteiden keksiminen (Sanderson, 2023).

Hallusinoinnin muita syitä ovat suuren kielimallin ylisovittuminen ja vinoumat koulutus-
datassa. Toisin kuin monissa muissa tekoälysovelluksissa, oikea vastaus ei aina ole
suurien kielimallien tapauksessa yksiselitteinen. Esimerkiksi kuvien luokittelussa teko-
älyllä koulutusdatan avulla voidaan saada selkeä vastaus, onko tekoälyn luokittelu oi-
kein vai väärin. Vastaavasti suurissa kielimalleissa syötteet voivat olla hyvin moninaisia
ja monimutkaisia, joten oikean tulosteen tunnistaminen ei ole samalla tapaa yksinker-
taista. (Amaratunga, 2023)

Mustan laatikon ongelma on se, että ihmiset eivät kykene ymmärtämään tekoälyn pää-
töksentekoa, vaan tärkeimpänä pidetään hyviä tuloksia (Rai, 2020). Mustan laatikon on-
gelman seurauksena voidaan Gusenbauerin (2023) mukaan pitää sitä, että ei voida tie-
tää miten tekoäly valitsee lähteensä. Joissain tilanteissa tekoäly voi jättää huomiotta hy-
vin olennaisia lähteitä ja suosia vain tiettyntyyppisiä lähteitä. Tällaisia vinoumia voi olla
vaikeaa havaita, koska tekoälyn tai suuren kielimallin toiminta ja painotukset eivät ole
ihmisen ymmärrettävissä. Ratkaisuksi vinoumiin Gusenbauer (2023) esittää puolueetto-
mien kolmansien osapuolten suorittamia tarkastuksia ja tutkimuksia selvittämään ja tuo-
maan esille suurten kielimallien rajoitteita ja puolueellisia vinoumia.

Konkreettinen esitys tavasta arvioida suurten kielimallien vinoumia on Oketunji et al.
(2023) luoma Large Language Model Bias Index (LLMBI), joka on esitetty mittari suurten
kielimallien vinoumille. Oketunji et al. (2023) mukaan suuret kielimallit saattavat taipua
samankaltaisiin ennakkoluuloihin kuin ihmisetkin. Muun muassa sukupuoliin liittyvät ja
rasistiset stereotypiat vaikuttavat suurten kielimallien päätöksentekoon. Esitetty LLMBI-

malli mittaa ihmisläheisempää vinoumaa, joten se on hyödyllinen suurin kielimalleihin perustuvien keskustelubottien vinoumien selvittämiseen. Tiedonhaun kannalta olennais-
ten tietäntyyppisten lähteiden suosimisen mittaaminen lienee vaikeampaa. Kuitenkin tie-
donhakua helpottavat suuria kielimalleja hyödyntävät tiivistelmiä luovat tekoälytyökalut
saattavat tehdä vastaavanlaisia vinoumia valitessaan tietoa tiivistelmään.

4. Tekoäly nykyisissä hakukoneissa

Tässä luvussa tarkastellaan jo olemassa olevia tekoälypohjaisia hakukoneita. Suuria kielimalleja hyödynnetään erityisesti akateemisissa hakukoneissa, kun taas verkon suurissa horisontaalisissa hakukoneissa hyödynnetään enemmän äänen- ja hahmon-tunnistusta. Tämän vuoksi luku on jaettu erillisiin alilukuihin akateemisille vertikaalisille hakukoneille ja verkon horisontaalisille hakukoneille. Tekoälyn suosio on kasvanut nopeasti viime vuosina, joten termiä ”tekoäly” käytetään paljon myös markkinointimielessä monenlaisista algoritmeista.

Nykyisissä tekoälyä hyödyntävissä hakukoneissa yleisenä ominaisuutena voidaan pitää luonnollisen kielen ymmärrystä. Suuri osa suosituista kielimalleja hyödyntävistä työkaluista pystyy käsittelemään ihmisten luomaa luonnollista kieltä, ja tuottamaan vastauksia luonnollisella kielellä (Gusenbauer 2023).

4.1. Akateemiset hakukoneet

Akateemisilla hakukoneilla tarkoitetaan hakukoneita, joiden avulla haetaan tieteellisiä artikkeleita ja julkaisuja. Useimmiten akateemiset hakukoneet ovat vertikaalisia hakukoneita, koska ne rajautuvat usein yhteen tai pieneen määrään tietokantoja. Esimerkiksi Tampereen yliopiston Andoria voidaan pitää akateemisena hakukoneena. Tieteellinen hakeminen on aikaavievää työtä, jonka tehostamiseksi on kehitetty erilaisia tekoälyratkaisuja. Tieteelliseen tutkimukseen liittyy olennaisesti uusien asioiden tutkimista ja löytämistä, joten on luonnollista, että myös akateemiset hakukoneet ovat edelläkävijöitä tekoälyn hyödyntämisessä. Tunnettuja jo käytössä olevia tieteelliseen tutkimukseen tarkoitettuja tekoälytyökaluja ovat muun muassa Elicit, Scite ja Consensus (Sanderson 2023).

Elicit-hakukone eroaa tavallisesta hakukoneesta siten, että hakusanan sijaan tulee käyttää tutkimuskysymystä. Elicit hyödyntää tekoälyä myös tutkimuskysymyksen luomiseksi. Käyttäjä voi hyödyntää niin sanottua aivoriihiätyökalua (engl. brainstorming), joka luo eri-

laisia tutkimuskysymyksiä käyttäjän antamien avainsanojen perusteella. Hakukoneen tulokset ovat samankaltaisia kuin muissakin tieteellisten tietokantojen hauissa, eli tuloksena on lista tieteellisiä artikkeleita. Tulokset saadaan hyödyntämällä semanttista samankaltaisuutta tutkimuskysymyksen suhteen. (Whitfield & Hofmann 2023)

Elicit-hakukoneessa hyödynnetään useita suuria kielimalleja, joka mahdollistaa esimerkiksi yhdestä tai useammasta tuloksesta automaattisesti luotavan tiivistelmän. Toinen mahdollisuus on samankaltaisten tulosten etsiminen, jolloin hakukone tarjoaa semanttista samankaltaisuutta hyödyntämällä löydettyjä valittuun lähteeseen liittyviä samankaltaisia tuloksia. Elicit hyödyntää Unpaywall-tietokannassa olevia tuloksia, mutta käyttäjä voi ladata sivulle myös muita papereita PDF-muodossa. (Whitfield & Hofmann 2023)

Scite on tieteelliseen hakemiseen tarkoitettu tekoälytyökalu, joka pyrkii organisoimaan lähdeviittaukset (Sanderson 2023). Tämä sisältää myös tiedon, missä tiettyä paperia on viitattu. Scite tarjoaa Sandersonin mukaan myös *Scite Assistant* -työkalun, joka toimii vastaavalla tavalla kuin yleisimmät suuria kielimalleja hyödyntävät chattibotit, kuten ChatGPT. Scite Assistant eroaa muista siten, että se antaa vastauksien lähteet ja viittaa lähteisiin tekstissä. Sanderson kertoo Scite:n tekevän yhteistyötä useiden julkaisufoorumien kanssa, joka mahdollistaa suuren määrän käytettäviä lähteitä. Scite pyrkii korjaamaan muiden suuriin kielimalleihin perustuvien keskustelubottien merkittävän luotettavuuden puutteen.

Consensus on kolmas tutkimukseen suunnattu tekoälytyökalu ja Scite tekee sen kanssa yhteistyötä. Consensus-työkalun ideana on esittää prosenttilukuna, kuinka suuri osa tutkimuksista on samaa ja eri mieltä annetun väitteen kanssa. (Sanderson 2023)

Sekä Elicit että Consensus käyttävät Sandersonin (2023) mukaan Semantic Scholar-hakukonetta lähteiden etsimiseen. Semantic Scholar on tekoälypohjainen akateeminen hakukone, jonka taustalla on voittoa tavoittelematon *Allen Institute for Artificial Intelligence* (Fricke 2018). Semantic Scholar –verkkosivuilla ominaisuuksiksi kerrotaan muun muassa tekoälyn luomia tiivistelmiä ja käyttäjän aiempaan selaushistoriaan perustuvia tekoälyn luomia suosituksia uusista tutkimuksista.

Akateemisten tekoälytyökalujen ongelmat ovat pitkälti samanlaisia kuin muissa tekoälysovelluksissa. Suuret kielimallit hallusinoivat olemattomia lähteitä ja tekoäly on musta laatikko, eli toimintaperiaate päätösten takana ei ole tiedossa. Ei ole esimerkiksi tiedossa, että suosivatko suuret kielimallit tietyn tyyppisiä lähteitä ilman järkevää syytä (Gusenbauer 2023). Esimerkiksi Elicit-hakukoneen tiivistelmien laatua on kritisoitu (Sanderson 2023). Tiivistelmät sisältävät myös epäolennaista tietoa. Väitteisiin kantaa ottavaa Consensus-työkalua on Sandersonin (2023) mukaan ollut syytä tarkastaa manuaalisesti, sillä se ei arvioi lähteiden laatua, vaan määrää.

4.2. Verkon hakukoneet

Verkon hakukoneet ovat yleensä horisontaalisia hakukoneita. Ne ovat myös yleisimmin käytettyjä hakukoneita. Verkon hakukoneiden yleisiä käyttöä helpottavia ominaisuuksia ovat esimerkiksi puheentunnistus ja käänteinen kuvahaku, jossa syötteenä käytetään hakusanan sijaan kuvaa.

Microsoft onnistui jopa kaksinkertaistamaan päivittäisten Bing-hakukoneen käyttäjien määrän ottaessaan käyttöön OpenAI:n ChatGPT-tekniikan hakukoneessaan. (Wang 2023). Tämä ei ollut kuitenkaan tarpeeksi iso muutos, jotta Microsoft voisi haastaa Googlen verkon hakukonemarkkinoilla. Tällaisesta suosion kasvusta voidaan kuitenkin päätellä, että hakukoneiden tekoälyominaisuuksille on kysyntää. On hyvin todennäköistä, että hakukoneissa tullaan jatkossa näkemään yhä enemmän tekoälyominaisuuksia. Se lienee järkevää jo pelkästään markkinoinnin näkökulmasta ottaen huomioon tekoäly-sanan trendikkyuden.

5. Multimodaalisuus

Tässä luvussa tarkastellaan, kuinka tekoälyn avulla voidaan mahdollistaa multimodaaliset ominaisuudet hakukoneissa. Multimodaalisuutta voidaan hyödyntää sekä syötteessä (engl. input) että tulosteessa (engl. output).

Multimodaalisuus tarkoittaa Kesslerin (2022) mukaan kommunikointia useiden moodien avulla. Moodeja voivat olla esimerkiksi teksti, ääni, kuva ja eleet. Moodit voivat Kesslerin (2022) mukaan olla tyypiltään lingvistisiä, visuaalisia, auditiivisia, gesturaalisia tai spatiaalisia. Jyväskylän yliopiston kielikompassi selittää moodien tyypit seuraavasti:

1. Lingvistinen moodi: kielellinen ilmaisu, esimerkiksi äänneistä koostuvat sanat ja lauseet
2. Auditiivinen moodi: kuuloaistiin perustuva moodi, esimerkiksi puhe ja musiikki
3. Spatiaalinen moodi: tilan käytön moodi, esimerkiksi verkkosivun asettelu
4. Gesturaalinen moodi: eleiden ja liikkeen moodi, esimerkiksi liikkuvat elementit verkkosivulla
5. Visuaalinen moodi: näköaistin moodi, esimerkiksi kuvat, kaaviot ja värit

Tekoälyn kannalta merkityksellisiä listatuista moodien luokista ovat ainakin lingvistinen, auditiivinen ja visuaalinen moodi, sillä tekoäly kykenee käsittelemään tekstiä (mukaan lukien luonnollista kieltä), ääntä, puhetta ja kuvia. Hakukoneiden kannalta myös spatiaalinen moodi voisi olla olennainen, koska hakutulokset voidaan asetella monella eri tapaa.

Internetissä lähes kaikki sivut hyödyntävät nykyään multimodaalisuutta. On vaikeaa löytää verkkosivua, joka koostuisi vaikkapa pelkästä tekstistä ilman minkäänlaisia kuvia tai kuvakkeita. Tekoälyn kehityksen myötä multimodaalisuutta voitaisiin käyttää myös hakemisen syötteenä. Tähän liittyviä tekoälyn ominaisuuksia ovat muun muassa tietokone-näkö sekä hahmon- ja äänentunnistus.

Suurten kielimallien kyvyt käsitellä visuaalisia syötteitä on kehittynyt kielimallista riippuen korkealle tasolle. Esimerkiksi GPT-4V –kielimalli pystyi ratkomaan visuaalisia ohjelmointitehtäviä yli 96 % tarkkuudella (Hou et al., 2024). Tämä tarkoittaa, että suurilla kielimalleilla on suuri potentiaali käyttää myös kuvan moodia syötteenä. Suuria kielimalleja hyödyntämällä hakukoneiden multimodaalisilla syötteillä voisi monissa tilanteissa helpottaa hakemista. Esimerkiksi tekstin ja kuvan yhdistäminen syötteessä tarkentaa hakua verrattuna vain yhden moodin käyttämiseen.

Konkreettinen esimerkki moodien yhdistämisestä haussa on Googlen Multisearch. Multisearch mahdollistaa multimodaalisen hakemisen kuvan ja tekstin avulla. Käyttäjä pystyy hakemaan verkosta ottamansa kuvan avulla, ja lisäämään hakusanoja tarkentamaan hakua (Shantanu, 2022). Vuonna 2022 Shantanun mukaan palvelu oli saatavissa vain englanniksi. Toisaalta suurten kielimallien kyky käyttää sulavasti eri kieliä mahdollistaa ennemmin tai myöhemmin Multisearch-ominaisuuden monilla eri kielillä.

OpenAI:n omien verkkosivujen (2023) mukaan GPT-kielimallit tukevat multimodaalista syötettä. Syötteenä voi käyttää tekstin lisäksi myös kuvia tai puhetta. Mobiililaitteilla on myös mahdollista keskustella ChatGPT-chattibotin kanssa puhumalla ja saada vastauksena tietokoneen tuottamaa puhetta. Tässä ominaisuudessa yhdistyy puheentunnistus ja teksti puheeksi –teknologia. Amaratungan (2023) mukaan suurilla kielimalleilla on multimodaalisia ominaisuuksia, joiden avulla ne pystyvät esimerkiksi löytämään yhteyksiä tekstin ja kuvan välillä, tai ymmärtämään puhuttua kieltä ja sen suhdetta kirjoitettuun tekstiin.

6. Yhteenveto

Hakukoneet käyttävät monipuolisesti erilaisia tekoälyominaisuuksia helpottamaan tiedon hakemista. Perinteisiin tekoälyominaisuuksiin lukeutuvat muun muassa äänen- ja hahmontunnistus, joita käytetään erityisesti multimodaalisuuden mahdollistamisessa. Yleisin hakukoneiden syötteissä käytetty moodi on teksti, mutta myös kuvia ja ääntä voidaan käyttää hakemisessa. Multimodaalisuutta on kuitenkin hyödynnetty vain vähän suosituimmista hakukoneista. Nykyisten hakukoneiden multimodaaliset syötteet rajautuvat lähinnä kuvien ja tekstin yhdistelmähakuun. Tulosteissa multimodaalisuus on paremmin edustettuna; yhdellä haulla löytyy tekstiä, kuvia ja videoita.

Nykyisistä hakukoneista suuria kielimalleja hyödyntävät erityisesti akateemiset hakukoneet. Tieteellisen yhteisön rooli edelläkävijänä heijastuu siten myös akateemisiin hakukoneisiin. Akateemisten tekoälyyn perustuvien hakukoneiden tarkoituksena on nopeuttaa akateemista tiedonhakuja, joka on usein hyvin aikaavievää työtä. Kuitenkin tiedeyhteisössä näihin tekoälyä hyödyntäviin hakukoneisiin on suhtauduttu varauksella, sillä tekoälyn ja erityisesti suurten kielimallien ongelmat ovat laajasti tiedossa. Nähtäväksi jää, miten tekoälyhakukoneet kehittyvät ja onnistuvat estämään tai lievittämään ongelmia hallusinoinnissa ja toistettavuudessa.

Suuria kielimalleja voitaisiin varmasti hyödyntää entistä enemmän myös verkon hakukoneissa. Edelläkävijänä voidaan pitää Bing-hakukonetta, mutta tekoälyominaisuudet tulevat varmasti lisääntymään myös muissa hakukoneissa. Googlella on varmasti tahtoa säilyttää ylivoimaisen markkina-asemansa, joten myös Googlen haussa tullaan suurella todennäköisyydellä enemmän tekoälyyn perustuvia ominaisuuksia.

Suurten kielimallien ongelmien vuoksi käyttäjän kannattaa ainakin toistaiseksi tarkistaa lähteet itse. Tämän kirjallisuuskatsauksen perusteella voidaan todeta, että esimerkiksi tekoälyn tuottamiin tiivistelmiin artikkeleista tulee suhtautua varauksella, koska tekoäly on saattanut ymmärtää tekstin väärin tai jättää huomioimatta olennaisia yksityiskohtia.

Suuri osa tiedosta verkossa on englanniksi. Toisaalta osa tiedosta on saatavilla vain yhdellä kielellä. Tekoälyn avulla voisi olla mahdollista hakea yhdellä haulla tietoa useilla eri kielillä, joka sitten käännettäisiin käyttäjälle mieleiselle kielelle. Konekääntäminen on

kohtalaisen laadukasta ja suuret kielimallit ymmärtävät sulavasti useita eri kieliä, mukaan lukien suomea.

Tekoäly kehittyy jatkuvasti, joten asiaan liittyvät tieto saattaa myös vanhentua nopeasti. Työssä käytetyt lähteet ovat pääsääntöisesti tuoreita ja ne ovat julkaistu viimeisen parin vuoden aikana, mutta niissä oleva tieto saattaa olla jo osin vanhentunutta. Erityisesti suuriin kielimalleihin perustuvat tekoälysovellukset kehittyvät hyvin nopeasti. Kuitenkin uusimmista tiedoista on vaikea löytää vertaisarvioituja lähteitä, joten myös jatkotutkimukselle olisi perusteita. Jatkotutkimuksissa voisi olla mielenkiintoista tarkastella erityisesti, millaisin keinoin suuriin kielimalleihin perustuvat hakukoneet pystyvät lieventämään suurten kielimallien ongelmia ja riskejä. Esimerkiksi suurten kielimallien hallusinointi voisi olla riittävän laaja aihe syvällisemmälle tutkimukselle. Myös multimodaalisuuden mahdollisuuksia voitaisiin tutkia tätä työtä tarkemmin myöhemmissä tutkimuksissa.

LÄHTEET

- Ahmad, Z., Kaiser, W. & Rahim, S. (2023). *Hallucinations in ChatGPT: An Unreliable Tool for Learning*. Rupkatha journal on interdisciplinary studies in humanities 15.4
<https://doi.org/10.21659/rupkatha.v15n4.17>
- Amaratunga, T. (2023). *Understanding Large Language Models: Learning Their Underlying Concepts and Technologies*. (ss. 81–117) Apress.
https://doi.org/10.1007/979-8-8688-0017-7_4
- Fricke, S. (2018) *Semantic Scholar*. Journal of the Medical Library Association 106.1 (ss. 145–147). <https://doi.org/10.5195/jmla.2018.280>
- Gusenbauer, M. (2023). *Audit AI Search Tools Now, before They Skew Research*. Nature (London) 617.7961 (ss. 439–439).
<https://doi-org.libproxy.tuni.fi/10.1038/d41586-023-01613-w>
- Hou, I., Man, O., Mettillie, S., Gutierrez, S., Angelikas, K. & MacNeil, S. (2024). *More Robots Are Coming: Large Multimodal Models (ChatGPT) Can Solve Visually Diverse Images of Parsons Problems*. 26th Australasian Computing Education Conference (ACE '24). Association for Computing Machinery (ss. 29–38).
<https://doi-org.libproxy.tuni.fi/10.1145/3636243.3636247>
- Jyväskylän Yliopisto, Kielikompassi: Multimodaalisuus (2024)
https://kielikompassi.jyu.fi/resurssikartta/netro/pankki/parametrit_moodi_multi.shtml (Haettu 5.3.2024)
- Karanikolas, N., Manga, E., Samaridi, N., Tousidou, E. & Vassilakopoulos, M. (2023) *Large Language Models versus Natural Language Understanding and Generation*. ACM International Conference Proceeding Series (ss. 278-290).
<https://doi.org/10.1145/3635059.3635104>
- Kessler, M. (2022) *Multimodality*. ELT journal 76.4 (ss. 551–554).
<https://doi.org/10.1093/elt/ccac028>
- Kielitoimiston sanakirja –verkkosivu (2024)
<https://www.kielitoimistonsanakirja.fi/#/teko%C3%A4ly> (Haettu 15.4.2024)

- Kulmanov, M., Smaili, F., Gao, X. & Hoehndorf, R. (2021) *Semantic Similarity and Machine Learning with Ontologies*. Briefings in bioinformatics 22.4
<https://doi.org/10.1093/bib/bbaa199>
- Lewandowski, D. (2023). *Understanding Search Engines*. Cham, Switzerland: Springer. (Luvut 2.4.1 ja 2.4.2)
- Oketunji, A., Anas, M & Saina, D. (2023). *Large Language Model (LLM) Bias Index – LLMBI*. Arxiv.org <https://doi.org/10.48550/arXiv.2312.14769>
- OpenAI –sivusto (2024)
<https://openai.com/blog/chatgpt-can-now-see-hear-and-speak>
 (Haettu 19.3.2024)
- Rai, A. (2020) *Explainable AI: From Black Box to Glass Box*. Journal of the Academy of Marketing Science 48.1 (ss. 137–141).
<https://doi.org/10.1007/s11747-019-00710-5>
- Sanderson, K. (2023). *AI Science Search Engines Are Exploding in Number - Are They Any Good?*. Nature (London) 616.7958 (ss. 639–640).
<https://doi-org.libproxy.tuni.fi/10.1038/d41586-023-01273-w>
- Semantic Scholar –verkkosivusto (2024) <https://www.semanticscholar.org/product>
 (Haettu 5.3.2024)
- Shantanu D. (2022). *Google multisearch: A new arena for advertisers?*. Exchange4media, 21 Dec. 2022. Gale Business: Insights.
<https://link.gale.com/apps/doc/A730796722/GBIB?u=tampere&sid=bookmark-GBIB&xid=1e3a9075> (Haettu 15.4.2024)
- Tikhonova, E. & Raitskaya, L. (2023). *ChatGPT: Where Is a Silver Lining? Exploring the Realm of GPT and Large Language Models*. Journal of language and education 9.3 (ss. 5–11). <https://doi.org/10.17323/jle.2023.18119>
- Wang, B. (2023) *Microsoft Bing With AI Doubles Bing Users*. Weblog post.
<https://lib-proxy.tuni.fi/login?qurl=https%3A%2F%2Fwww.proquest.com%2Fblogs-podcasts-websites%2Fmicrosoft-bing-with-ai-doubles-users%2Fdocview%2F2785210913%2Fse-2%3Faccountid%3D14242>
 (Haettu 20.5.2024)

Whitfield, S. & Hofmann, M. (2023). *Elicit: AI Literature Review Research Assistant*.
Public Services Quarterly 19.3 (ss. 201–207).
<https://doi-org.libproxy.tuni.fi/10.1080/15228959.2023.2224125>