



Vaasan yliopisto
UNIVERSITY OF VAASA

Markus Panttila

Neocloudi vai tekoälypilviyritys?

Tarkastelussa tekoälykäsitteet ja -termit Ylen tekoälyaiheisissa
verkkoartikkeleissa

Markkinoinnin ja viestinnän akateeminen yksikkö
Teknisen viestinnän pro gradu -tutkielma
Teknisen viestinnän maisteriohjelma

Vaasa 2026

VAASAN YLIOPISTO**Markkinoinnin ja viestinnän akateeminen yksikkö**

Tekijä:	Markus Panttila		
Tutkielman nimi:	Neocloudi vai tekoälypilviyitys? : Tarkastelussa tekoälykäsitteet ja -termit Ylen tekoälyaiheisissa verkkoartikkeleissa		
Tutkinto:	Filosofian maisteri		
Oppiaine:	Tekninen viestintä		
Työn ohjaaja:	Niina Nissilä		
Valmistumisvuosi:	2026	Sivumäärä:	114

TIIVISTELMÄ:

Tekoäly on ilmiönä merkityksellinen sen erikoisalan ulkopuolellakin tekoälyn yhteiskunnallisten vaikutusten takia, ja tekoälyilmiöstä kirjoittavilla journalisteilla on oma roolinsa kansalaisten tekoälylukutaidon kehittämisessä sekä ylläpitämisessä. Siksi tässä tutkielmassa tarkastellaan tekoälyn ilmiötä siitä viestinnällisestä näkökulmasta, miten Ylen verkkoartikkelit rakentavat tekoälyä koskevaa käsitteistöä ja millä tavoin käytetty termistö voi mahdollisesti edistää tai vaikeuttaa ilmiön ymmärrettävyyttä. Tutkimuksen tavoitteena on selvittää, millaisia tekoälykäsitteitä ja -termejä ilmenee Ylen yleistajuisissa tekoälyaiheisissa verkkoartikkeleissa. Tutkimusta ohjaavat kaksi tutkimuskysymystä, jotka ovat 1. Mitä tekoälykäsitteitä ja -termejä ilmenee Ylen tekoälyä käsittelevissä verkkoartikkeliteksteissä? 2. Miten tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit? Tutkimusaineisto koostuu vuonna 2025 julkaistuista tekoälyä käsittelevistä Ylen verkkoartikkeleista. Aineistosta muodostettiin tekstikorpus Sketch Engine -ohjelmalla, jolla korpukselle suoritettiin myös puoliautomaattinen terminpoiminta terminlistojen muodostamiseksi.

Tutkimusmenetelminä analyysissä käytettiin systemaattista käsiteanalyysiä sekä termianalyysiä, mikä asemoi tutkimuksen osaksi terminologisen tutkimuksen viitekehystä. Terminlistojen käsitteistä muodostettiin systemaattisella käsiteanalyysillä yleistajuisen aineiston tekoälykäsitteiden kokonaiskäsitejärjestelmää kuvaava graafinen satelliittimalli. Satelliittimallista ilmenee käsitesuhteiden kautta tekoälykäsitteiden kytkös tekoälyn pääkäsitteeseen sekä toisiin tekoälykäsitteisiin. Systemaattinen käsiteanalyysi osoitti, että Ylen verkkoartikkeleissa tekoälyn käsitteistöä rakennetaan etenkin geneeriseen käsitesuhteeseen sekä toimintasuhteeseen liittyvien käsitteiden kautta. Tämä tarkoittaa erilaisten tekoälyn sovellusten tyypittelyn ohella sitä, että tekoälyn ilmiöstä rakennetaan kuvaa etenkin toiminnan näkökulmasta.

Käsiteanalyysin tuloksena muodostetusta kokonaiskäsitejärjestelmästä poimittiin viisi tekoälytermiä termianalyysiin, jossa termien laatua tarkasteltiin hyvän termin kriteerien näkökulmista. Termianalyysi osoitti, että tekoälytermit ovat ongelmallisia etenkin käsitesisältöön ja kieleen liittyvien tekijöiden osilta, mikä viittaa käsitesisältöä tarkemmin kuvaavien omakielisten termien tarpeeseen. Tekoälyilmiön ymmärrettävyyttä voitaisiin parantaa verkkoartikkeliteksteissä keskittymällä etenkin termien omakielisyyteen ja läpinäkyvyyteen. Termianalyysin lopuksi muodostettiin vaihtoehtoisia termejä analyysissä ongelmallisiksi havaituille tekoälytermeille. Vaihtoehtoisia termejä on mahdollista käyttää jatkossa esimerkiksi suuntaa antavana ohjenuorana termien muodostamiselle tekoälystä kirjoitettaessa.

AVAINSANAT: käsiteanalyysi, systemaattinen käsiteanalyysi, tekoäly, tekoälykäsitteet, tekoälytermit, termianalyysi, terminologia

Sisällys

1	Johdanto	6
1.1	Tavoite	7
1.2	Aineisto	10
1.3	Menetelmä	14
2	Terminologinen tutkimus	17
2.1	Tarkoite ja käsite	18
2.2	Käsitepiirteet	21
2.3	Käsitesuhteet ja käsitejärjestelmät	23
2.3.1	Geneeriset käsitesuhteet	23
2.3.2	Ontologiset käsitesuhteet	25
2.4	Termi	28
2.4.1	Termin määritelmän häilyvyys	29
2.4.2	Termin määritelmä tässä tutkielmassa	31
2.5	Systemaattinen käsiteanalyysi	32
2.5.1	Käsiteanalyysin tasot	33
2.5.2	Satelliittimalli	34
2.6	Termianalyysi	36
2.6.1	Hyvän termin kriteerit	36
2.6.2	Termin muodostus	40
3	Tekoäly	41
3.1	Tekoälyn määritelmä	41
3.2	Generatiivinen tekoäly	47
3.3	Kielimalli	48
3.4	Tekoälyagentit ja agenttinen tekoäly	48
3.5	Tekoälyn tutkimista journalismin kontekstissa	50
3.6	Terminologista tutkimusta tekoälystä	54
4	Tekoälyaiheisten verkkoartikkelien käsitteiden analyysi	57
4.1	Terminpoiminnan toteuttaminen	57

4.2	Satelliittimallin muodostaminen	59
4.3	Satelliittimalli verkkoartikkelien tekoälykäsitteistä	61
4.4	Satelliittimallin geneeriset käsitesuhteet	63
4.4.1	Laaja kielimalli satelliittimallissa	71
4.4.2	Generatiivinen tekoäly satelliittimallissa	73
4.5	Satelliittimallin toimintasuhteet	77
4.5.1	Toimijat	77
4.5.2	Syyt	80
4.5.3	Välineet	80
4.5.4	Paikat	81
4.5.5	Tulokset	82
4.6	Satelliittimallin muut käsitesuhteet	83
4.6.1	Ainesosa- ja koostumussuhteet	84
4.6.2	Ominaisuussuhde	85
4.6.3	Riippuvuus- ja korrelaatio-suhde	86
4.6.4	Sijaintisuhde	87
4.7	Yhteenveto käsiteanalyysin tuloksista	88
5	Termianalyysi tekoälytermeistä	93
5.1	Omakielisyyys	94
5.2	Läpinäkyvyys	95
5.3	Johdonmukaisuus	96
5.4	Rakenne ja käyttö	97
5.5	Yhteenveto termianalyysin tuloksista	98
6	Päätäntö	102
	Lähteet	107

Kuviot

Kuvio 1. Esimerkki satelliittimallin merkintätavoista.	61
Kuvio 2. Kokonaiskuva aineiston tekoälykäsitteistä.	62
Kuvio 3. Tekoälykäsitteiden geneeriset käsitesuhteet.	64
Kuvio 4. Supistettu versio geneerisistä käsitesuhteista.	65
Kuvio 5. Kielimallin ja laajan kielimallin käsitesuhteet.	71
Kuvio 6. Generatiivisen tekoälyn käsitesuhteet.	74
Kuvio 7. Tekoälyn toimintasuhteet.	78
Kuvio 8. Muut käsitesuhteet.	84

Taulukot

Taulukko 1. Tutkimuskysymykset, -aineisto, -menetelmä ja -tulos.	15
---	----

1 Johdanto

Tekoäly on viime vuosina noussut julkisen keskustelun yhdeksi merkittäväksi puheenaiheeksi. Yleisesti tekoäly käsitteenä tarkoittaa teknologiaa, jolla pyritään jäljittelemään ihmisenkaltaista älykkyyttä ja siihen liittyviä toimintoja, kuten kielen tuottamista (ks. esim. Haleem ja muut, 2022, s. 120; Sanastokeskus, n.d.b). Tekoäly käsitteenä ei koostu yhdestä tietystä teknologiasta vaan sen taustalla on pitkä historia useita eri menetelmiä ja teknologioita. Juurensa tekoäly juontaa viime vuosisadan puoliväliin. Vuonna 1950 Alan Turing julkaisi artikkelin, joka käsitteli älykkäiden koneiden luomista ja niiden älykkyyden testaamista (Haenlein & Kaplan, 2019, s. 6–7), ja tekoäly terminä itsessään ilmeni ensimmäisen kerran vuonna 1956 järjestetyn Dartmouth Summer Research Project on Artificial Intelligence -tiedekonferenssin yhteydessä.

Viimeaikainen julkinen tekoälykeskuselu on kuitenkin seurausta nykymuotoisen generatiivisen tekoälyn kehityksen tuottamien tekoälysovellusten yleistymisestä ja niiden käyttöönotosta esimerkiksi työpaikoilla ja kouluissa. Generatiivisella tekoälyllä tarkoitetaan teknologiaa, joka hyödyntää syväoppimismalleja ihmisenkaltaisen sisällön, kuten kuvien ja tekstien tuottamiseen kehoitteiden, kuten ohjeiden ja kysymysten, perusteella (Lim ja muut, 2023, s. 2). Suositut tekoälysovellukset, esimerkiksi ChatGPT, Copilot, Claude ja Gemini ovat niin sanottuja laajoja kielimalleja, generatiivisen tekoälyn alalajeja, jotka pyrkivät simuloimaan ihmisenkaltaisia ajatusprosesseja luonnollista kieltä tuottamalla (Chen & Chan, 2024, s. 9101, 9103). Kansalainen voi kohdata generatiivisen tekoälyn arjessa esimerkiksi sosiaalisessa mediassa erilaisten tekoälysisältöjen kuluttajana tai itse tekoälysovelluksen käyttäjänä esimerkiksi sisällöntuotannon apuvälineenä tai keskustelukumppanina.

Vaikka tekoäly on saanut alkunsa jo yli 70 vuotta sitten, nykymuotoinen generatiivinen tekoäly ja sen sovellukset ovat aihepiirinä ajankohtaisia ja se on teknologiana jatkuvassa kehityksessä. Tekoälystä kirjoitetaan erikoisanalan julkaisujen ja tekoälytutkimusten ohella runsaasti myös yleistajuiseksi tarkoitetuissa laajalle yleisölle suunnatuissa uutislähteissä, kuten Ylen verkkoartikkeleissa. Tekoälystä, sen käytöstä ja vaikutuksista kirjoitetaan

paljon yleistajuisissa medioissa, koska aiheena se koskettaa monen ihmisen elämää eri tavoilla esimerkiksi koulutukseen liittyen tai työelämän eri aloilla ja työtehtävissä. Yhtäältä tekoälyn kerrotaan helpottavan ja tehostavan työprosesseja. Toisaalta tekoälyn pelätään pahimmassa tapauksessa aiheuttavan työttömyyttä korvaamalla ihmistyöntekijät kokonaan tai vähintään pienentämällä ihmistyöntekijöiden tarvetta merkittävästi. Yhtä kaikki, tekoälyn kehitys ja sen soveltaminen yhteiskunnan eri kehyksissä mullistavat jatkuvasti lähes jokaisen ihmisen elämää.

Tekoäly on siten ilmiö, johon sopeutuminen ja jonka ymmärtäminen koskettavat isoa joukkoa kansalaisista – tekoälyn vaikutukset ja tekoälystä uutisoitavat aihepiirit tekevät tekoälystä kaikkia koskettavan tutkimuskohteen. Kansalaisten on tästä näkökulmasta katsoen tärkeää ymmärtää, minkälaisista käsitteistä tekoälystä viestittäessä puhutaan ja minkälaisilla termeillä näistä käsitteistä viestitään yleistajuisissa teksteissä. Siksi olen tässä tutkielmassa keskittynyt tarkastelemaan tekoälystä kertovia yleistajuisia verkkoartikkeleita terminologisesta näkökulmasta.

1.1 Tavoite

Tutkimukseni tavoitteena on selvittää, millaisia tekoälykäsitteitä ja -termejä ilmenee Ylen yleistajuisissa tekoälyaiheisissa verkkoartikkeleissa. Käsitteellä tarkoitetaan jostakin ilmiöstä, kuten tekoälystä, ajatuksella muodostettua mielikuvaa (Suonuuti, 2024, s. 11) ja termillä tarkoitetaan erikoisalan käsitteen nimitystä (Sanastokeskus TSK, 2006, s. 22). Tutkielmassa tekoälyä tarkastelen siitä näkökulmasta, miten Ylen verkkoartikkelit rakentavat tekoälyä koskevaa käsitteistöä ja millä tavoin käytetty termistö voi mahdollisesti edistää tai vaikeuttaa ilmiön ymmärrettävyyttä. Tutkimustani ohjaavat kaksi tutkimuskysymystä:

1. Mitä tekoälykäsitteitä ja -termejä ilmenee Ylen tekoälyä käsittelevissä verkkoartikkeliteksteissä?
2. Miten tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit?

Aiempi tutkimus siitä, miten tekoölyaiheita käsitellään journalismissa, on keskittynyt tutkimaan tekoölyä etenkin kehysten ja diskurssien näkökulmasta (ks. esim. Duberry & Hamidi 2021; Hollanek ja muut, 2025; Nguyen, 2023; Nguyen & Hekman, 2024; Xian ja muut, 2024). Tämän painotuksen takia olen valinnut aikaisemmasta tutkimuksesta poikkeavan terminologisen tutkimuksen lähestymistavan tekoölyaiheisen journalismin tutkimiseen, koska sillä on mahdollista luoda uutta tietoa tutkimusaiheesta. Terminologisella tutkimuksella tarkoitetaan erikoisalojen käsitteiden ja niiden termien tutkimiseen keskittyvää tieteenalaa (Cabr , 1999, s. 37; Nuopponen, 1999, s. 1; Sager, 1990, s. 2). Vaikka aiempi tekoölyaiheisen journalismin tutkimus ei keskityk n niink n terminologiseen n k kulmaan, on silti huomioitava, ett  terminologisessa tutkimuksessa teko ly  ja sen mahdollisuuksia on k sitelty pidemp n kuin monella muulla alalla kauan ennen suurten kielimallien l pimurtoa (ks. esim. Ahmad, 2001; Nedobity, 1985), ja tutkimusta tehd n edelleen (ks. esim. Lewandowska-Tomaszczyk & Pawlowski, 2025).

Teko lyk sitteiden ja termien k sitteily  mediassa on t rke   tutkia terminologisesta n k kulmasta siksi, ett  kansalaisten on t rke  ymm rt   teko ly  ja n en, ett  journalismilla on oma roolinsa kansalaisten teko lylukutaidon kehitt misess  sek  yll pit misess . Teko lylukutaidolla tarkoitan yleist  ymm rryst  teko lyn toiminnasta ja kyky  arvioida teko lyn yhteiskunnallisia vaikutuksia (ks. esim. Hollanek ja muut, 2025, s. 6935). Teko lylukutaidon m  ritelm n voidaan toisaalta sis llytt   my s kyky k ytt   teko ly  tehokkaalla ja vastuullisella tavalla (Hollanek ja muut, 2025, s. 6935). Aikaisemmassa journalismin kontekstissa tehdyss  teko lytutkimuksessa korostetaankin journalismin t rke   asemaa tiedon levitt misess  ja yleis n ymm rryksen muokkaamisessa sek  yll pit misess  uusien teknologioiden kuten generatiivisen teko lyn kohdalla (ks. esim. Xian ja muut, 2024, s. 1). Journalismilla ja medialla n hd n olevan keskeinen rooli teko lyn merkityksen muodostamisessa yhteiskunnassa, koska media rakentaa aktiivisesti diskursseja teko lyst  muun muassa teko lyn hy tyjen ja riskien n k kulmista (ks. esim. Nguyen & Hekman, 2024, s. 438–439).

Tekoälyä on mielekästä tutkia terminologisesta näkökulmasta myös siksi, että suomen kielellä tehtyä terminologista tutkimusta tekoälystä on olemassa vasta verrattain vähän. Englannin kielellä toteutettua tutkimusta tekoälyn terminologiasta on runsaammin ja sitä on tehty useista eri näkökulmista, esimerkiksi lääketieteen, psykologian, sosiaalisten tieteiden, ihmistermien käytön eettisyyden tai toimijuuden näkökulmista (ks. esim. Cao ja muut, 2025; Deroy, 2023; Haskell & Clark, 2025).

Suomalaisena tekoälyn terminologisena tutkimuksena voidaan pitää esimerkiksi Sanastokeskuksen Tietotekniikan termitalkoot -projektissa syntyneitä tekoälykäsitteiden käsitteitä (ks. Sanastokeskus, n.d.b), jonka muodostamisessa on analysoitu tekoälykäsitteiden sisältöjä, niiden termejä ja annettu termisuosituksia. Suomalaisesta terminologisesta tekoälytutkimuksesta voidaan nostaa myös esiin esimerkiksi Laatun (2025) pro gradu -tutkielma, joka on kirjallisuuskatsaus generatiivisen tekoälyn hallusinaatiokäsitteestä, hallusinaatio-termin muodostumisesta sekä eri termivaihtoehtojen arvioinnista. Laatun (2025) kirjallisuuskatsaus edustaa oman tutkielmani aihetta ja suuntausta lähellä olevaa suomalaista tutkimusta, mutta lähestyy tekoälyn terminologiaa eri näkökulmasta erilaisen aineiston sekä tiukemman rajauksen suhteen keskittyen ainoastaan tekoälyn hallusinaation käsitteeseen.

Voidaan siis argumentoida, että tutkielmani aineisto ja näkökulma luovat uutta tietoa aiheesta. Vaikka tekoälykäsitteiden terminologista muodostusta ja laatua on jo tutkittu vieraskielisissä tutkimuksissa (ks. esim. Alshtaiwi, 2025; Graziani ja muut, 2023; Helfrich, 2024; Kozak, 2026), vastaavaa tutkimusta on syytä tehdä myös suomeksi. Kuten myöhemmin 1.3 luvussa esittelen, muun muassa läpinäkyvyys, omakielisyys ja kieliopillinen moitteettomuus ovat hyvän termin kriteereitä. Näin ollen esimerkiksi englanninkielisiä tutkimustuloksia tekoälytermien laadusta ei voida täysin suoraan soveltaa suomen kieleen, sillä etenkin termin omakielisyys ja kieliopillinen moitteettomuus (ks. esim. Nykänen, 1997) asettavat tekoälytermeille vaatimuksia, jotka pätevät ainoastaan käyttökielen, tässä tapauksessa suomen kielen, kontekstissa.

Lisäksi käytännön tarvetta tekoälykäsitteiden ja termien selvittämiseksi edustaa osin myös Kauppisen (2025) kolumni Aamulehdessä, jossa kirjoittaja nostaa esiin tekoäly-termin käytön epämääräisyyden. Kolumnissaan Kauppinen (2025) toteaa, että tekoäly-termin käyttö on merkitykseltään epäselvää, sillä esimerkiksi lehtijutuista on pääteltävissä, että tekoäly-termillä saatetaan tarkoittaa muun muassa laajoja kielimalleja, koneoppimista tai tavallista sääntöihin pohjautuvaa koodiakin. Kauppinen (2025) argumentoi, että tämä voi luoda virheellisiä mielikuvia lukijalle, mistä hän tarjoaa esimerkkinä tilanteen, jossa tutkija luo uuden antibiootin tekoälyllä; lukija voi ymmärtää tilanteen virheellisesti niin, että tutkija antaisi kehoitteita generatiiviselle tekoälylle, vaikka kyse on tosiasiaassa koneoppimisen hyödyntämisestä antibiootin kehittämisessä. Kolumnissaan Kauppinen (2025) peräänkuuluttaa tarvetta selkeyttää tekoälyn merkityssisältöä yritysviestinnässä sekä epäsuorasti myös median vastuuta kyseisen merkityssisällön selkeyttämisessä.

Oletuksena tutkielmassani on se, että tekoälykäsitteistä kirjoitettaessa käytetään vieraskielisiä tai vieraaseen kieleen pohjautuvia termejä. Tämä perustuu olettamalleni siitä, että tekoäly on nopeasti kehittyvä teknologian erikoisala, jossa syntyy nopeaan tahtiin uusia käsitteitä, jotka tarvitsevat myös uusien termien luomista (ks. Cabré, 1999, s. 4). Eryityisesti IT-alalla englannin kielellä on hallitseva asema, minkä takia monet termit ovat englanniksi (Plechko ja muut, 2022, s. 1), eli voidaan olettaa vieraskielisten termien käytön olevan myös tekoälyalalle tyypillinen piirre. Tästä syystä odotan analyysin tuloksilta sitä, että aineisto sisältää tekoälykäsitteistä käytettäviä termejä, jotka ovat ongelmallisia omakielisyyden näkökulmasta ja sitä kautta myös läpinäkyvyyden näkökulmasta, mikä voi mahdollisesti vaikeuttaa ilmiön ymmärtämistä.

1.2 Aineisto

Tutkimukseni aineistoksi olen valinnut vuonna 2025 julkaistut Ylen verkkoartikkelit, joiden tekstit käsittelevät tekoälyä. Aineistoon olen hyväksynyt artikkelit, jotka on julkaistu Ylen verkkosivuilla vuonna 2025, ja jotka käsittelevät tekoälyä jollain tapaa. Hyväksyin Ylen tekoälyartikkelit tutkimuksen aineistoon käsittelykontekstista riippumatta. Tämä

tarkoittaa sitä, että tarkastelun kohteeksi on valikoitunut esimerkiksi tekoälyteknologiaa koskevat uutiset, sen soveltamista koskevat oppaat ja tekoälyn vaikutuksia pohtivat kolumnit.

Pelkkä tekoäly-sanan ilmeneminen artikkelissa ei ole kuitenkaan tutkielmani kontekstissa riittävä aineistonvalintakriteeri – Yle käyttää useissa verkkoartikkeleissaan tekoälyllä luotua tiivistelmää, jossa on maininta siitä, että tiivistelmä on tuotettu tekoälyn avulla. Uutisissa saattaa siis esiintyä tekoäly-sana ilman, että itse uutinen kuitenkaan käsittelee tekoälyä aiheena millään tavalla. Tutkimuksen ulkopuolelle jäivät myös artikkeleissa mahdollisesti esiintyvät erilaiset kuvat ja videot tutkimuksen analyysin keskittyessä tekstisisältöön.

Tekoälyaiheiset artikkelit etsin aineistooni Ylen tekoäly-aihetunnisteella merkittyjen verkkoartikkelien perusteella. Tekoäly-aihetunnisteella merkityt Ylen artikkelit on löydettävissä linkistä <https://yle.fi/t/18-215158/fi> tai hakemalla artikkeleita hakusanalla *tekoäly*. Toteuttamani aineistonkeräyksen perusteella aineistoon valikoitui yhteensä 179 tekoälyä aiheenaan käsittelevää verkkoartikkeliä. Aineistosta valtaosa koostuu uutisista, mutta mukana on myös esimerkiksi oppaiksi, analykseiksi, haastatteluiksi ja kolumneiksi tunnistettavia artikkeleita.

Olen valinnut Ylen tekoälystä kertovat verkkoartikkelit tutkimukseni kohteeksi kahdesta syystä. Pääasiallinen syy valinnalle liittyy Ylen vastuuseen valtiollisena toimijana. Valtioomisteisena julkista palvelua tarjoavana yhtiönä Ylen lakisääteisiin tehtäviin kuuluvat muun muassa kotimaisen kulttuurin tuottaminen, luominen, kehittäminen ja säilyttäminen (Yle, n.d.). Kieli on yksi merkittävimmistä kulttuurin osa-alueista. Voidaan argumentoida, että yksityisomisteisiin media-yhtiöihin verrattuna Ylellä on suurempi vastuu tuottaa hyvää kieltä ja ylläpitää hyvän kielen säilymistä. Tämän vastuun piiriin voidaan siten laskea myös Ylen verkkoartikkeleissaan käyttämät termit, kuten tutkimukseni kohteena olevien tekoälykäsitteiden termit. Tätä valintaperustetta tukee myös esimerkiksi se, että Yle on vastineessaan Helsingin Sanomissa julkaistulle mielipidekirjoitukselle kiteyttänyt

Ylen tehtävän erityisesti kieleen, kulttuuriin ja demokratiaan (Hirvonen, 2018). Vastineessaan silloinen Ylen nuorisomedioiden päällikkö Antti Hirvonen korostaa, että hyvän ja elävän kielen käyttö, vaaliminen ja ylläpito ovat Ylen tehtävän ytimessä.

Toisaalta samassa vastineessa Hirvonen (2018) puolustaa osittain Ylen hyvän ja elävän kielen tavoitteiden vastaisestikin esimerkiksi englannin kielestä kantautuvien lainasanojen, kuten *FOMO:n* (*fear of missing out* eli paitsi jäämisen pelko) käyttöä nuorille suunnatuissa sisällöissä. Tätä perustellaan sillä, että käsitteiden nimien suomentaminen menettäisi sanojen painoarvon, tunnelatauksen ja tunnistettavuuden. Yksi hyvän termin vaatimuksista on kuitenkin omakielisyys (ks. esim. Isohella & Nuopponen, 2016, s. 229). Vastineessa ilmenevästä ristiriitaisuudesta huolimatta voidaan väittää, että Ylellä on vastuu tuottaa hyvää kieltä sisältäviä tekstisisältöjä ja ylläpitää hyvän, elävän kielen säilymistä.

Toissijainen, mutta silti merkittävä syy Ylen verkkoartikkelien valikoitumiselle on se, että kaupallisiin toimijoihin verrattuna Yle ei aseta verkkoartikkeleitaan erillisen maksumuurin taakse. Tämä erilaisten uutisten, oppaiden ja kolumnien vapaa saatavuus mahdollistaa sen, että Ylen verkkoartikkelit tavoittavat potentiaalisesti isomman yleisön kuin kaupallisen median verkkoartikkelit, joista osa on lukittu maksujärjestelmän taakse. Tämä aineiston vapaa saatavuus tekee myös tutkimuksestani helpommin toistettavamman, kun artikkelit ovat kaikkien saavutettavissa niin kauan kuin Yle säilyttää niitä omalla sivustollaan.

Aineistovalinnan kohdalla on otettava huomioon se, että aineiston edustama tekstityyppi, eli yleistajuiset informointiin pyrkivät tekstit, asettavat omat reunaehdonsa artikkelien sisällölle. Yleistajuisessa viestinnässä, kuten uutisissa, erikoisalaan liittyvä erikoistunut tieto pyritään kertomaan usein tiukasti rajatun tekstimäärän rajoissa siten, että tekstin sisältö on mahdollisimman suuren lukijayleisön ymmärrettävissä ilman erikoisalan osaamista (Raevaara, 2016, luku 3). Yleistajuisissa teksteissä kirjoittaja on harvemmin erikoisalan asiantuntija (Raevaara, 2016, luku 3), mutta journalistin työhön voidaan

katsoa kuuluvaksi riittävä ymmärrys esimerkiksi tekoälyaiheista, jotta journalisti kykenee esittämään monimutkaiset kokonaisuudet ymmärrettävässä muodossa maallikolle (Dandurand ja muut, 2023, s. 4). Yleistajuisten tekstien vastakohtana voidaan pitää asiantuntijatekstejä, jotka määrittelen viestinnäksi, jossa erikoisalan asiantuntijat viestivät toisilleen erikoisalaan liittyvästä tiedosta ja käyttävät termejä, joiden merkityssisältö on pääasiassa kaikille osapuolille tuttua ilman erillistä selvitystä (Gotti, 2011, s. 17–18).

Toisin sanoen Ylen verkkoartikkelit edustavat yleistajuisia tekstejä, joita kirjoitettaessa on otettu huomioon yleisön laaja tausta. Tämä tarkoittaa, että Ylen tekoälyaiheiset verkkoartikkelit on kirjoitettu tavalla, jolla myös sellainen lukija, jolla ei ole aikaisempaa tekoälyn erikoisalan tietämystä taustalla, kykenee ymmärtämään artikkelissa käsitellyt asiat. Tämän lähtökohdan katson osaltaan asettavan oletuksen myös siitä, että aineistoissa käytetyt tekoälytermit tulisi myös olla ymmärrettäviä lukijalle.

Aineistoon keräämistäni verkkoartikkeleista muodostin yhtenäisen tekstimassan eli korpuksen verkkosivupohjaisella Sketch Engine -ohjelmalla, joka on korpuksen hallintaan ja tekstin analyysiin keskittynyt analyysityökalu. Verkkoartikkeleista luomani korpuksen sanakoko on yhteensä 77630 sanaa. Korpuksen luomisen jälkeen suoritin korpukselle Sketch Enginellä puoliautomaattisen terminpoininnan termilistojen muodostamiseksi. Puoliautomaattinen terminpoininta tarkoittaa aineistosta tietokoneavusteisesti toteutettavaa termilistauksen tuottamista, josta tutkija sitten poistaa manuaalisesti sellaiset kielelliset elementit, joita ei lasketa termiksi tutkimuksen kontekstissa (Pasanen, 2009, s. 43–44). Puoliautomaattisessa terminpoininnassa ohjelma poimii korpuksesta termejä termilistaan erilaisiin sääntöihin pohjautuvien algoritmien perusteella (ks. Pasanen, 2009, s. 47, 60). Tämä eroaa perinteisemmästä manuaalisesta terminpoininnasta, jossa tutkija omaa intuitiota käyttäen poimii itse aineistosta termilistaan ne sanat, joita pitää termeinä, jolloin termilistaan ei myöskään päädy ylimääräisiä sanoja tutkijan siivottavaksi (Pasanen, 2009, s. 44–45).

Puoliautomaattisesti toteutettuna terminpoiminta vaatii siten tietokoneen, tietokoneohjelman, jolla poiminta suoritetaan, aineiston tallennettuna sähköisessä muodossa, minkä lisäksi aineistotekstiä on esivalmisteltava terminpoimintaohjelmalle käytettäväksi sopivaan muotoon (Pasanen, 2009, s. 49). Puoliautomaattisen terminpoiminnan etuja ovat muun muassa sen nopeus ja kyky toteuttaa sille ohjelmoidut tehtävät luotettavasti. Puoliautomaattisen terminpoiminnan heikkoutena on se, että toisin kuin manuaalisessa terminpoiminnassa, terminpoimintaohjelma kerää termilistaan paljon tutkijan tarkoitukseen kelpaamattomia termiehdokkaita ja samalla ohjelma saattaa jättää poimimatta termejä, jotka tutkija muuten hyväksyisi termilistaansa (Pasanen, 2009, s. 49–50).

Tietokoneohjelman suorittaman terminpoiminnan jälkeen siivosin termilistan ylimääräisistä sanoista, joita en pitänyt tekoälyterminä tämän tutkielman kontekstissa, jotta termilistaan ei päätyisi ylimääräisiä kielellisiä elementtejä (Pasanen, 2009, s. 45). Termilistan siivoamiseen käytin manuaaliselle terminpoiminnalle ominaista tutkijan oman tulkinnan käyttöä termistä (Pasanen, 2009, s. 44). Siihen, mikä termiksi poimitaan, vaikuttaa tutkijan tietämys aiheesta, konteksti sekä tavoitteet. Termien siivoamisen jälkeen lopulliseen termilistaan päätyi 51 yksisanaista termiä ja 10 monisanaista termiä: yhteensä 61 tekoälytermiä.

1.3 Menetelmä

Tutkielmani on luonteeltaan kvalitatiivinen eli laadullinen tutkimus. Laadullinen tutkimus keskittyy tarkastelemaan ihmisten keskeisestä sosiaalisesta toiminnasta syntyviä merkityksiä ja merkityssuhteita, jotka muodostavat merkityskokonaisuuksia, esimerkiksi erilaisia tapahtumia, toimintoja, ajatuksia tai rakenteita (Vilka, 2025, luku 5). Laadullisessa tutkimuksessa tutkija pyrkii oman tulkintansa avulla luomaan ymmärrystä tai tuomaan aineistosta esiin jotakin, mikä ei ole välittömän aistihavainnon saavutettavissa.

Tutkielmassani pyrin saamaan vastaukset tutkimuskysymyksiini kahta terminologista analyysimenetelmää käyttäen, jotka ovat systemaattinen käsiteanalyysi sekä

termianalyysi. Taulukossa 1 kuvaan menetelmiä, joilla vastaan tutkimuskysymyksiin sekä niitä tuloksia, joita aineiston analyysin seurauksena on odotettavissa.

Taulukko 1. Tutkimuskysymykset, -aineisto, -menetelmä ja -tulos.

Tutkimuskysymys	Aineisto	Menetelmä	Tulos
1. Mitä tekoälykäsitteitä ja -termejä ilmenee Ylen tekoälyä käsittelevissä verkkoartikkeliteksteissä?	Vuonna 2025 julkaistut Ylen tekoälyaiheiset verkkoartikkelit	Systemaattinen käsiteanalyysi	Satelliittimalli eli graafinen kuvaus aineistosta havaittavista tekoälykäsitteistä ja niiden muodostamasta käsitejärjestelmästä
2. Miten tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit?	Läpinäkyvyyden ja omakielisyyden näkökulmista ongelmallisimmat tekoälytermit luomas-tani satelliittimal-lista	Termianalyysi	Ongelmallisten tekoälytermien arviointi ja vaihtoehtoisten termien muodostaminen

Ensimmäiseen tutkimuskysymykseen pyrin saamaan vastauksen systemaattisella käsiteanalyysillä. Systemaattinen käsiteanalyysi on käsitteiden sisältöjen ja käsitesuhteiden selvittämiseen käytetty analyysimenetelmä (ks. Nuopponen, 2020, s. 97). Teen puoliau-tomaattisella terminpoiminnalla aineistosta muodostamieni termilistojen pohjalta aineiston kokonaiskäsitejärjestelmää kuvaavan graafisen satelliittimallin, johon sijoitan tekoälyn pääkäsitteeseen erilaisilla käsitesuhteilla liittyvät tekoälykäsitteet. Satelliittimallin muodostan Miro-ajatuskarttaohjelmalla. Käsiteanalyysillä muodostamani satelliittimalli kuvastaa aineistossa ilmeneviä tekoälyn alakäsitteitä sekä käsitteiden välisiä suhteita. Systemaattinen käsiteanalyysi vastaa siten ensimmäiseen tutkimuskysymykseen siitä, mitä tekoälykäsitteitä ja -termejä aineistossa ilmenee. Esittelen systemaattisen käsiteanalyysin tarkemmin luvussa 2.5 ja satelliittimallin luvussa 2.5.2.

Systemaattisen käsiteanalyysin valintaa menetelmäksi ohjaavat kaksi syytä. Pääasiallinen syy on se, että käsiteanalyysillä on mahdollista muodostaa aineiston tekoälykäsitteistä kokonaiskäsitejärjestelmä, joka tuo järjestelmällisesti esiin aineistosta löydettyjen

käsitteiden suhteen sekä itse tekoälyn pääkäsitteeseen että muihin ylä-, ala- ja vieruskäsitteisiin. Näiden käsitesuhteiden selvittäminen ja ymmärtäminen on myös oleellista termianalyysissä termin laadun analysoinnin kannalta, mikä siten tukee termianalyysin suorittamista. Toinen syy käsiteanalyysin valintaan liittyy tutkielmassani käyttämäni termin määritelmään. Käytän tutkielmassani termistä määritelmää, joka poikkeaa perinteisestä terminologian tutkimuksen mukaisesta termin määritelmästä. Oman termin määritelmäni kannalta onkin oleellista todistaa käsiteanalyysillä käsitteiden suhde tekoälyn erikoisalaan. Omaan termin määritelmäni syvennyn luvussa 2.4.2.

Toiseen tutkimuskysymykseen vastaan termianalyysillä. Termianalyysin valintaan vaikutti se, että termianalyysi on analyysimenetelmä, jolla voidaan muun muassa tarkastella ja arvioida käsitteen suhdetta termiin sekä termin laatua (ks. Nuopponen & Pilke, 2016, s. 18), mikä tekee menetelmästä sopivan tämän tutkielman tarpeisiin. Termianalyysin avulla tutkin aineistossa käytettyjen tekoälytermien laatua hyvän termin kriteerien näkökulmasta. Valitsen aineistosta muodostamastani satelliittimallista omakielisyyden ja läpinäkyvyyden osalta ongelmallisimmat termit, joille toteutan tarkemman termianalyysin. Terminmuodostamiseen tarkoitettujen hyvän termin kriteerien mukaisesti hyvän termin tulee olla omakielisyyden ja läpinäkyvyyden lisäksi muun muassa johdonmukainen ja erottuva (Isohella & Nuopponen, 2016, s. 229–232; Nykänen, 1997). Esittelen hyvän termin kriteerit tarkemmin luvussa 2.6.1. Termianalyysin päätteeksi muodostan analysoituille tekoälytermeille vaihtoehtoiksi hyvän termin kriteereihin pohjautuvia termivaihtoehtoja.

2 Terminologinen tutkimus

Tutkielmassani keskityn selvittämään tekoälykäsitteitä ja arvioimaan tekoälytermien laadukkuutta Ylen verkkoartikkeleissa, ja tutkimuskysymyksiini on mahdollista saada vastauksia systemaattisella käsiteanalyysillä ja termianalyysillä. Tutkimuskysymykset ja niihin vastaavat menetelmät asemoivat siten tutkimukseni teoreettisen viitekehyksen vahvasti terminologisen tutkimuksen piiriin. Tässä pääluvussa keskityn taustoittamaan terminologista tutkimusta sekä tutkielmalleni olennaisia terminologisessa tutkimuksessa esiintyviä yleisiä käsitteitä, menetelmiä ja työkaluja.

Terminologinen tutkimus on tieteenala, joka keskittyy tunnistamaan erikoisaloihin liittyvien todellisuuksien osioita tunnistamalla erikoisalatiedon käsitteitä, antamalla niille määritelmiä ja nimeämällä niitä termeillä (ks. esim. Cabré, 1999, s. 37; Nuopponen, 1999, s. 1; Sager, 1990, s. 2), ja se on kehittynyt tarpeesta ohjata erikoisalojen kielenkäyttöä standardoimalla erikoisalojen sanastoa (Nuopponen, 1999, s. 1). Terminologisen tutkimuksen perustana on pyrkiä erottamaan käsite sen tarkoitteesta ja termistä ja selvittämään, millä tavalla tutkimuksen kohteena olevan ilmiön käsitteet, termit sekä käsitejärjestelmät muodostuvat ja minkälaisia suhteita käsitteiden välillä on (Nuopponen, 1999, s. 1–2; Nuopponen, 2020, s. 95, 97–103). Olennaisena osana terminologista tutkimusta on myös tarkastella käsitteiden määrittelyä ja termien laatua niiden taustalla vallitsevien periaatteiden näkökulmasta (Nuopponen, 1999, s. 2; Nuopponen, 2020, s. 99–100).

On hyvä huomioda, että itse terminologia-termillä on polyseeminen merkitys (ks. esim. Nuopponen, 2020, s. 95; Sager, 1990, s. 3). Nuopponen (2020, s. 95) mukaan terminologialla tarkoitetaan tyypillisesti erikoisalan termeistä koottua sanastoa, mutta terminologialla voidaan myös tarkoittaa terminologista tutkimusta toimintana. Sagerin (1990, s. 3) mukaan terminologialla voidaan lisäksi viitata terminologisen tutkimuksen taustalla oleviin teorioihin. Tämän moniselitteisyyden takia Nuopponen (2020, s. 95) suosittelee käyttämään tarkempia ilmaisuja: terminologia-alasta puhuttaessa kontekstin mukaan joko terminologiaoppi, terminologinen tutkimus tai terminologian teoria, ja alan käytännön toiminnasta puhuttaessa sanastotyö, termityö tai terminologiatyö.

Terminologisen tutkimuksen teorioita ja menetelmiä hyödynnetään erikoisalan tiedon analysointiin, jonka perusteella erikoisalan tiedosta on mahdollista muodostaa järjestelmällinen yleiskuva esitettävään muotoon (Nuopponen, 1999, s. 3). Käyttötarkoituksia terminologian teorialle on useita ja niistä hyötyvät sanastotyöntekijöiden, kääntäjien ja viestinnän opiskelijoiden ohella myös muut erikoisalan tiedon parissa työskentelevät (Nuopponen, 1999, s. 1–7). Esimerkiksi erilaisissa viestinnän konteksteissa työskentelevät ammattilaiset, kuten toimittajat, voivat hyötyä taidosta viestiä lukijalle selkeällä tavalla käsittelyn kohteena olevan erikoisalatiedon taustalla vallitsevista käsitteistä ja käsittejärjestelmistä, luoda kuvaavia määritelmiä tai muodostaa uusia termejä (Nuopponen, 1999, s. 3–5).

Seuraavaksi esittelen terminologiselle tutkimukselle olennaiset pääkäsitteet aloittaen tarkoitteesta ja käsitteestä, jonka jälkeen esittelen käsitepiirteet, käsitesuhteet ja käsittejärjestelmät sekä termin. Näiden tutkimuksen pääkäsitteiden esittelemisen jälkeen siirryn taustoittamaan tässä tutkielmassa käyttämiäni terminologisen tutkimuksen tutkimusmenetelmiä, eli systemaattista käsiteanalyysiä, sen tuloksena syntyvää satelliittimalia sekä termianalyysiä hyvän termin kriteereineen.

2.1 Tarkoite ja käsite

Tarkoite on reaali maailman konkreettinen tai abstrakti todellinen tai kuviteltu ilmiö, joka on mahdollista osoittaa, käsittää tai kuvitella (Nuopponen, 2020, s. 97; Sanastokeskus TSK, 2006, s. 10; Suonuuti, 2024, s. 11), ja joka on reaali maailmassa oleva vastine käsitteelle (Sanastokeskus TSK, 2006, s. 10). Tässä tutkielmassa keskeisenä tarkoitteena oleva tekoälyn tarkoite on ilmiönä aineeton, mutta esimerkiksi tekoälyn tiettyjä vaikutuksia on mahdollista havaita aisteilla fyysisesti. Muun muassa generatiivisen tekoälyn avulla tuotettuja kuvia tai tekstejä voidaan tarkastella myös fyysisinä objekteina varsinkin, jos ne on painettu fyysiseen formaattiin, kuten sanomalehtien sivuille.

Käsite on ihmisen ajatteluprosessin seurauksena abstrahoinnilla muodostunut ajattelun tai tiedon yksikkö (ks. esim. Cabré, 1999, s. 42; Nuopponen, 2020, s. 97; Sager, 1990, s. 22), joka on tarkoitteesta johdettu mielikuva ihmisen ajattelussa (Suonuuti, 2024, s. 11). Abstrahoinnilla tarkoitetaan ihmisen mielen toimintaa, jossa ajattelulla pyritään erottamaan reaali maailman konkreettinen tai abstraktinen olio sen muodosta (Tieteen termipankki, n.d.a) eli ajattelemalla pyritään ikään kuin tuomaan esiin tarkoitteen keskeinen luonne, jotta siitä voidaan muodostaa ajatuksella käsiteltäviä mielikuvia. Tekoäly käsitteenä on siten ihmisten ajattelun tasolla muodostunut käsitys tai mielikuva tekoälystä ilmiönä. Terminologisessa tutkimuksessa käsitteillä viitataan erityisesti erikoisalan tietoon (L'Homme, 2020, s. 24), kuten tutkielmani kohteena olevaan tekoälyalan tietoon. Erikoisaloilla käsitetieto kattaa suuren määrän asiantuntijatietoa, ja saman alan asiantuntijoilla voidaan karkeasti kärjistettynä katsoa olevan samanlainen tarkka ja erikoistunut käsitys erikoisalansa keskeisistä käsitteistä. Tällaisen erikoisalatiedon saavuttaminen edellyttää luonnollisesti asiantuntijatiedon opiskelua ja sen syvällistä ymmärtämistä.

Käsitteen intensiolla eli sisällöllä tarkoitetaan käsitepiirteiden muodostamaa joukkoa, jonka avulla tietty käsite voidaan asettaa osaksi käsitteen ekstensiota eli käsitteen alaa, jolla tarkoitetaan jotain tiettyä käsitettä vastaavista tarkoitteista muodostettua ryhmää (Sager, 1990, s. 24; Sanastokeskus TSK, 2006, s. 10). Käsitepiirteellä tarkoitetaan käsitteen ominaisuutta (ks. Nuopponen, 2020, s. 97). Tietyn käsitteen sisältö on uniikki joukko käsitepiirteitä, joista vähintään yksi käsitepiirre erottaa käsitteen sen lähellä olevasta toisesta käsitteestä (Sager, 1990, s. 24). Käsitteen laajuuteen vaikuttaa käsitepiirteiden määrä; laajempaan alaan viittaavat käsitteet sisältävät suhteessa vähemmän käsitepiirteitä kuin kapeaan alaan viittaavat käsitteet (Sager, 1990, s. 24).

Käsitteet voidaan jakaa yksilökäsitteisiin ja yleiskäsitteisiin (Nuopponen, 2020, s. 98; Sanastokeskus TSK, 2006, s. 11–12; Suonuuti, 2024, s. 11). Sellainen käsite, joka vastaa tiettyyn uniikkiin tarkoitteeseen, on yksilökäsite. Yleiskäsite on sen sijaan sellainen käsite, joka jakaa jonkin tarkoitteiden joukon kanssa yhtenäisiä piirteitä, niin että ne voidaan kategorioida samaan ryhmään kuuluvaksi. Tämän tutkielman kontekstissa

yksilökäsitteenä voidaan pitää esimerkiksi Microsoftin Copilotia, joka on käsitteenä yksittäinen, uniikki generatiivisen tekoälyn ohjelmisto, kun taas generatiivinen tekoäly on käsitteenä laaja-alaisempi yleiskäsite, jonka alle Copilot ja muut samoja määritteleviä käsitepiirteitä sisältävät tekoälysovellukset kuuluvat. Vastaavasti tekoäly on käsitteenä generatiivista tekoälyä vielä laaja-alaisempi yleiskäsite.

Koska käsitteet muodostetaan ihmisen ajattelun tasolla, käsitteisiin liittyy olennaisesti subjektiivinen ulottuvuus. Sager (1990, s. 16–17) tuo esiin, että henkilöt, joilla on toisinaan vastaava tietotaso aiheesta, eivät kuvaile samaa käsitettä täysin samalla tavalla tai täysin samanlaisella tietosisällöllä. Hänen mukaansa se, että joukko voi olla yhtä mieltä käsitteen nimityksestä, on seurausta käsitteen idealisoidusta ja käytännöllisestä yhteisöön sekä erikoisaloihin liittyvästä tietorakenteesta, joka määräytyy sosiaalisten normien mukaan. Sagerin (1990, s. 17) mukaan sosiaalinen normi, johon kaikki yksilöllinen tieto on suhteessa, edustaa yhteisön kollektiivista hiljaista hyväksyntää nauttivaa näkemystä, joka asettaa rajoja käsitteiden sisällöllisille rakenteille yhteisöllisellä tasolla.

Tätä käsitteiden subjektiivista aspektia sivuaa myös Cabré (1999, s. 43), jonka mukaan tutustuessaan erikoisalan tietoon erikoisalan osaajat muodostavat tästä tiedosta vaihtelevia käsitteellisiä rakenteita, toisin sanoen käsitejärjestelmiä. Cabré (1999, s. 43) mukaan nämä vaihtelevat käsitejärjestelmät heijastavat yhden tai useamman erikoisalan osaajan tietämystä aiheesta, minkä takia tarkkarajaisia tai pysyviä käsitejärjestelmiä on vaikea rakentaa, ja käsitejärjestelmien tulkinnassa on aina tilaa laajoille näkökulmille esimerkiksi käsitesuhteista. Myös L'Homme (2020, s. 90) tuo esiin, että käsitteiden tarkastelu eri tiedonalueiden näkökulmista tai tarkastelu saman tiedonalueen sisällä olevien tiettyjen aiheiden näkökulmista vaikuttavat siihen, miten käsitteitä rajataan, kategorisoidaan ja määritellään.

Jos käsitteisiin ja niiden muodostumiseen liittyy tällainen subjektiivinen taso, saman subjektiivisuuden on silloin myös ulotuttava käsiteanalyysiin. Toisin sanoen käsitteisiin liittyvä subjektiivisuuden taso tekee käsitysanalyysin tuloksena muodostetusta

käsitejärjestelmästä myös osittain tutkijan omaan tulkinnanvaraan ja taustaan nojaavan subjektiivisen näkemyksen aineistosta löydettävästä käsitejärjestelmästä. Jos kaksi saman tietotason saavuttanutta henkilöä ei anna yhdelle käsitteelle juuri täysin samaa sisältöä, on silloin myös lähes mahdotonta, että käsiteanalyysin tuloksena syntyisi täysin identtisiä käsitejärjestelmiä käsitesuhteineen tekijästä riippumatta.

2.2 Käsitepiirteet

Käsitteet muodostuvat tarkoitteen ominaisuuksien kokonaisuudesta, joka tarkoitteesta on havaittavissa tai joita oletetaan tarkoitteessa olevan (Cabré, 1999, s. 42; Nuopponen, 2020, s. 97). Mielen tasolla tarkoitteen ominaisuuksia käsitellään tiedonosasina, joita kutsutaan käsitepiirteiksi, ja käsitepiirteiden kokonaisuudesta muodostetaan yhdessä käsite eli mielikuva tarkoitteesta (Cabré, 1999, s. 42; Suonuuti, 2024, s. 11–12). Käsitepiirteinä voidaan pitää mitä tahansa käsitteessä havaittavissa tai kuviteltavissa olevaa ominaisuutta (Sager, 1990, s. 24).

Käsitepiirteet jakautuvat Nuopposen (2020, s. 98–99) mukaan sisäisiin ja ulkoiisiin käsitepiirteisiin, joita analysoimalla käsitteet on mahdollista erottaa yläkäsitteestä, vieruskäsitteestä sekä sen alle sijoittuvista alakäsitteistä. Sisäiset käsitepiirteet ovat tarkoitteessa havaittavia ominaisuuksia. Tällaisia sisäisiä ominaisuuksia ovat esimerkiksi rakenne, koostumus, koko, määrä, materiaali, laatu, väri, muoto ja olemassaolon muoto. Ulkoiset käsitepiirteet ovat tarkoitteen ominaisuuksia, jotka muodostuvat vasta, kun tarkoitetta verrataan muihin tarkoitteisiin. Tällaisia ulkoisia ominaisuuksia ovat esimerkiksi alkuperä, sijainti, käyttötarkoitus, syy, ikä ja kesto. Nuopposen (2020, s. 98–99) määritelmästä poiketen sijoittaisin koon sisäisen käsitepiirteen lisäksi myös ulkoiseksi käsitepiirteeksi; vaikka tarkoitteesta itsestään olisi havaittavissa fyysiset mitat, kuten pituus ja paino, tarkoitteen voisi havaita esimerkiksi kooltaan isoksi vasta sitten, kun tarkoitteen kokoa vertaillaan muiden tarkoitteiden rinnalla.

Käsitteet edustavat abstraktia yleistystä tarkoitteiden jakamista käsitepiirteistä, joiden perusteella käsitteet on mahdollista kategorisoida omiin luokkiinsa (Cabr , 1999, s. 43; L'Homme, 2020, s. 23–24; Sager, 1990, s. 25). Käsitteen kuuluminen samaan kategoriaan on riippuvainen siitä, jakaako käsite muiden kategoriaan kuuluvien käsitteiden kanssa riittävästi yhteisiä käsitepiirteitä, ja vastavuoroisesti vaadittavien käsitepiirteiden puuttuminen tarkoittaa, ettei käsite voi kuulua samaan kategoriaan (L'Homme, 2020, s. 23–24). Useimpien käsitteiden rajat ovat likimääräisiä ja hajanaisia, mutta käsitteiden järjestelmällisellä tarkastelulla käsitteiden rajat on mahdollista keinotekoisesti määrittellä (Cabr , 1999, s. 43).

Käsitteen kategorisoinnin kannalta kaikki käsitepiirteet eivät ole oleellisia (Sanastokeskus, 2020, s. 14). Suonuutin (2024, s. 12) mukaan lähes jokainen käsite sisältää tyypillisesti runsaasti yleisiä käsitepiirteitä, jotka ovat yleisyytensä takia epäkelvoja tietyn käsitteen erottamiseen toisista käsitteistä. Tästä syystä käsitteiden kategorisoinnissa tulee myös kiinnittää huomiota yhteisten käsitepiirteiden lisäksi erottaviin käsitepiirteisiin. Erottavat käsitepiirteet ovat niitä tarkoitteen ominaisuuksia, jotka määrittävät käsitteen joko yksin tai yhdessä muiden käsitepiirteiden kanssa ja erottavat nimensä mukaisesti tietyn käsitteen muista käsitteistä (Suonuuti, 2024, s. 12).

On huomioitava, että se, mitä käsitepiirteitä otetaan tarkastelun kohteeksi ja mistä näkökulmasta käsitettä tarkastellaan, vaikuttaa käsitteen kategorisointiin (L'Homme, 2020, s. 88–89). Esimerkiksi generatiivista tekoälyä voidaan tarkastella tekoälyn alalajina, mutta sen käsiteala muuttuu, kun generatiivista tekoälyä lähdetään tarkastelemaan esimerkiksi laajemmin tietokoneteknologian näkökulmasta tai työelämän työvälineiden näkökulmasta. Tällöin käsitteen rajat ja asema käsitejärjestelmässä muuttuvat, ja käsite voi saada uuden ryhmittelyn myötä uuden yläkäsitteen lisäksi mahdollisesti myös uusia vierus- ja alakäsitteitä.

2.3 Käsitesuhteet ja käsitejärjestelmät

Käsitteet eivät terminologisessa tutkimuksessa ole muusta todellisuudesta irrallisia tiedon yksiköitä, vaan käsitteet ovat aina olemassa suhteessa toisiin käsitteisiin (Nuopponen, 2020, s. 100) muodostaen käsitejärjestelmiä (Suonuuti, 2024, s. 13). Käsitesuhteet ovat ajattelun tasolla muodostettuja käsitteiden välisiä kytköksiä, jotka yhdistävät käsitteet toisiinsa erilaisten loogisten yhteyksien tai löyhempien assosiaatioiden kautta (Nuopponen & Pilke, 2016, s. 33), ja jotka yhdessä luovat käsitteiden verkostoja, joista muodostuu käsitejärjestelmiksi kutsuttuja kokonaisuuksia tai osittaisia käsitejärjestelmiä (Nuopponen, 2020, s. 100). Käsitejärjestelmät ovat ihmisen ajatustoiminnan kautta luotuja abstrakteja konstruktioita käsitteiden yhteyksistä tai siitä, miten niiden ajatellaan olevan yhteydessä toisiinsa (Nuopponen & Pilke, 2016, s. 33).

Käsitejärjestelmä muodostuu käsitepiirteiden kautta, jotka luovat käsitesuhteita eri käsitteiden välille (Sanastokeskus, 2020, s. 14). Käsitesuhteita voidaan muodostaa tarpeen mukaan käytännössä mikä tahansa määrä, ja käsitteiden ala sekä käsitesuhteiden monimutkaisuus ovat käsiteanalyysissä riippuvaisia tutkijan käytännön tarpeista ja tavoitteista (Sager, 1990, s. 29). Erilaiset käsitesuhteet voidaan jakaa kahteen ryhmään kuuluvaksi, eli geneerisiin käsitesuhteisiin ja ontologisiin käsitesuhteisiin (Nuopponen, 2020, s. 101).

2.3.1 Geneeriset käsitesuhteet

Geneeriset käsitesuhteet ja käsitejärjestelmät muodostuvat tarkoitteiden jakamien yhteisten ominaisuuksien ja erojen perusteella, ja käsitteet jaetaan geneeristen käsitesuhteiden perusteella luokkiin, mitä voidaan pitää samana kuin tyypittelyä tai genrejaottelea (Nuopponen, 2020, s. 101; Nuopponen & Pilke, 2016, s. 33). Nuopponen ja Pilkeen (2016, s. 33) mukaan geneeriset käsitesuhteet ovat tavallisimpia ja samalla tärkeimpiä käsitesuhteita terminologisessa analyysissä ja niitä kutsutaan myös loogisiksi tai hierarkisiksi käsitesuhteiksi. Sanastokeskus TSK (2006, s. 16) määrittelee hierarkkisen eli

geneerisen käsitesuhteen käsitesuhteeksi, jossa käsitteen sisältö koostuu toisen käsitteen sisällöstä sekä yhdestä tai useammasta erottavasta käsitepiirteestä.

Geneerisissä käsitesuhteissa ja -järjestelmissä hierarkkisuus syntyy sillä, että geneeristen käsitesuhteiden avulla käsitteet järjestetään toisiinsa nähden ylempään, alempaan tai viereiseen järjestykseen, minkä perusteella käsitteitä kutsutaankin käsitejärjestelmässä yläkäsitteeksi, alakäsitteeksi tai vieruskäsitteeksi (Nuopponen & Pilke, 2016, s. 33). Käsitesuhteet järjestetään hierarkkisesti siten, että vähäsisältöisemmät mutta alaltaan laajemmat käsitteet ovat hierarkiassa ylempänä kuin enemmän sisältöä sisältävät mutta alaltaan kapeammat käsitteet (Nuopponen & Pilke, 2016, s. 34). Yläkäsitteen sisältö sisältyy alakäsitteeseen, jonka sisällössä on lisäksi vähintään yksi käsitettä rajaava käsitepiirre (Nuopponen, 2022, s. 10; Sanastokeskus TSK, 2006, s. 12–13). Samalla hierarkian tasolla olevia käsitteitä kutsutaan vieruskäsitteiksi. Eli vieruskäsitteet ovat alisteisia yläkäsitteelle, mutta eivät muille samalla tasolla oleville käsitteille.

Alakäsite siis perii sen yläkäsitteen käsitepiirteet ja samalla siinä on oltava vähintään yksi yläkäsitteestä ja vieruskäsitteistä erottava olennainen käsitepiirre (L'Homme, 2020, s. 181–182; Sanastokeskus TSK, 2006, s. 16; Sanastokeskus, 2020, s. 22). Käsitejärjestelmissä on monesti useita tasoja, ja yläkäsitteen ja suoraan sen alaisuudessa olevan alakäsitteen suhde pätee myös sen alla oleviin alakäsitteisiin eli yläkäsitteen ominaisuudet periytyvät alakäsitteiden alakäsitteisiin (L'Homme, 2020, s. 182). Tästä esimerkkinä L'Homme (2020, s. 182) kuvailee viulua kielisoittimen alakäsitteeksi, jolloin viulu perii kielisoittimen käsitepiirteet, kuten sen, että viulu tuottaa ääntä kielten värinän avulla. Kielisoittimet puolestaan kuuluvat musiikki-instrumenttien alaisuuteen, joiden käsitepiirteisiin kuuluu se, että sillä soitetaan musiikkia, mikä periytyy geneeristen käsitesuhteiden kautta myös viululle (L'Homme, 2020, s. 182).

2.3.2 Ontologiset käsitesuhteet

Ontologiset käsitesuhteet ja käsitejärjestelmät muodostuvat muista tarkoitteiden välisistä reaali maailmassa havaittavissa olevista yhteyksistä, kuten sijaintiin tai aikaan liittyvistä yhteyksistä (Nuopponen, 2020, s. 101; Nuopponen, 2022, s. 3). Terminologisessa kirjallisuudessa tai tutkimuksessa ei aina tehdä eroa erityyppisten ontologisten käsitesuhteiden välille ja niitä kutsutaan tyyppillisesti assosiatiivisiksi käsitesuhteiksi ilman tarkempaa kategorisointia (Nuopponen & Pilke, 2016, s. 37). Esimerkiksi sanastotyössä muut käsitesuhteet kuin hierarkkiset tai koostumussuhteet luetaan assosiatiivisiksi suhteiksi (ks. esim. Suonuuti, 2024, s. 17).

Tutkielmassani otan huomioon tarkemmin myös erityyppiset ontologiset käsitesuhteet, joita esimerkiksi Nuopponen (2020) esittelee kuuluvaksi ontologisiin käsitesuhteisiin, koska assosiatiiviset käsitesuhteet pitävät sisällään useita erilaisia käsitetiedon kannalta tärkeitä käsitesuhteita (L'Homme, 2020, s. 154–155). Tutkielmani aineiston luonteen näkökulmasta onkin tärkeää, että en rajaa analyysilleni oleellista tietoa pois keinotekoisesti keskittymällä pelkästään geneeriseen käsitesuhteeseen ja koostumussuhteeseen. Ontologisiin käsitesuhteisiin kuuluvat ainesosasuhde, alkuperäsuhde, arvosuhde, edustusuhde, kausaalisuuhde, kehityssuhde, koostumussuhde, lisäosasuhde, ominaisuussuhde, riippuvuus- ja korrelaatio suhde, sijaintisuuhde, temporaalinen käsitesuhde, toimintasuuhde ja transmissiosuhde (Nuopponen, 2020, s. 102; Nuopponen & Pilke, 2016, s. 37–42).

Analyysissäni huomioin kaikki yllä mainitut ontologiset käsitesuhteet, mutta kaikkiin käsitesuhteisiin ei löytynyt niihin sopivia käsitteitä aineistosta. Ontologisista käsitesuhteista analyysini kannalta olennaisimmiksi käsitesuhteiksi muodostuivatkin lopulta ainesosasuhde, koostumussuhde, ominaisuussuhde, riippuvuus- ja korrelaatio suhde, sijaintisuuhde ja toimintasuuhde, jotka esittelen seuraavaksi tarkemmin. Analyysini tulosten kannalta vähemmän oleellisia ontologisia käsitesuhteita edustavat siten alkuperäsuhde, arvosuhde, edustusuhde, kausaalisuuhde, kehityssuhde, lisäosasuhde, temporaalinen

käsitesuhde sekä transmissiosuhde, joita tässä tutkielmassa ei ole tarpeen käsitellä mainintaa syvemmin.

Ainesosasuhde on käsitesuhde, joka tuo tarkastelun kohteeksi käsitteestä ne asiat, ainesosat tai materiaalit, joista ilmiö on valmistettu tai joista se muodostuu (Nuopponen, 2020, s. 102). Abstraktien ilmiöiden kohdalla on huomioitava, että ainesosasuhteen ja koostumussuhteen välillä voi olla hankalaa tai mahdotonta tehdä eroa käsitesuhteissa (Nuopponen, 2011, s. 8).

Koostumussuhde on olemassa käsitteiden välillä, joiden tarkoitteet viittaavat samaan kokonaisuuteen tai ovat osia samasta kokonaisuudesta, ja käsitettä analysoidaan siitä näkökulmasta, miten käsitekokonaisuus voidaan jakaa erilaisiin osatekijöihin, osakokonaisuuksiin tai komponentteihin (L'Homme, 2020, s. 149; Nuopponen, 2020, s. 102; Nuopponen & Pilke, 2016, s. 38). Koostumussuhteen tarkastelun avulla käsitteiden muodostama kokonaisuus voidaan jakaa osatekijöihin, joka myös voidaan jakaa yhä pienempiin osiin (Nuopponen & Pilke, 2016, s. 38). Toisin kuin geneerisissä suhteissa koostumussuhteiset käsitteet eivät välttämättä jaa yhteisiä käsitepiirteitä, mutta koostumussuhteessa käsitteet ovat siitä huolimatta myös hierarkkisia toisiinsa nähden geneeristen käsitesuhteiden tavoin (L'Homme, 2020, s. 149).

Ominaisuussuhde keskittyy käsitteen ja sen ominaisuuksien välisiin suhteisiin ja se auttaa tarkastelemaan käsitteiden keskeisiä ominaisia piirteitä (Nuopponen, 2022, s. 14; Nuopponen, 2020, s. 102; Nuopponen, 2011, s. 8–9). Generatiivisen tekoälyn ominaisuuksiin voitaisiin laskea esimerkiksi tekstisisällön tuottamisen kohdalla tehokkuus tai nopeus.

Riippuvuus- ja korrelaationsuhde auttaa tarkastelemaan käsitettä siitä näkökulmasta, mistä ilmiön olemassaolo on riippuvainen ja mitkä ilmiöt vaikuttavat siihen (Nuopponen, 2020, s. 102). Riippuvuus- ja korrelaationsuhde on siten vastavuoroinen kahden käsitteen välillä eli jos yksi käsite muuttuu, toisessa käsitteessä tapahtuu myös muutos

(Nuopponen, 2022, s. 21). Tekoälyn kontekstissa esimerkiksi generatiivisen tekoälyn kohdalla voitaisiin muodostaa riippuvuussuhde dataan, jolla tekoälymalleja ohjelmoidaan – ilman ison datamäärän saatavuutta generatiivisen tekoälymallin kehittäminen edistyneempään versioon voisi olla haastavaa tai jopa mahdotonta.

Sijaintisuhteella tuodaan esiin paikka, jossa ilmiö sijaitsee tai jossa sen voidaan havaita tai koeta esiintyvän (Nuopponen, 2020, s. 102). Lisäksi sijaintisuhteella on mahdollista tarkastella myös ilmiön sisäisten ilmiöiden esiintymissijainteja. Tekoälyn kohdalla ilmiö voidaan havaita monessa eri kontekstissa. Tekoäly voi sijaita esimerkiksi tekoälynkehittäjien suunnittelupöydän lisäksi työpaikoilla työtä tehostavana työkaluna tai uutisissa kirjoituksen pääasiallisena aiheena.

Toimintasuhte kuvaava toimintaa sekä toimintaan liittyviä erilaisia näkökulmia (Nuopponen, 2020, s. 102; Nuopponen & Pilke, 2016, s. 42), ja toimintasuhteella voidaan tarkastella sitä, minkälaisia ilmiöitä voidaan osoittaa toimijan ja toiminnan välillä tai toiminnan ja objektin välillä (Nuopponen & Pilke, 2016, s. 42). Toimintasuhdetta selvittäessä vastataan sellaisiin kysymyksiin, joissa selvitetään muun muassa sitä, kuka toimintaa suorittaa, minkälaisia motivaatioita ja tavoitteita toiminnan suorittamisen taustalla on tai missä ja koska toimintaa toteutetaan (Nuopponen, 2020, s. 102; Nuopponen & Pilke, 2016, s. 42). Toimintasuhteilla on mahdollista tuoda yhteen erillisiä käsitejärjestelmiä tai osasia erilaisista käsitejärjestelmistä (Nuopponen, 2011, s. 10).

Nuopponen ja Pilkkeen mukaan (2016, s. 37–38) osa ontologisista käsitesuhteista, kuten ainesosassuhde ja sijaintisuhte, sopii paremmin konkreettisiin ilmiöihin ja toiset käsitesuhteet, kuten kausaalisuhde ja temporaalinen käsitesuhde, sopivat luontevammin abstraktien ilmiöiden välisten suhteiden kuvailemiseen. Jotkin käsitesuhteet, kuten koostumussuhde, sopivat luontevasti niin konkreettisten kuin abstraktienkin ilmiöiden käsitesuhteiden kuvailuun.

Käsitesuhdetyypit antavat tilaa joustavalle tulkinnalle, mikä heijastuu niiden nimeämisessä, jotta niitä on mahdollista käyttää luontevasti erilaisten erikoisalojen yhteydessä (Nuopponen, 2020, s. 102–103). Nuopponen (2020, s. 102) antaa tästä tulkinnan joustavuudesta esimerkkinä lisäosa-käsitesuhteen, jossa on luontevaa puhua erilaisten koneiden lisäosista kuten tietokoneen hiirestä kuin myös abstraktimpienkin lisäosasuhteiden välillä kuten sähköpostin ja liitteen kohdalla. Nuopponen (2020, s. 102–103) kannustaa-kin muissa kuin terminologisissa tutkimuksissa käyttämään suhdenimityksiä suhdetyyppien hahmottamisessa ja itse analyysissä käyttämään jotain muuta luontevampaa nimitystä tarvittaessa. Koska tutkielmani on luonteeltaan terminologinen tutkimus, pitäydyn analyysissäni yllä esitellyissä suhdetyyppien nimityksissä.

2.4 Termi

Tässä tutkielmassa tarkastelen käsitteitä tekstipohjaisesta aineistosta kielen kautta. Tämä tarkoittaakin sitä, että vaikka käsitteet muodostetaan ajattelulla mielikuviksi, käsitteet eivät ole pelkästään ajatuksen tasolle jääviä mielikuvia ilmiöistä, sillä niistä on mahdollista viestiä muun muassa termien ja määritelmien avulla.

Terminologisessa tutkimuksessa termi voidaan määritellä kielellisenä nimityksenä erikoisalan yleiskäsitteelle (ks. esim. Cabré, 1999, s. 35; Pasanen, 2015, s. 179; Sanastokeskus TSK, 2006, s. 22). Jotta termillä voidaan kuvata erikoisalan yleiskäsitettä, käsitteen on oltava olemassa tai synnyttävä ennen termin muodostamista (Cabré, 1999, s. 42; Sager, 1990, s. 22). Toisin sanoen termi siis viittaa olemassa olevaan erikoisalan tietosisältöön eli erikoisalan käsitteeseen, joka on vuorostaan muodostettu ajatuksen tasolla reaali maailman ilmiöstä eli tarkoitteesta. Termit syntyvät erikoisaloilla esimerkiksi silloin, kun erikoisalan asiantuntijat luovat uuden käsitteen uuden löydöksen yhteydessä, kun vanhan käsitteen tietosisältöä järjestetään uudelleen, kun suunniteltujen teollisten kehityksien tuloksena syntyy uusia käsitteitä, jotka tarvitsevat uusia nimityksiä, ja uusia termejä voi syntyä satunnaisten havaintojenkin kohdalla (ks. Nykänen, 1997; Sager, 1990, s. 61–62).

Termien yhteydessä ilmenee usein myös määritelmiä. Määritelmällä tarkoitetaan käsitteen kielellistä kuvausta, jossa kerrotaan käsitteen sisällöstä tietoa niin, että käsite erottuu lähikäsitteistä (ks. esim. Sanastokeskus TSK, 2006, s. 19). Hyvän määritelmän piirteinä voidaan pitää muun muassa sitä, että määritelmä paljastaa käsitteen sisällön, määritelmä tuo esiin käsitepiirteitä, jotka auttavat erottamaan kuvailtavan käsitteen ylä- ja alakäsitteistä ja että määritelmä yhdistää käsitteen termiin jollain tavalla (Nuopponen, 2020, s. 99).

2.4.1 Termin määritelmän häilyvyys

Tutkielmassani etsin ja analysoin tekoälykäsitteitä ja -termejä sellaisesta aineistosta, jonka tekstit eivät ole tekoälyn asiantuntijoiden kirjoittamia. Aineiston yleistajuisuus tuo esiin ongelmallisuuden tutkielmani termin määrittelylle ylempänä antamani termin määritelmän näkökulmasta. On mahdollista, että satelliittimalliin ja sitä kautta termianalyysiin päätyy nimityksiä, jotka eivät perinteisessä terminologisessa mielessä välttämättä täytä termin määritelmää, koska ne eivät ole esimerkiksi erikoisalan asiantuntijoiden luomia ja/tai alalla käytössä olevia tiukasti erikoisalan käsitetietoon viittaavia nimityksiä. Siksi tarjoan seuraavassa 2.4.2-alaluvussa oman vaihtoehtoisen termin määritelmän, jota käytän tämän tutkielman kontekstissa.

Tämän tutkielman kontekstissa on siis otettava huomioon se, että termin määrittely ei ole aivan yksiselitteistä, sillä termin määritelmälle ei ole todellisuudessa yksimielisyyttä terminologian alalla (L'Homme, 2020, s. 55). Termille on pyritty antamaan sen erikoisala-kytköksen ja yleiskäsitteeseen viittaamisen lisäksi erilaisia rajauksia. Tällaisia rajauksia ovat esimerkiksi kontekstisidonnaisuuden puuttuminen, vaatimus tarkasta määritelmästä, vaatimus termin vakiintuneesta statuksesta ja aktiivisesta käytöstä erikoisalalla tai rajaus, jonka mukaan erikoisalojen ulkopuolella ei voi olla termejä (ks. esim. Pasanen, 2009, s. 27–29; Pasanen, 2015, s. 180; Satokangas, 2021, s. 69–72). Tutkijoiden mukaan ne eivät kuitenkaan ole osoittautuneet ongelmattomiksi rajanvedoiksi. Termit

esimerkiksi siirtyvät erikoiskielestä yleiskieleen ja yleiskielestä erikoiskieleen, termejä määritellään kontekstin mukaisesti kirjavin käytännöin ja usein erikoisalan käsitteiden abstraktin luonteen takia niille on vaikea tehdä tarkkoja rajanvetoja, mistä syystä erikoisalan käsitteiltä puuttuu monesti myös tarkat määritelmät (Pasanen, 2009, s. 27–28; Satokangas, 2021, s. 80).

Pasanen (2009, s. 27) mukaan terminologisessa tutkimuksessa ei ole kyetty antamaan termille tarkkaa määritelmää, jonka perusteella termi olisi mahdollista tunnistaa varmasti mistä tahansa sanasta tai sanaliitosta. L’Homme (2020, s. 59) näkemyksen mukaan termiä itsessään ei ole edes olemassa, sillä kielellisestä yksiköstä muodostuu termi vasta silloin, kun se suhteutetaan aihealueeseen sekä sen rajaukseen, mistä käsin termiä tarkastellaan. Tämän seurauksena myös tavalliset sanat voivat muodostua termeiksi erikoisaloilla, ja sama sana voi olla terminä useilla eri erikoisaloilla, jolloin termi on myös altis eri erikoisalojen näkökulmien vaikutuksille (L’Homme, 2020, s. 59).

L’Homme (2020, s. 66) tuokin esiin sen näkökulman, että termi on suhteellinen käsite, sillä termin kohdalla tulee ottaa huomioon erilaisia tekijöitä, joista tärkeimpiä ovat kielen ulkoiset tekijät eli se erikoisala ja erikoisalan rajaus, johon termi liittyy sekä käyttötarkoitus, jonka takia termejä kerätään. Hänen mukaansa tämä johtaa siihen, että asiantuntijat tuottavat erilaisia termilistoja tilanteessa, jossa asiantuntijoita pyydetään poimimaan termejä samasta tekstistä (ks. myös Pasanen, 2009, s. 45). Sen lisäksi erilaiset termiin liittyvät teoriat ja määritelmät myös vaikuttavat siihen, mitä termeinä ylipäätään pidetään (L’Homme, 2020, s. 66). Myös Pasanen (2009, s. 45) toteaa, että termien tunnistaminen on aina subjektiivista.

Termin häilyvyydestä kirjoittaa myös Satokangas (2021, s. 69–72). Satokangas (2021, s. 71) päätyy termin määritelmän tarkastelussa tulkintaan, jonka mukaan termi on rajaukseltaan sumea kimppu piirteitä, joista erilaiset asiat korostuvat näkökulman vaihtuessa. Satokankaan mukaan terminologisen tutkimuksen kontekstissa on kuitenkin selvää se,

että käsiteanalyysissä termi on tunnus käsitettä edustavalle ajatuskokonaisuudelle. Myös tässä tutkielmassa käyttämäni termin määrittely alkaa tästä lähtökohdasta.

Satokangas (2021, s. 71) hyödyntää termin määritelmän häilyvyyttä ja tarjoaa omaa tutkimustaan tukevan määritelmänsä termille. Satokangas (2021, s. 71–72) määrittelee termin sen kirjoittajan toiminnan, tarkoitusperän sekä kontekstin näkökulmista ja pitää termiä kielellisenä ilmauksena, jonka sen kirjoittaja esittää erikoistuneena tiedonalakohtaiseen käyttöön. Satokangas (2021, s. 71–72) hyväksyy tutkimuksessaan sanan termiksi tulkintakehyksen perusteella, eli käsitteen nimitys on termi, jos aineisto käsittelee erikoisalan ilmiötä, käsitteen nimitys kytkeytyy erikoisalan ilmiöön ja jos nimitys esitetään erikoistuneena erikoisalalla.

2.4.2 Termin määritelmä tässä tutkielmassa

Omassa tutkielmassani nojaan Satokankaan (2021) tavoin termin määritelmän häilyvyyteen. Tutkimukseni termianalyysin kohdalla on huomioitava tyyppillisestä poikkeava termin määrittely, ja tässä tutkielmassa laajennan termin käsitteen koskevan laajempaa alaa kuin mitä tyyppillisesti terminologisessa tutkimuksessa termillä tarkoitetaan. Kuten 2.4 Termini-luvun alussa esittämästäni määritelmästä käy ilmi, termillä viitataan perinteisesti erikoisalan käsitteen kielelliseen nimitykseen. Mutta esimerkiksi L’Hommen (2020) ja Satokankaan (2021) esiintuoma termin suhteellisuus ja Satokankaan (2021) poikkeava termin määrittely tarjoavat mahdollisuuden laajemmalle termin tulkinnalle tämän tutkielman kontekstissa.

Tässä tutkielmassa termillä tarkoitan erikoisalan yleiskäsitteen nimitystä sekä sellaisia yleiskäsitteiden nimityksiä, jotka ovat kytköksissä jollain tavalla tekoälyn erikoisalan tietoon, mutta joita ei perinteisesti terminologisessa tutkimuksessa laskettaisi termiksi. Määrittelen termin seuraavasti: termi on kielellinen nimitys erikoisalan yleiskäsitteelle sekä erikoisalaan liittyvälle tai liitettävälle yleiskäsitteelle. Tämä rajaukseltaan sallivampi

termin määritelmä antaa tutkimukseni näkökulmasta tarvittavaa liikkumavaraa termin analyysin toteuttamiseen.

Kytköksen tekoälyn erikoisalan ja käsitteen sekä sen nimityksen välillä todistan systemaattisen käsiteanalyysin tuloksena aineistosta muodostamallani käsitejärjestelmää kuvaavalla satelliittimallilla. Satelliittimallissa käsitteillä on yhteys niin tekoälyn pääkäsitteeseen kuin toisiin samassa käsitejärjestelmässä sijaitseviin käsitteisiin geneeristen sekä erilaisten ontologisten käsitesuhteiden kautta. Esimerkkinä termianalyysin kohteeksi valitsemastani aineistosta löytyvästä tekoälykäsitteestä on *tekoälyhype*, jota ei perinteisessä terminologisessa mielessä välttämättä laskettaisi termiksi vaan esimerkiksi uudissanaksi tai muotisanaksi. *Tekoälyhypen* sisällytän termianalyysin kohteeksi, koska siitä on vedettävissä yhteys tekoälyn erikoisalaan, ja koska kyseisen käsitteen nimitys vaatii termin laadun tarkempaa tarkastelua etenkin omakielisyyden osalta.

2.5 Systemaattinen käsiteanalyysi

Systemaattinen käsiteanalyysi on käsitteiden terminologinen analyysimenetelmä, jolla pyritään selvittämään tutkimuksen kohteena olevien käsitteiden sisältöjä ja käsitteiden välisiä suhteita hahmottaen näin kuvaa analyysin kohteena olevasta käsitejärjestelmästä sekä käsitteiden paikasta käsitejärjestelmän hierarkiassa (Nuopponen, 2003, s. 14; Nuopponen, 2020, s. 97). Systemaattisella käsiteanalyysillä voidaan lisäksi tuoda esille ja kuvata yhteys käsitteiden ja käsitteitä nimittävien termien välille (Nuopponen, 2020, s. 97). Käsiteanalyysin suorittamisen tuloksena syntyy kuvaus käsitejärjestelmästä tai -järjestelmistä (Nuopponen, 2009, s. 314). Käsiteanalyysissä käsitejärjestelmien visuaalisessa esittämisessä käytetään hyödyksi esimerkiksi satelliittimallia. Käsiteanalyysistä tekee systemaattisen eli järjestelmällisen se, kun analyysiprosessissa käsitetietoa järjestellään yhdenmukaiseksi loogiseksi kokonaisuudeksi, joka pyritään saamaan selville käsitteiden välisiä suhteita selvittämällä (Nuopponen, 2020, s. 100).

Käsiteanalyysillä luodaan käsitejärjestelmän ja käsitesuhteiden hahmottamisen ohella myös perustaa käsitelmärittelyille sekä termien laadunarvioinnille esimerkiksi omakieliyyden tai synonyymisuuden osalta (Nuopponen, 2003, s. 14). Vaikka käsitteiden määritelmät liittyvätkin luontevasti käsiteanalyysiin käsitteiden ja niiden termien kautta, eivät määritelmät ole sinänsä osa käsiteanalyysiprosessia vaan määritelmien laatimista voidaan pitää käsiteanalyysistä erillisenä työprosessina, jota hyödynnetään esimerkiksi sanastotyössä (Nuopponen, 2009, s. 314).

Systemaattisessa käsiteanalyysissä muodostan aineistossa esiintyvistä tekoälykäsitteistä kokonaiskäsitejärjestelmän. Käsitteiden sijoittelussa otan huomioon 2.3 Käsitesuhteet ja käsitejärjestelmät -alaluvuissa esittelemäni generisen käsitesuhteen sekä erilaiset ontologiset käsitesuhteet, joiden perusteella käsitteet sijoittuvat luomaani käsitejärjestelmään. Aineiston perusteella muodostamaani kokonaiskäsitejärjestelmää kuvaan graafisesti satelliittimallilla. Satelliittimallin esittelen 2.5.2 Satelliittimalli -alaluvussa.

2.5.1 Käsiteanalyysin tasot

Käsiteanalyysiä voidaan tehdä joko todellisuuden tason, käsitteiden tason tai ilmaisutason näkökulmista (Nuopponen, 2003, s. 18–19). Todellisuuden tasolla käsiteanalyysiä tehdään tarpeesta jäsentää todellisuuden tilaa, esimerkiksi uuden ilmiön tarkastelun kohdalla. Käsitteiden tasolla tutkija käyttää käsiteanalyysiä tutustuakseen alan todellisuuteen jo olemassa olevien käsitteiden kautta tai esimerkiksi tarpeesta selventää tai muokata valmiiksi olemassa olevia erikoisalan käsitteitä.

Ilmaisutasolla käsiteanalyysiä tehdään termien ja tekstien selvittämisen, tulkitsemisen tai esimerkiksi muokkaamisen motivoimana (Nuopponen, 2003, s. 18–19). Oma tutkimukseni sijoittuu käsitteiden tason ja ilmaisutason välimaastoon, sillä kiinnostukseni kohdistuu verkkoartikkelien tekoälykäsitesisällöistä muodostuvan käsitejärjestelmän analysointiin, käsitteitä nimittävien tekoälytermien laadun tarkastelemiseen sekä tekoälytermien muokkaamiseen vaihtoehtoisia tekoälytermejä luomalla.

2.5.2 Satelliittimalli

Käsitejärjestelmiä voi visualisoida esitettävään muotoon eri tavoin, esimerkiksi puu- ja kampakaaviomalleilla tai yksinkertaisilla taulukoilla, joissa tekstin avulla kuvataan käsitejärjestelmän eri osioita käsitesuhdetyyppineen (Nuopponen, 2020, s. 102–103). Tutkielmassani käytän käsitejärjestelmän visuaaliseen kuvaamiseen satelliittimallia. Satelliittimalli on käsitejärjestelmän visuaalinen esitysmuoto, eräänlainen ajatuskartta, jossa erityinen painopiste on käsitejärjestelmien ja käsitesuhteiden hierarkkisessa esittämisessä (Nuopponen, 2016, s. 190; Nuopponen, 2020, s. 103). Käsitesuhteiden esiintuomisen ja kuvaamisen ohella satelliittimallilla voidaan selventää myös käsitepiirteitä eli käsitteiden ominaisuuksia.

Satelliittimalli on rakenteeltaan yksinkertainen, sillä se koostuu noodeista, viivoista, teksteistä ja värikoodeista, joskin kaikki miellekarttaohjelmat eivät mahdollista värikoodien käyttöä (Nuopponen, 2020, s. 103). Satelliittimallin keskiössä on keskusnoodi, johon yhdistetään viivalla muut sen alaisuuteen sijoittuvat käsitteet, jotka on merkitty omiin satelliittinoodeihinsa. Käsitesuhteita voidaan merkitä käsitteitä toisiinsa yhdistäviin viivoihin joko tekstillä, värikoodeilla tai vaihtoehtoisesti käsitteiden väliin sijoitettavilla apunooodeilla, joihin käsitesuhteen suhdetyyppi tai muu jakoperuste merkitään tekstillä.

Satelliittimallin keskellä keskusnoodissa on pääkäsite, jonka ympärille muut käsitteet kietoutuvat, mutta keskusnoodiin kytketyt satelliittinoodit saavat alleen myös muita käsitteitä, jotka voivat saada alaisuutensa toisia käsitteitä erilaisten käsitesuhteiden perusteella, mikä voi toistua yhä uudelleen ja uudelleen (Nuopponen, 2016, s. 194). Tämä tarkoittaa käytännössä sitä, että käsitteiden sijainti käsitejärjestelmän hierarkiassa on suhteellinen – käsite, joka on pääkäsite yhdelle voi olla samanaikaisesti alakäsitteenä jollekin muulle käsitteelle (ks. esim. Suonuuti, 2024, s. 14), ja samalla kaikki käsitejärjestelmän käsitteet ovat alisteisia keskusnoodiin valitulle pääkäsitteelle.

Satelliittimallissa otetaan tarkastelun kohteeksi kerrallaan vain yksi pääkäsite, joka rajaa tarkastelun kohteena olevaa alaa (Nuopponen, 2011, s. 5; Nuopponen 2016, s. 193). Prosessin alussa pääkäsitteeksi valittu käsite voi edustaa tutkittavaa aluetta kokonaisuudessaan (Nuopponen, 2016, s. 193). Valittu käsite voi olla korkean tason abstrakti käsite tai muulla tavalla keskeinen käsite tutkinnan kohteena olevalla alalla ja käsitteen on kyettävä tuomaan käsitesuhteilla yhteen siihen liittyvät alakäsitteet (Nuopponen, 2011, s. 5; Nuopponen 2016, s. 193).

Luonteeltaan satelliittimalli on prosessina joustava ja dynaaminen; sitä voidaan laajentaa, muokata ja tarkentaa moneen kertaan analyysiprosessin aikana (Nuopponen, 2011, s. 5). Satelliittimallin luominen alkaa alustavasti miellekartan omaisen hahmotelman tekemisestä, jota lähdetään jalostamaan tiedon lisääntyessä tarkkarajaisempaan esitykseen useiden eri kehitysvaiheiden kautta (Nuopponen, 2020, s. 103). Toisin sanoen satelliittimallin luominen on iteratiivinen prosessi, jossa pyritään lähestymään jatkuvasti tarkempaa ja selkeämpää kuvausta käsitejärjestelmästä.

Muissa käsitekarttojen esittämistavoissa voidaan käyttää ristiin meneviä viivoja yhdistämään käsitteet toisiinsa (Nuopponen, 2016, s. 197–198), mutta satelliittimallissa ristiin merkitsemistä vältetään siten, että sama käsite saa ilmetä satelliittimallissa useamman kuin yhden kohdan tai suhdetyypin yhteydessä (Nuopponen, 2016, s. 198; Nuopponen, 2020, s. 105). Tällä esitystavalla pyritään ennen kaikkea ilmaisun selkeyteen, jotta satelliittimallia tehdessä voidaan välttyä lukemista ja tulkintaa hankaloittavilta noodien välillä tai ylitse ristiin meneviltä viivoilta. Selkeyden lisäksi käsitteiden merkitseminen useampaan kuin yhteen kohtaan auttaa tarkastelemaan tutkimuksen kohteena olevaa ilmiötä monipuolisesti eri näkökulmista (Nuopponen, 2020, s. 105).

Vaikka satelliittimalli esittää hyvin tehtynä kattavan ja selkeän kuvauksen kohteena olevasta käsitejärjestelmästä, satelliittimalli vaatii aina kirjallisen selityksen sen tueksi, jotta kuvion ymmärtävät muutkin kuin vain sen laatija (Nuopponen, 2020, s. 118–119). Tästä syystä erilaiset terminologisessa tutkimuksessa vakiintuneet merkintätavat kuin myös

tutkijan omat merkintätavat ja ratkaisut on aina selitettävä lukijalle erikseen myös tekstimuodossa.

Satelliittimallin luomiseen voi käyttää kynää ja paperia tai lukuisia ajatuskarttoja mahdollistavia karttaohjelmia. Tutkielmassani käytän satelliittimallin luomiseen vapaasti saatavilla olevaa verkkopohjaista Miro-ajatuskarttaohjelmaa. Mirolla on mahdollista luoda joustavia ja helposti muokattavia miellekarttoja isoista kokonaisuuksista. Miron muokausmahdollisuudet ovat myös riittävän joustavia esimerkiksi siihen, että yhtä aikaa on mahdollista tehdä pienemmästä kokonaisuudesta omaa osittaista käsitejärjestelmäkokonaisuutta, jonka voi myöhemmin liittää osaksi pääkäsitettä kokonaiskäsitejärjestelmässä. Ohjelma mahdollistaa värikoodauksen, mutta käytän selkeyden ja luettavuuden takia värikoodauksen sijaan apunoodeja käsitesuhteiden välissä käsitesuhteiden kuvaamiseen.

2.6 Termianalyysi

Tässä tutkielmassa termianalyysillä tarkoitan termien analyysimenetelmää, jolla määritellään käsitteiden ja termien välinen suhde sekä arvioidaan termin laatua sekä muodostetaan termejä (Nuopponen & Pilke, 2016, s. 18). Termianalyysillä tarkastelen siten analyysin kohteeksi valitsemieni tekoälytermien sopivuutta niiden käsitteisiin ja käsitteiden tietosisältöihin, arvioin termien laatua sekä muodostan vaihtoehtoisia termejä tilalle.

2.6.1 Hyvän termin kriteerit

Termianalyysissä termin laatua arvioin hyvän termin kriteerien näkökulmista, joilla tarkastellaan termin ominaisuuksia. Hyvän termin kriteerejä ovat omakielisyys, läpinäkyvyys, johdonmukaisuus, tarkoituksenmukaisuus, erottuvuus, lyhyys, produktiivisuus, kielellinen sopivuus ja kielellinen moitteettomuus (Isohella & Nuopponen, 2016, s. 229; Nuopponen & Pilke, 2016, s. 65–69; Nykänen, 1997).

Omakielinen termi koostuu kielen omista elementeistä (Nuopponen & Pilke, 2016, s. 69) ja Nykäsen (1997) mukaan omakielisyys hyvän termin kriteerinä tarkoittaa suomen kielen kontekstissa sitä, että suomenkielinen termi on parempi kuin muusta kielestä lainattu lainasana (ks. myös Nuopponen & Pilke, 2016, s. 68). Toisaalta kyky tuottaa uusia termejä pelkästään omakielisesti on rajallista, minkä takia uusien käsitteiden termien luomisen yhteydessä on joskus turvaututtava yleiskielen ja muiden erikoisalojen sanastojen lisäksi myös esimerkiksi vieraskielestä peräisin oleviin ilmaisiin (Nuopponen & Pilke, 2016, s. 71).

Läpinäkyvä termi heijastaa käsitteen keskeistä sisältöä ainakin yhden olennaisen käsitepiirteen kautta (Nuopponen & Pilke, 2016, s. 68). Läpinäkyvyys tarkoittaa, että termin itsessään tulee kertoa käsitteen olennaisimmat käsitepiirteet, ja termin osien tulisi viestiä oikea mielikuva käsitteestä (Isohella & Nuopponen, 2016, s. 230; Nykänen, 1997). Läpinäkyvää termiä voidaan tulkita myös niin, että henkilö ymmärtää termin riippumatta siitä, onko hän asiantuntija vai maallikko (Nuopponen & Pilke, 2016, s. 58). Tässä tutkielmassa läpinäkyvyys ajatellaan siten, että läpinäkyvä termi kuvailee jo itsessään käsitettä niin hyvin, että määritelmää tai taustatietoa ei välttämättä edes tarvita käsitteen ymmärtämiselle.

Johdonmukaisuudella Nykänen (1997) tarkoittaa termin suhdetta erityisesti käsitejärjestelmään. Johdonmukainen termi osoittaa käsitteen paikan käsitejärjestelmässä ja antaa tietoa ylä-, ala- ja vieruskäsitteistä (Nuopponen & Pilke, 2016, s. 67). Termin johdonmukaisuudella voidaan tarkoittaa Isohellan ja Nuopposen (2016, s. 232) tavoin myös termin johdonmukaista käyttöä tekstissä, mutta tässä tutkielmassa johdonmukaisuudella tarkoitetaan termin suhdetta sen käsitteeseen ja käsitejärjestelmään.

Termin tarkoituksenmukaisuudella Nykänen (1997) viittaa siihen, ettei termi saisi käytökontekstissa herättää haitallisia mielikuvia. Sen sijaan Isohellan ja Nuopposen (2016) määritelmässä negatiivisen konnotaation puuttuminen on vain yksi

tarkoituksenmukaisuuden ominaisuuksista. Tarkoituksenmukaisuudella Isohella ja Nuopponen (2016, s. 231) tarkoittavat termin soveltuvuutta tehtäväänsä, ja kuvaavat siihen liittyviä erilaisia ominaisuuksia. Negatiivisen konnotaation puuttumisen lisäksi tarkoituksenmukainen termi ottaa huomioon kohderyhmänsä ja käyttökontekstinsa, joiden perusteella voidaan myös arvioida kielen tasoa, rekisteriä, neutraaliutta sekä kielen rakennetta selkeyden ja muistettavuuden kannalta.

Termin erottuvuudella tarkoitetaan sitä, kuinka termi erottuu samalla erikoisalalla käytetyistä tai yleisesti tunnetuista termeistä kirjoituksen ja ääntämisen osalta (Isohella & Nuopponen, 2016, s. 229; Nykänen, 1997). Erottuva termi on siis sellainen, jota ei sekoiteta helposti toiseen termiin, ja jos termejä erottaa esimerkiksi vain yksi kirjain, on silloin riskinä, että termejä ymmärretään ja käytetään helposti väärin (Nuopponen & Pilke, 2016, s. 67).

Hyvä termi on myös lyhyt. Tarkkaa rajausta termin lyhyydestä ei kuitenkaan ole, ja mentäessä yhä alemmas käsitesuhteiden hierarkiassa, sitä pidempiä termit yleensä ovat, koska ne ilmaisevat paikkaansa käsitejärjestelmässä (Nuopponen & Pilke, 2016, s. 66). Sen lisäksi mitä monimutkaisempi käsite on, sitä pidempi sitä kuvaava termi usein on. Suomen kielessä myös yhdyssanojen käyttö kasvattaa termin pituutta, mikä voi johtaa lyhenteiden muodostamiseen. Isohella ja Nuopponen (2016, s. 229–230) huomauttavat, että terminmuodostuksessa termin taloudellinen ilmaisu onkin tasapainottelua lyhyen ei-läpinäkyvän termin ja pitkän läpinäkyvän termin välillä. Liian lyhyet termit ja lyhenteet hämmentävät termin sisältöä, kun taas liian pitkät termit voivat olla sisällöllisesti läpinäkyviä, mutta ne voivat hankaloittaa viestintää ja johtaa epävirallisten lyhenteiden käyttöön.

Termin produktiivisuus tarkoittaa, että termistä on vaivatonta muodostaa johtamalla uusia sanoja suomen kielen sääntöjen mukaisesti (Isohella & Nuopponen, 2016, s. 229–230; Nykänen, 1997). Kielellisellä sopivuudella tarkoitan sitä, että termi on helppoa sekä yksiselitteistä ääntää, kirjoittaa ja taivuttaa (Nykänen, 1997). Kielellisesti moitteeton termi

on kielelliseltä ulkoasultaan virheetön ja se noudattaa kieliopin sääntöjen lisäksi sanan- ja terminmuodostusmalleja (Isohella & Nuopponen, 2016, s. 229).

Termiä analysoitaessa tai uutta termiä muodostettaessa on otettava huomioon, että kaikkia hyvälle termille asetettuja vaatimuksia on usein vaikeaa täyttää (Isohella & Nuopponen, s. 229), eivätkä monet termit niitä täytäkään (Nykänen, 1997). Yhtä periaatetta noudattaessa joutuu usein tekemään kompromisseja toisen periaatteen kustannuksella (Nuopponen & Pilke, 2016, s. 65). On lisäksi huomioitava, että hyvän termin kriteereille ei ole olemassa myöskään selkeää tärkeysjärjestystä (Nykänen, 1997). Eli termianalyysin tulokset saattavat siten vaihdella sen mukaan, kuinka kukin termin analysoija itse arvottaa hyvän termin kriteerejä. Tämä kriteerien arvottamisen vaihtelu tekee käsitteen subjektiivisuuden ohella termianalyysistä osittain subjektiivisen analyysimenetelmän.

Tutkielmani kontekstissa erityistä painoarvoa asetan termin omakielellisyydelle sekä läpinäkyvyydelle, joita pidän tärkeimpinä hyvän termin kriteereinä. Näkemystäni mukaan molemmat hyvän termin ominaisuudet ovat erottamattomasti yhteydessä toisiinsa – termi ei voi olla läpinäkyvä, jos se ei ole omakielinen. Näkemystäni tukee osittain myös Nuopponen ja Pilke (2016, s. 68), joiden mukaan omakielinen termi on pääsääntöisesti aina läpinäkyvämpi kuin vieraskielinen termi, eli yhteys omakielisyyden ja läpinäkyvyyden välillä on havaittavissa.

Perustelen näkemystäni omakielisyyden ja läpinäkyvyyden painotuksesta sillä, että sellaiselle termin käyttäjälle, joka osaa vain suomea, vieraskielinen termi ei voi olla läpinäkyvä, koska läpinäkyvä termi kuvaa itsessään käsitteen tietosisällön ja tämän havaitsemiseksi sekä ymmärtämiseksi vaadittaisiin kyseisen vieraan kielen taitamista. Vieraskielisen termin kohdalla käsitteen sisällön päättelemisen termistä vaatii siten myös muun kuin äidinkielen taitoa, esimerkiksi riittävää englannin kielen taitoa, jota kaikilla termin käyttäjillä ei välttämättä ole. Yllä olevaa perustelua mukaillen myös termin johdonmukaisuus jää vierasta kieltä osaamattoman käyttäjän pimentoon, koska

vieraskielisestä termistä on silloin lähes mahdotonta hahmottaa käsitteen sijaintia ja suhdetta käsitejärjestelmässä.

2.6.2 Termin muodostus

Kun olen analysoinut valitsemiani termejä hyvän termin kriteerien näkökulmista, siirryn lopuksi muodostamaan niille vaihtoehtoisia termejä, jotka täyttävät hyvän termin kriteerit paremmin erityisesti omakielisyyden ja läpinäkyvyyden osa-alueilta. Uusien vaihtoehtoisten termien luomisprosessissa otan huomioon termin muodostuksen periaatteita. Termin muodostuksessa on huomioitava, että kielellisesti termi voi olla yksisanainen tai moniosainen eli termit voivat koostua yhden sanan lisäksi myös yhdyssanasta tai sanaliitosta (Nuopponen & Pilke, 2016, s. 61–62; Nykänen, 1997; Suonuuti, 2024, s. 31), ja termissä voi kirjainten lisäksi olla myös numeroita, merkkejä ja symboleja (Nuopponen & Pilke, 2016, s. 61–62).

Yksisanainen termi voi olla yhdistämätön termi eli sana, jota ei voi jakaa pienempiin osiin (esim. *äly*), johdostermi, joka on muodostettu lisäämällä kantasanaan johdin (esim. *älykyys*), yhdyssanatermi (esim. *tekoäly*) tai lyhenne (esim. *AI* eli *artificial intelligence*) (Nuopponen & Pilke, 2016, s. 62; Sanastokeskus TSK, 2006, s. 23). Yhdyssanat rakentuvat perusosasta ja määriteosasta (Hongisto, 2023, s. 2). Esimerkiksi *tekoälyn* kohdalla *äly* on yhdyssanan perusosa ja *teko* on määriteosa, joka kuvailee perusosaa muun muassa siten, että kyse on keinotekoisesti valmistetusta älystä tai älyn jäljitelmästä.

Sanaliitot ovat monisanaisia termejä ja ne koostuvat kahdesta tai useammasta sanasta, jotka on kirjoitettu erilleen (Sanastokeskus TSK, 2006, s. 23), kuten *generatiivinen tekoäly* -sanaliittotermissä. Sanaliitot voivat sisältää esimerkiksi substantiiveja, adjektiiveja, verbejä, adverbeja ja partikkeleita (Nuopponen & Pilke, 2016, s. 62; Suonuuti, 2024, s. 31). Esimerkiksi *generatiivinen tekoäly* -sanaliitto koostuu *generatiivinen*-adjektiivista ja *tekoäly*-substantiivista, joka on samalla myös yhdyssana, eli yhdyssanat voivat siten ilmetä yhdyssanatermien ohella myös sanaliittotermeissä.

3 Tekoäly

Vaikka tekoäly on nyt esillä median puheenaiheena etenkin generatiivisen tekoälyn muodossa, tekoälyteknologioita on ollut käytössä jo pitkään. Tekoäly ja sen sovellukset ovat digitalisaation myötä arkipäiväistyneet niin, että emme välttämättä tule edes huomaneeksi sitä. Esimerkiksi ihan tavallisissa, lähes kaikkien käytössä olevissa mobiililaitteissa kuten älypuhelimissa on nykypäivänä monenlaisia tekoälyyn perustuvia toiminnallisuuksia. Tekoäly on läsnä myös monenlaiseen arkiseen asiointiin liittyvien toimintojen sekä palvelujen yhteydessä, kuten ostamisessa, pankkipalveluissa, navigoinnissa ja mediassa.

Tässä pääluvussa taustoitan tekoälyä ilmiönä esittelemällä tekoälyn määritelmiä, tekoälyyn liittyviä keskeisimpiä käsitteitä ja tekoälystä tehtyä aikaisempaa tutkimusta sekä journalismin että terminologisen tutkimuksen konteksteista. Aloitan kertomalla tekoälyn erilaisista määritelmistä. Määritelmien esittelyn jälkeen tarjoan oman määritelmäni tekoälystä. Sen jälkeen esittelen tekoälykäsitteitä, kuten generatiivisen tekoälyn ja kiellimallin. Lopuksi siirryn taustoittamaan tekoälystä tehtyä aikaisempaa tutkimusta journalismin ja terminologian näkökulmista.

3.1 Tekoälyn määritelmä

Tekoälyn käsitteelle on olemassa monia erilaisia määritelmiä. Määritelmät asettavat tekoälyn käsitteen erilaisiin käsitejärjestelmiin, jotka rajaavat käsitettä eri tavoilla. Esittelen tässä alaluvussa tekoälylle annettuja erilaisia määritelmiä ensin instituutiokontekstissa, jonka jälkeen keskityn tekoälyn määritelmiin tekoälytutkimuksen kontekstissa. Lopuksi annan esiteltyihin määritelmiin pohjautuvan oman määritelmäni tekoälylle.

Euroopan parlamentin (2023) määritelmässä tekoäly on *koneen kyky* käyttää sellaisia taitoja, jotka voidaan katsoa kuuluvan ominaisesti ihmisen älykkyyteen. Tällaisina taitoina määritelmässä pidetään esimerkiksi päättelykykyä, kykyä oppia, suunnitella tai luoda. Määritelmässä tekoälylle annetaan kaksi eri olomuotoa, jotka ovat erilaiset ohjelmistot

ja ruumiillistettu tekoäly, kuten robotit ja itseohjautuvat autot. Euroopan unionin (2024) tekoälysäädös määrittelee tekoälyn *konepohjaiseksi järjestelmäksi*, joka toimii vaihtelevalla autonomian tasolla päätellen sille annetuista syötteistä tavoitteiden mukaisia tuotoksia kuten sisältöä, ennusteita, suosituksia tai päätöksiä. Eli EU:n sisällä tekoäly määritellään ainakin kahdella eri tavalla: yhdessä määritelmässä tekoäly käsitteenä on koneen kyky ja toisessa määritelmässä tekoäly on konepohjainen järjestelmä.

Myös Suomen valtiovarainministeriö määrittelee tekoälyn *tietojärjestelmän kyvyksi* suorittaa tyypillisesti ihmismäistä älykkyyttä vaativia tehtäviä, kuten oppimista, ongelmanratkaisua ja kielen ymmärtämistä (Tammes-Peters, 2025, s. 8). Tekoälyn määrittelemineen koneen kyvyksi on kuitenkin ongelmallista. Se voi antaa vääristyneen kuvan käsitteestä, että tekoäly voisi olla minkä tahansa koneen yksi ominaisuuksista, jonka kone voi mahdollisesti opetella tai sille voidaan antaa. Tämä ei luonnollisesti ole kaikille koneille mahdollista, sillä tekoäly vaatii toimiakseen tietokoneille ominaisia komponentteja, kuten laskentaan vaadittavia prosessoreita. Tietojärjestelmän kyvyksi määrittelemineen sen sijaan asettaa tekoälyn käsitteen tietotekniikan teknologian yläkäsitteen alaisuuteen.

Julkista rahoitusta saava ja Sanastokeskuksen ylläpitämä erikoisalojen termejä sekä niiden määritelmiä sisältävä TEPA-Termipankki antaa tekoälylle kaksi erilaista määritelmää. Yhdessä määritelmässä tekoäly on *ohjelma*, joka jäljittelee ihmisen älykkyyttä vaativia toimintoja (Sanastokeskus, n.d.b). Toisessa määritelmässä tekoäly on *teknologia*, jonka avulla koneet ja ohjelmat voivat jäljitellä ihmisenkaltaisen älykkyyden mukaisia toimintoja, kuten kielen tuottamista. Molemmat määritelmät asettavat tekoälyn tietotekniikan teknologian pääkäsitteen alaisuuteen eri painotuksin.

Tekoäly määritellään myös tekoälyyn ja sen käyttöön keskittyvissä tutkimuksissa eri tavoilla. Esimerkiksi Huang ja Rust (2021, s. 31) määrittelevät tekoälyn *koneiksi*, jotka osoittavat ihmisälykkyydelle ominaisia piirteitä. Määrittelyssään he erottavat tekoälyn yleisestä informaatioteknologiasta, sillä tekoälyt pitävät sisällään oppimiseen, yhdistämiseen ja soveltamiseen kykeneviä teknologioita. Vaikka Huang ja Rust (2021) asettavat

tekoälyn loogisesti informaatioteknologian piiriin, ongelmallista tässä määritelmässä on se, että tekoäly on asetettu koneen käsitteeksi; osoittavatko koneet ihmisälykkyydelle ominaisia piirteitä vai onko ihmisälykkyydelle ominaisten piirteiden taustalla edistyneitä tietoteknologisia ohjelmia ja matemaattisia laskukaavoja, jotka saavat koneen toimimaan kuvatusen kaltaisesti?

Haleem ja muut (2022, s. 120) määrittelevät tekoälyn *tietojenkäsittelytieteen teknologiaaksi*, jonka avulla tietokone voidaan opettaa ymmärtämään ihmisille ominaista kommunikaatiota ja käyttäytymistä sekä jäljittelemään näitä kykyjä. Lisäksi he tuovat esille, että tekoäly on kokoelma useita erilaisia teknologioita, jotka kykenevät toteuttamaan ihmisenkaltaista älykkyyttä vaativia tehtäviä. Esimerkiksi kommunikointiin keskittyvät tekoälyagentit ovat Guzmanin ja Lewisin mukaan (2020, s. 72) seurausta etenkin luonnollisen kielen käsittelyn sekä luonnollisen kielen tuottamisen teknologioiden kehittymisestä. Tekoälyagentit ovat autonomisia järjestelmiä, jotka suorittavat kapeita tehtäviä rajatuissa digitaaliympäristöissä (ks. esim. Ren ja muut, 2025, s. 126). Guzman ja Lewis (2020, s. 72) asettavat Haleemin ja muiden (2022) tavoin tekoälyn teknologian alakäsitteeksi – tekoäly on ihmisen kehittämä *teknologia*, joka on suunniteltu suorittamaan ihmisen älykkyyteen liitettyjä tehtäviä.

Myös Xing ja muut (2026) tuovat määritelmässään esille, että tekoäly käsitteenä koostuu useista eri teknologisista tekijöistä. Xingin ja muiden (2026, s. 2) mukaan tekoäly viittaa *koneisiin tai järjestelmiin*, jotka pystyvät suorittamaan ihmisen älykkyyttä vaativia kognitiivisia funktioita, kuten vuorovaikutusta, oppimista ja ongelmanratkaisua. Heidän mukaansa tekoäly koostuu useista eri teknologioista, kuten koneoppimisesta, älykkäistä roboteista, luonnollisen kielen käsittelystä ja konenäöstä. Yllä mainittujen ominaisuuksien lisäksi Xing ja muut esittävät, että tekoälyteknologian luonteeseen kuuluvat monimutkaisuus, epävarmuus ja sosiaalisuus.

Banh ja Strobel (2023) tarjoavat yllä esitellyistä määritelmistä hieman poikkeavan määritelmän. Banhin ja Strobelin (2023, s. 62) mukaan tekoäly on *kattotermi*, joka kattaa

alleen erilaisia laskennallisia algoritmeja, jotka kykenevät suorittamaan tyypillisesti ihmisälyä vaativia tehtäviä. Tällaisiksi tehtäviksi he laskevat esimerkiksi luonnollisen kielen ymmärtämisen, päätöksenteon ja kokemuksesta oppimisen. Banhin ja Strobelin (2023) määritelmästä kuitenkin puuttuu tekoälykäsitteen asettaminen osaksi tarkempaa pääkäsitettä. Tekoälyä ei määritelmässä aseteta suoraan algoritmiksi vaan tekoälyn kerrotaan ainoastaan sisältävän niitä. Samalla tekoälyn määrittelemisen kattotermiksi ei myöskään kerro itse käsitteen merkityssisällöstä sitä, mihin pääkäsitteeseen tekoäly kuuluu. Määritelmä on kuitenkin onnistunut siltä osin, että se kertoo lukijalla tekoälyn koostuvan useista eri algoritmeista eli laskennallisista menetelmistä. Määritelmästä ja sen kontekstista on silti tulkittavissa, että tekoäly asettuu tässä määritelmässä tietokoneisiin liittyvän teknologian alakäsitteeksi.

Banh ja Strobel (2023, s. 62) ovat luoneet tekoälykäsitteestä yksinkertaisen hierarkkisen käsitejärjestelmän, jonka graafisessa esittelyssä tutkijat asettavat tekoälyn pääkäsitteen alle koneoppimisen käsitteen, jonka alakäsitteeksi he asettavat syväoppimisen. Syväoppimisen alakäsitteeksi he asettavat generatiivisen tekoälyn. Koneoppiminen on Banhin ja Strobelin (2023, s. 63) mukaan tekoälyn osa-alue, joka keskittyy kehittämään algoritmeja, jotka kykenevät datan avulla ratkaisemaan itsenäisesti tehtäviä ilman eksplisiittistä ohjelmointia, mitä he kutsuvat oppimiseksi.

Koneoppimiseen käytetään useita erilaisia oppimisprosesseja, jotka riippuvat käytetyn datan luonteesta sekä toivotuista lopputuloksista (Banh & Strobel, 2023, s. 63). Ohjattu oppiminen on yksi oppimisprosessi, jossa algoritmeja koulutetaan nimettyjen datajoukkojen avulla luokittelemaan tai ennustamaan dataa, jolloin algoritmi oppii yhdistämään syötteet tulosteisiin luoden uusia ennusteita. Muita esimerkkejä koneoppimisessa käytetyissä oppimisprosesseista ovat ohjaamaton oppiminen ja vahvistusoppiminen. Ohjaamattomalla oppimisella pyritään löytämään merkitsemättömästä datajoukosta piilossa olevia rakenteita ja kuvioita. Vahvistusoppimisessa pyritään optimaaliseen päätöksentekoon olemalla vuorovaikutuksessa ympäristön kanssa sekä kumulatiivisen hyödyn maksimointiin pitkällä aikavälillä yrityksen ja erehdyksen kautta. Yhteistä kaikilla

koneoppimisalgoritmeilla on se, että ne pyrkivät käsittelemään dataa erottelevasti esimerkiksi luokitteluun tai päätöksentekorajojen määrittämiseksi.

Syväoppiminen on koneoppimisen edistyneempi osa-alue, joka hyödyntää keinotekoisia neuroverkkoja monimutkaisten datarepresentaatioiden mallintamiseen sekä automaattiseen korrelaatioiden ja kuvioiden havaitsemiseen suurista datajoukoista (Banh & Strobel, 2023, s. 63). Neuroverkoilla tarkoitetaan laskennallisia malleja, joiden rakenne on saanut mallinsa ihmisaivoista, ja ne koostuvat toisiinsa kytketyistä keinotekoisien neurotien kerroksista. Banh ja Strobel (2023) osoittavatkin luomallaan käsitejärjestelmällään, että tekoäly käsitteenä on iso kokonaisuus, jonka alaisuuteen kuuluvat useat erilaiset tekniikat, joilla kone pyritään saamaan tuottamaan erilaisia ihmisen toimintaa muistuttavia toimintoja.

Yksi merkittävä tekijä, jolla tekoälyn erilaiset sovellukset eroavat käsitteenä perinteisistä automatisoiduista tietokoneohjelmista on tekoälyn mahdollisuus toimia autonomisesti eli itsenäisesti. Frank ja muut (2023, s. 156) määrittelevät tekoälyn autonomian tekoälyteknologian kyvyksi suorittaa ihmisiltä johdettuja tehtäviä ilman erillistä ihmisen puuttumista kyseiseen toimintaan. Tekoälyn autonomia mahdollistaa tekoälyn sopeutumisen ympäristöönsä ennakoivasti toimien aikaisemmin oppimansa sekä ympäristön havainnoinnin pohjalta.

Tekoäly voi toimia autonomian suhteen eri tasoisesti. Hauptmanin ja muiden (2024, s. 438) mukaan tekoälyn autonomian tutkijat jakavat tekoälyagentin autonomian tason kolmeen tasoon: ei autonomiaa, osittainen autonomia ja korkea autonomia. Hauptman ja muut (2024, s. 438) korostavat, että tekoäly, joka vaatii ihmisten panosta päätösten tai toimintojen suorittamiseen, ei ole autonominen, koska sillä ei ole itsenäistä roolia. Esimerkiksi kuluttajien ostokäyttäytymisen kontekstissa matalan autonomian tekoälyt voivat Frankin ja muiden (2023, s. 156) mukaan tarjota esimerkiksi personoituja suosituksia aiemman ostohistorian perusteella, kun taas korkean autonomian tekoälyt voivat tehdä

päätöksiä kuluttajan puolesta esimerkiksi tilaamalla automaattisesti sellaisia tuotteita, joita kuluttaja käyttää usein.

Tekoälytutkimuksessa tarjottavat tekoälyn määritelmät asettavat tekoälyn pääasiassa tietotekniikan pääkäsitteen alakäsitteeksi. Tukeudun tutkielmassani juuri tähän näkökulmaan, jossa tekoäly ei ole itse kone, koneen ominaisuus tai kyky vaan tekoäly on erilaisista tietokoneisiin liittyvistä teknologioista ja algoritmeista koostuva laaja tietotekninen teknologiakokonaisuus. Esittelemieni määritelmien pohjalta tarjoan oman tekoälyn määritelmäni, joka nojaa ylempänä esiteltyihin määritelmiin. Tässä tutkielmassa tekoälyllä tarkoitetaan useista eri teknologioista ja laskennallisista algoritmeista koostuvaa tietotekniikan teknologiaa, jolla saadaan aikaan tyypillisesti ihmisen älykkyyden mahdollistamia toimia, esimerkiksi keskustelua, tekstin tuottamista, analyysiä, todennäköisyyslaskentaa, oppimisen ja soveltamisen kaltaisia toimia, ja joka voi toimia erilaisilla autonomian tasoilla.

Tekoälyn määrittelemisen tietotekniikan teknologiaksi rajaa käsitteen tarkemmin kuin pelkäksi teknologiaksi määrittelemisen ja antaa johdonmukaisen kuvan siitä, että tekoäly on teknologian pääkäsitteen alle kuuluva alakäsite, joka on vahvasti kytköksissä tietokoneteknologiaan ja tietotekniikkaan liittyviin käsitteisiin. Pelkkään teknologian pääkäsitteeseen määrittelemisen asettaisi tekoälyn erilaisten laaja-alaisten teknologioiden kirjojen joukkoon, vaikkapa auton moottorien tai aseteknologian joukkoon, mikä vaikeuttaisi käsitteen tarkempaa rajaamista ja hahmottamista.

Tekoäly käsitteenä pitää sisällään laajan kokonaisuuden erilaisia teknologioita ja niiden erilaisia sovelluksia esimerkiksi päättelyn kaltaiseen toimintaan, tekstintuotantoon tai analyysiin liittyen. Tästä syystä *tekoäly* onkin terminä kattotermi, joka ei itsessään kerro kovinkaan tarkasti siitä, minkälaisia teknologioita tai sovelluksia sillä tarkoitetaan, kun viestinnässä käytetään *tekoäly*-termiä. Jotta tekoäly käsitteenä voidaan ymmärtää tarkemmin, on syytä perehtyä sen eri alalajeihin ja teknologioihin niiden taustalla.

Seuraavaksi esittelen generatiivisen tekoälyn, joka on tekoälyn pääkäsitettä huomattavasti tarkkarajaisempi käsite, jolla on selkeämmät ominaisuudet ja toiminnot.

3.2 Generatiivinen tekoäly

Vaikka tekoäly eri muotoineen on ollut olemassa jo kauan, tekoälystä uutisointi lisääntyi merkittävästi vuonna 2022 julkaistun ChatGPT-sovelluksen myötä. ChatGPT:tä voidaan pitää esimerkkinä generatiivisesta tekoälystä, jota voidaan soveltaa erilaisiin tarkoituksiin eri aloilla (Ray, 2023, s. 121). Generatiivinen tekoäly on tekoälymalli, jolla on mahdollista luoda erilaisia sisältöjä, esimerkiksi tekstiä, kuvia ja musiikkia, olemassa olevan datan sisältämien rakenteiden ja kaavojen perusteella.

Generatiivisia tekoälymalleja kutsutaan myös syväoppimismalleiksi, jotka koulutetaan suurilla datamäärillä (Huang & Rust, 2024, s. 5–6; Ray, 2023, s. 121). Generatiiviset tekoälymallit hyödyntävät tätä suurta datamäärää mahdollistaen ihmismäisten vastausten ja sisältöjen tuottamisen käyttäjän kehoitteiden, eli erilaisten tekoälymallille esitettävien ohjeiden ja kysymysten (Lim ja muut, 2023, s. 2), perusteella (Huang & Rust, 2024, s. 5–6; Ray, 2023, s. 121). Suuri datamäärä mahdollistaa myös generatiivisen tekoälyn nopean reagoinnin ja kommunikatiivisen vastaustyylin käyttäjän kehoitteisiin (Huang & Rust, 2024, s. 5–6).

Generatiivisen tekoälyn pääpiirteisiin lukeutuvat muun muassa reaktiivisuus ja usein myös multimodaalisuus (Sapkota ja muut, 2026, s. 6). Reaktiivisuus tarkoittaa, että generatiiviset tekoälyt eivät toimi autonomisesti vaan toimiakseen ne ovat riippuvaisia syötteistä, eli käyttäjän antamista kehoitteista. Reaktiivisuus myös tarkoittaa generatiivisen tekoälyn kohdalla sitä, ettei niillä ole pysyvää muistia aiemmasta järjestelmän tilasta eikä tavoitteiden seurantamekanismeja. Multimodaalisuudella tarkoitetaan generatiivisen tekoälyn kykyä tuottaa erilaisia tulosteita, kuten tekstiä, koodia tai kuvia, tai kykyä tuottaa esimerkiksi kuvasta tekstiä tai tekstistä kuvia.

3.3 Kielimalli

Generatiivisen tekoälyn esimerkkinä mainitun ChatGPT:n ihmismäistä tekstiä tuottava ominaisuus on teknologialtaan niin sanottu laaja kielimalli, joka on suunniteltu luonnollisen kielen prosessointiin liittyviin tehtäviin, esimerkiksi tekstin tuottamiseen ja kielen ymmärtämiseen (Ray, 2023, s. 121). Laajoista kielimalleista käytetään usein *large language model* -sanoista muodostettua *LLM*-lyhennettä, ja laajaa kielimallia voidaan pitää generatiivisen tekoälyn alalajina (ks. esim. Chen & Chan, 2024, s. 9101, 9103). Esimerkiksi ChatGPT:tä voidaan pitää multimodaalisena kielimallina (Ren ja muut, 2025, s. 126), sillä se kykenee tekstin lisäksi tuottamaan myös kuva- ja videosisältöä.

Laajat kielimallit ovat neuroverkkojen sovelluksia, jotka käsittelevät dataa järjestyksessä ja niiden koulutusprosessissa kielimallit pyrkivät ennustamaan jaksossa ilmenevien sanojen järjestyksiä aikaisemman kontekstin perusteella (Chen & Chan, 2024, s. 9103). Koulutusprosessi hyödyntää laajoja tekstiaineistoja, joista kielimalli ”oppii” toistuvia kaavoja tilastollisesti toistuvien sanojen yhteisesiintymisten perusteella. Tämän koulutusaineistossa toistuvien sanojen yhteisesiintymisten todennäköisyyslaskennan perusteella kielimalli luo kieliopillisesti ja semanttisesti yhtenäisiä uusia tekstejä. Yksinkertaistettuna voitaisiin siis sanoa, että kielimallilla tuotetut tekstit ovat todennäköisyyslaskennalla tiettyyn järjestykseen asetettuja sanajaksoja, joiden järjestyksen todennäköisyys pohjautuu valtavaan tietomäärään, jolla kielimalli on ohjelmoitu.

3.4 Tekoälyagentit ja agenttinen tekoäly

Tekoälyagentit ovat autonomisia järjestelmiä, jotka on suunniteltu kapeiden ja tarkasti määriteltyjen tehtävien suorittamiseen rajatuissa, mutta dynaamisissa digitaalisissa ympäristöissä (Ren ja muut, 2025, s. 126; Sapkota ja muut, 2026, s. 4–5). Tekoälyagenttien ominaisuuksiin kuuluvat kyky havaita strukturoituja ja epästrukturoituja syötteitä, tehdä päätelmiä kontekstuaalisesta informaatiosta ja tehdä toimia tavoitteiden

saavuttamiseksi ihmistoimijoiden puolesta (Sapkota ja muut, 2026, s. 4). Tekoälyn määritelmässä korostamani tekoälyn mahdollisuus toimia autonomisesti on tekoälyagenttien keskeinen käsitepiirre. Autonomia mahdollistaa tekoälyagenttien toimimisen käyttöönoton jälkeen joko täysin ilman ihmistä tai ainoastaan minimaalisella ihmiskäyttäjän puuttumisella sen toimintaan (Sapkota ja muut, 2026, s. 5).

Autonomian ohella tekoälyagentit eroavat perinteisistä automaatio-ohjelmista niiden mukautuvuuden osalta. Siinä missä automaatio-ohjelmat noudattavat niille ennalta määrättyjä työkulkuja, tekoälyagentit sen sijaan reagoivat älykkäästi ympäristöönsä, mikä ansiosta ne voivat tulkita dynaamisia syötteitä ja tuottaa niihin vastaavia tulosteita (Sapkota ja muut, 2026, s. 4). Tekoälyagentteja on käytössä muun muassa asiakaspalvelun automatisoinnissa, tuottavuuden tukityökaluna, organisaatioiden tiedonhaussa ja päätöksentekojärjestelmissä.

Tekoälyagenttien kapeat automaatiomahdollisuudet tehtävien suhteen ja skaalautuvuusrajoitukset monimutkaisissa ja monitasoisissa ympäristöissä ovat johtaneet kehittyneemmän autonomisen tekoälyparadigman syntymiseen, jota kutsutaan agenttiseksi tekoälyksi (Sapkota ja muut, 2026, s. 7). Agenttisella tekoälyllä tarkoitetaan autonomisia järjestelmiä, jotka pystyvät itsenäisesti suorittamaan monimutkaisia tavoitteita minimaalisella ihmisvalvonnalla muuttuvissa ja epävarmoissa ympäristöissä (Ren ja muut, 2025, s. 128). Agenttiset tekoälyt erottuvatkin siis tekoälyagenteista tehtävien, tavoitteiden ja toimintaympäristöjen monimutkaisuudessa.

Ren ja muut (2025, s. 128) määrittelevät agenttisen tekoälyn toimijuuden järjestelmän kyvyksi saavuttaa monimutkaisia tavoitteita dynaamisissa ympäristöissä siten, että toimintaa valvotaan suoraan vain rajatusti. Ren ja muut (2025, s. 128) jakavat tekoälyn toimijuuden neljään ulottuvuuteen, jotka ovat tavoitteen monimutkaisuus, ympäristön monimutkaisuus, mukautuvuus ja itsenäinen toteutus. Tekoälyn toimijuus on asteittainen skaala kiinteän luokittelun sijaan, mikä tarkoittaa, että tekoälyjärjestelmät eivät täysin jakaannu agenttisten ja ei-agenttisten järjestelmien välille, vaan toimijuuden aste

vaihtelee tekoälyjärjestelmien autonomisten kykyjen, mukautuvuuden ja tavoitteellisen päättelykyvyn kehittyessä (Ren ja muut, 2025, s. 129). Näiden kykyjen kehittyessä riittävän korkealle tasolle, tekoälyagentit muuttuvat tekoälyagenteista agenttisiksi tekoälyjärjestelmiksi.

3.5 Tekoälyn tutkimista journalismin kontekstissa

Tekoälyn tutkiminen journalismin kontekstissa on jakaantunut karkeasti kahteen tutkimussuuntaan, joista yksi suunta keskittyy tutkimaan tekoälyteknologian vaikutusta journalismin toteuttamiseen, journalistiyhteisöön sekä journalistisen työn luonteeseen, ja toinen, tälle tutkielmalle relevantimpi, suunta keskittyy selvittämään sitä, miten journalistit tiedottavat tekoälystä (ks. Hollanek ja muut, 2025, s. 6394). Tekoälyaiheiden uutisoinnin tutkimiseen keskittynyt tutkimussuunta pyrkii selvittämään sitä, miten median tekoälyn kuvaukset muokkaavat yleistä ymmärrystä tekoälystä ja miten ne vaikuttavat tekoälyteknologian kehitykseen ja sen sääntelyyn (Hollanek ja muut, 2025, s. 6394). Oma tutkimusaiheeni asettuu osaksi tätä jälkimmäistä tutkimussuuntaa, jonka viimeaikaista tutkimusta esittelen tässä alaluvussa.

Tekoälyaiheiden esittämistä sekä käsittelyä journalismissa on tutkittu erilaisten viitekehysten ja aiheiden näkökulmista. Tekoälyaiheisen journalismin tutkimuksessa on käsitelty esimerkiksi perinteisen median tapoja hillitä tekoälykeskusteluun olennaisesti liittyviä kiistanalaisuuksia, sukupuolen representaatioita tekoälyaiheisten uutisten kuvissa, tekoälyn ja ChatGPT:n representaatioita Iso-Britannian uutisissa sekä tekoälyn kehystystä ja toimijuutta lääketieteelliseen kontekstiin asettuvissa uutisissa (ks. esim. Bunz & Braghieri, 2022; Chen ja muut, 2024; Dandurand ja muut, 2023; Roe & Perkins, 2023).

Aikaisempi tutkimus median tekoälyaiheiden käsittelystä on painottunut erityisesti kehystyksen tutkimiseen (ks. esim. Duberry & Hamidi, 2021; Hollanek ja muut, 2025; Nguyen, 2023; Nguyen & Hekman, 2024; Xian ja muut, 2024). Monissa näissä kehystyksiin keskittyvissä tutkimuksissa eräänlaisena toistuvana teemana voidaan pitää teknologian

lukutaidon, datalukutaidon tai tekoälylukutaidon merkitysten korostumista niin itse journalistien osalta uutisten kirjoittamisprosessissa kuin myös yleisön osalta uutisten lukemisen ja ymmärtämisen näkökulmista. Lisäksi tutkimuksissa pidetään erityisen tärkeänä journalistien roolia kyseisten teknologiaan liittyvien lukutaitojen edesauttamisessa ja ylläpitämisessä kansalaisten keskuudessa.

Kehystykseen keskittyvässä tekoälyjournalismitutkimuksen suuntauksessa esimerkiksi Xian ja muut (2024) ovat tutkineet sitä, minkälaisia aiheita ja mihin sävyyn generatiivisesta tekoälystä uutisoidaan ja kuinka uutisten julkaisut sijoittuvat ajallisesti ja tilallisesti eri maiden välillä. Xian ja muut (2024, s. 1) esittävät, että medialla on tärkeä asema tiedon levittämisessä, ja media muovaa yleisönsä käsityksiä uusista teknologioista, mikä ylläpitää samalla suuren yleisön teknologian lukutaitoa. He toteavat, että suuren yleisön tietämys generatiivisesta tekoälystä ja sen taustalla olevasta teknologiasta on vasta orastavalla tasolla.

Xianin ja muiden (2024, s. 1) mukaan analysoimalla median uutisointia generatiivisesta tekoälystä voidaan ymmärtää, kuinka yleisön altistuminen median tulkinnoille tekoälyä koskevista aiheista saattaa muokata yleistä mielipidettä sekä yleisön arvioita generatiivisesta tekoälystä. Xian ja muut (2024, s. 10) saivat analyysissään selville, että eri maiden välillä on suuria eroja siitä, mistä kehystystä generatiivisesta tekoälystä uutisoidaan. Yhdysvallat ja Intia keskittyvät uutisoinnissaan liiketoimintaan ja teknologian kehitykseen, kun taas Iso-Britannia, Australia ja Kanada käsittelevät generatiiviseen tekoälyyn liittyviä sääntely- ja turvallisuusulottuvuuksia, minkä lisäksi Australian aineistosta nousi esiin myös koulutusnäkökulma.

Nguyen (2023) käsittelee tutkimuksessaan sitä, kuinka media kehystää dataan liittyviä riskejä uutisoidessaan tekoälystä ja massadatasta. Nguyen (2023, s. 26) tuo esiin neljä riskiluokkaa, jonka kautta tutkitut uutiset kehystävät datariskejä ja siihen liittyviä eettisiä haasteita: yksityisyyden loukkaaminen sekä valvonta, datavinouma sekä algoritmiin perustuva syrjintä, kyberturvallisuus ja informaatiohäiriö. Nguyen pitää kategorisointia

lähtökohtana sellaisten lähestymistapojen kehittämiseksi, jotka tekevät dataan liittyvistä riskeistä ymmärrettäviä monimuotoisille yleisöille ja parantavat suuren yleisön datalukutaitoa.

Datalukutaidon Nguyen (2023, s. 7) määrittelee maallikkojen käsitteelliseksi tiedoksi ihmisen toiminnasta muodostettavan datan käytöstä ja automatisoinnista sekä kyvyksi muodostaa kriittisiä mielipiteitä aiheesta. Datalukutaito käsittelee hänen mukaansa muun muassa sitä, miten datakäytännöt ja algoritmiset järjestelmät vaikuttavat esimerkiksi yksilön tietoturvaan, miten organisaatiot kerryttävät valtaa datan ja tekoälyn avulla ja miten kyseinen valta vaikuttaa yhteiskuntaan sekä sitä, miten kansalainen voi suojella itseään haitallisilta datakäytännöiltä. Nguyen (2023, s. 7–8, 26) korostaa median tärkeää roolia ja yhteyttä kansalaisten datalukutaidossa. Näen, että datalukutaito on oleellisesti yhteydessä tekoälylukutaitoon, koska sen hallitseminen auttaa ymmärtämään tekoälyä sekä arvioimaan tekoälyn yhteiskunnallisia vaikutuksia.

Nguyen on yhdessä Hekmanin kanssa tutkinut myös sitä, miten mediat kehystävät tekoälyä uutisissaan (Nguyen & Hekman, 2024). Tutkimuksen keskiössä ovat kysymykset siitä, miten uutismediat tekevät tekoälyn lukijalle ymmärrettäväksi, mistä tekoäly tulee, mihin tekoäly vaikuttaa sekä minkälaisia hyötyjä ja riskejä tekoälyyn liittyy (Nguyen & Hekman, 2024, s. 438). Nguyen ja Hekman (2024, s. 439) nostavat uutismedian keskeiseen rooliin tekoälyteknologian merkityksen muodostamisessa yhteiskunnallisella tasolla korostaen uutismedioiden rakentavan aktiivisesti diskursseja tekoälyn hyödyistä ja riskeistä. Nguyenin (2023) aikaisemman tutkimuksen tavoin Nguyen ja Hekman (2024, s. 440, 449) yhdistävät tekoälyuutisoinnin uutisten tekijöiden ja uutisten lukijoiden datalukutaitoon; journalistit toimivat teknologian kriittisinä havainnoitsijoina, joiden on ymmärrettävä tekoälyn ominaisuudet ja vaikutukset ainakin käsitetasolla, mutta ei välttämättä teknisen tiedon tasolla, jotta lukijat ymmärtäisivät aiheen paremmin.

Datalukutaitoa lähellä olevaa tekoälylukutaitoa on tutkittu journalistien näkökulmasta. Hollanek ja muut (2025) tutkivat toimittajien näkemyksiä tekoälylukutaitoon liittyvistä

huolista, tarpeista ja vastuista. Tekoälylukutaidon Hollanek ja muut (2025, s. 6395) määrittelevät monitahoiseksi tiedosta ja kyvyistä koostuvaksi taitokokonaisuudeksi: tekoälylukutaito on yleinen ymmärrys tekoälyjärjestelmien toiminnasta, kyky käyttää tekoälyjärjestelmiä tehokkaasti ja vastuullisesti sekä kyky arvioida tekoälyn laajempia yhteiskunnallisia vaikutuksia. Tutkimukseen osallistuneet toimittajat korostivat kriittisen tekoälylukutaidon tärkeyttä etenkin vastuullisen journalismin edellytyksenä (Hollanek ja muut, 2025, s. 6403). Kriittinen tekoälylukutaito painottaa etenkin tekoälyn ymmärtämistä laajempina teollisena ja poliittisena kenttänä, jolla on suoria ja epäsuoria pitkän aikavälin vaikutuksia niin yhteiskuntaan kuin ympäristöön.

Vaikka tutkimuksissa ei suoranaisesti mainittaisi tekoälylukutaitoa, siihen viitataan usein myös epäsuorasti. Esimerkiksi Dandurand ja muut (2023, s. 4) tuovat esiin sen, että journalistin työhön kuuluu kyky esittää monimutkaiset kokonaisuudet ymmärrettävässä ja kiinnostavassa muodossa maallikoille, ja tässä prosessissa tekoälyaiheen hyvä tuntemus katsotaan perusedellytykseksi. Aiheen hyvä tuntemus mahdollistaa Dandurandin ja muiden mukaan esimerkiksi haastattelutilanteissa merkityksellisten kysymysten esittämisen haastattelun kohteena olevalle teknologia-asiantuntijalle. Epäsuorasti tekoälylukutaitoon ja siihen liittyviin rooleihin viittaavat myös Roe ja Perkins (2023), jotka tutkivat tekoälyn ja ChatGPT:n representaatioita uutismediassa. Roen ja Perkinsin (2023, s. 8) mukaan heidän tutkimuksensa alleviivaa median, asiantuntijoiden ja yleisön vastuuta ylläpitää ja edistää tasapainoista tietoon perustuvaa monipuolista keskustelua tekoälystä ja sen mahdollisista vaikutuksista. Heidän mukaansa kyseinen tutkimus toimii ajankohtaisena muistutuksena siitä, että tekoälyn medianäkyvyyttä on pystyttävä arvioimaan kriittisesti. Nämä kuvaukset ovat linjassa tekoälylukutaidon vaatimusten ja ominaisuuksien kanssa.

Omaa tutkimusaiheittani lähimpänä on Duberryn ja Hamidin (2021) tutkimus, joka selvittää sisällönanalyysin keinoin, kuinka eurooppalaiset ja yhdysvaltalaiset lehdet määrittivät sekä kehystivät tekoälyä uutisissaan koronapandemian aikana, koska he sivuavat myös tekoälystä käytettyjä termejä. Duberryn ja Hamidin (2021, s. 767–770) mukaan

tekoälystä puhuttiin koronapandemian aikaisissa uutisissa monin erilaisin termein, kuten *koneoppimisena*, *algoritmina*, *robottina* tai *syväoppimisena tekoälyn* ja tekoälyn lyhenteen (*AI*) ollessa yleisimmin käytettyjä termejä. Termejä, kuten *algoritmia* tai *syväoppimistä*, käytettiin sellaisissa konteksteissa, joissa pyrittiin kuvailemaan jotain tiettyä tekoälyn kyvykkyyttä tai ominaisuutta tarkemmin.

Duberryn ja Hamidin (2021, s. 768–769) mukaan ihanteellista olisi, jos uutisissa mainitaisiin, kuinka erilaiset tekoälytermit liittyvät toisiinsa, mutta selitys yhteyksille oli sisälletty vain pieneen osaan uutisista. Uutisissa ei myöskään tarjottu määritelmää sille, mitä tekoälyllä tarkoitettiin, minkä he näkevät ongelmallisena. Syytä määritelmien puuttumisesta Duberry ja Hamidi (2021, s. 769) eivät kuitenkaan aseta medialle. He huomauttavat, että edes tekoälyyn keskittyneet asiantuntijatkaan eivät ole yksimielisiä tekoälyn määritelmästä, mutta toivovat, että lukija voisi silti tukeutua uutista lukiessaan jonkinlaiseen tekoälyn määrittelevään lähteeseen. Utisten selkeys paranisi heidän mukaansa myös, jos niissä tehtäisiin selvä ero sille, käsitelläänkö tekoälystä puhuttaessa tekoälyn erikoisalaa, tekoälyn toiminnallisuuksia vai tekoälyn soveltamista ja sovelluksia.

3.6 Terminologista tutkimusta tekoälystä

Tekoälyä tutkitaan paljon teknologisesta sekä tekoälyn käytännön sovellusten näkökulmista, mutta tekoälystä on tehty aiempaa tutkimusta myös terminologisesta näkökulmasta. Tutkimuksia tarkastellessa korostuu se, että tekoälytermien kritiikissä keskitytään etenkin antropomorfististen termien käytön ongelmallisuuteen. Antropomorfismilla tarkoitetaan sitä, kun ihmisille ominaisia käsitteitä ja ominaisuuksia liitetään elottomiin objekteihin (ks. esim. Tieteen termipankki, n.d.b).

Deroy (2023) käsittelee tutkimuksessaan sitä, voiko tekoälyä kuvailla ihmistermein. Deroy (2023, s. 881–882) nostaa esiin *luotettava*-sanankäytön tekoälyn kuvailussa argumentoiden, että luotettavuuden käsite pitää sisällään sellaisia ominaisuuksia, kuten rehellisyyden, hyvántahtoisuuden sekä moraalisten ja sosiaalisten normien

noudattamisen, joihin ainoastaan ihminen on kykenevä. Deroy (2023, s. 888) päätyy analyysissään siihen, että tekoälyn kuvailulle ihmistermein on olemassa vain vähän puoltavia argumentteja ja toteaa, että ihmistermien käyttö luo epäselvyyttä siitä, mitä tekoälyllä on oikeastaan mahdollista edes tehdä. Lisäksi Deroy (2023, s. 888) kyseenalaistaa *älykkyyden* termin käytön tekoälyn kontekstissa ja tuo esiin, että myös älykkyys on ihmisyyteen nojaava termi, jonka arviointiin tulisi mahdollisesti käyttää samoja perusteita kuin *luotettavuus*-sanon käyttöön kohdistuvassa kritiikissä.

Maleki ja muut (2024) ovat tutkineet tekoälyn *hallusinaatio*-termin käyttöä eri tiedejulkaisuissa ja heidän tutkimuksensa korostaa termin käytön epäjohdonmukaisuutta. Malekin ja muiden mukaan (2024, s. 133) *hallusinaatio*-termistä ei ole yhdenmukaista määritelmää. *Hallusinaatio*-termillä viitataan tekoälyn tekemiin virheisiin tai tekstisisältöihin, jotka voivat kuulostaa uskottavilta, mutta eivät ole tarkkoja tai totuudenmukaisia (Maleki ja muut 2024, s. 136). Termi koetaan ongelmalliseksi, koska sen käytön nähdään leimaavan kielteisesti mielenterveysongelmista kärsiviä ja koska termi on harhaanjohtava tekoälyn aistihavaintokyvyn puutteen takia – tekoälyn tekemät virheet ovat peräisin datasta ja kehoitteista eikä aistiärsykkeistä (Maleki ja muut, 2024, s. 133). Tämä on linjassa Deroy'n (2023) esiin tuoman kritiikin kanssa siitä, että tekoälykäsitteiden kuvailu ihmistermein voi olla harhaanjohtavaa ja ongelmallista.

Myös Rehak (2023) suhtautuu kriittisesti tekoälystä käytettäviin ihmistermeihin. Toisin kuin esimerkiksi matematiikan saralla, jossa useimmat tekniset termit ovat tunnistettaviksi sellaisiksi, tekoälyn alalla turvaudutaan antropomorfismiin (Rehak, 2023, s. 89). Tieteessä voidaan käyttää arkikielestä tuttuja sanoja, mikä muuttuu silloin ongelmalliseksi, jos samalla sanalla on täysin eri merkitykset tai jopa päinvastaiset merkitykset; esimerkiksi *läpinäkyvyys* tarkoittaa tietokonetieteissä näkymättömyyttä, kun taas politiikan tutkimuksessa se tarkoittaa päinvastaisesti näkyvää (Rehak, 2023, s. 89–90). Tästä arkikielen käytöstä seuraa Rehakin mukaan ongelmia silloin, kun termit siirretään huolimattomasti toisiin aloihin tai takaisin arkikieleen esimerkiksi tiedejournalismin muodossa.

Esimerkkejä tällaisista viestinnän näkökulmasta ongelmallisista tekoälyyn liitetystä termeistä on Rehakin (2023, s. 89) mukaan *tunnistaminen, oppiminen, toimiminen, päättäminen, muistaminen, ymmärtäminen ja älykkyys*. Esimerkiksi oppimista Rehak (2023, s. 93–94) kuvailee kognitiiviseksi ja sosiaalisiksi käsitteeksi, jolla tarkoitetaan sitä prosessia, jossa ihminen kerryttää uutta tietoa. Hänen mukaansa tähän uuden tiedon kerryttämiseen liittyvät yksilön lisäksi oppimisen taustalla olevat ryhmät, kulttuurit, kontekstit ja laaja määrä erilaisia oppimisen prosesseja, mikä eroaa merkittävästi siitä, kuinka tekoälyneuroverkkoa ohjelmoidaan manuaalisesti tai automaattisesti testidatalla.

Epäkuvaavien termien käytössä Rehak (2023, s. 96) näkee vaarana sen, että yleistyessään harhaanjohtavat antropomorfistiset metaforat voivat luoda vääristynyttä tietoa, joka muodostuu pohjaksi esimerkiksi poliittisissa, sosiaalisissa tai johtamisen päätöksissä. Rehak (2023, s. 98–99) tiedostaa, että tässä vaiheessa moni tekoälytermi on vakiintunut tieteessä, joten niiden kaikkien uusiminen olisi epärealistista. Hänen mukaansa tekoälytermien tulisi olla mahdollisimman lähellä käsitteen teknologista todellisuutta, mutta teknisten termien tulisi esimerkiksi välttää yleiskielisiä metaforina ymmärrettäviä konnotaatioita, antropomorfismeja ja populaaritieteessä laajasti käytettäviä käsitteitä, joilla on kuitenkin täysin eri merkitys.

4 Tekoölyyaiheisten verkkoartikkelien käsitteiden analyysi

Tässä tutkielmassa tavoitteena on selvittää, millaisia tekoölykäsitteitä ja -termejä ilmenee Ylen yleistajuisissa tekoölyyaiheisissa verkkoartikkeleissa. Tämä luku keskittyy tekoölykäsitteiden selvittämiseen systemaattisen käsiteanalyysin avulla, jolla vastaan näin olen ensimmäiseen tutkimuskysymykseen siitä, mitä tekoölykäsitteitä ja -termejä ilmenee Ylen tekoölyä käsittelevissä verkkoartikkeliteksteissä.

Ennen analyysin tulosten kuvaamista kerron aluksi aineiston muodostamisesta ja aineistolle suorittamastani puoliautomaattisesta terminpoiminnasta sekä terminpoimintaan liittyvistä päätöksistä ja rajauksista. Sen jälkeen kuvaan lyhyesti käsitejärjestelmän graafiseen kuvaamiseen käyttämäni satelliittimallin luomisprosessia ja kerron, kuinka satelliittimallin merkintöjä tulee tulkita. Lopuksi siirryn esittelemään käsiteanalyysin tuloksia esittelemällä ensin havainnollistavan kokonaiskuvan käsitejärjestelmästä, jonka jälkeen siirryn esittelemään käsitejärjestelmän osioita yksityiskohtaisemmin.

4.1 Terminpoiminnan toteuttaminen

Toteutin terminpoiminnan puoliautomaattisesti Sketch Engine -ohjelmalla. Terminpoimintaa varten aineisto oli ensin valmisteltava sellaiseen muotoon, jolla ohjelma kykenee lukemaan aineistoa ja poimimaan termejä siitä. Lisäsin verkkoartikkelien linkit Sketch Enginen korpuksenluontisovellukseen, joka muodosti erillisistä verkkoartikkeleista yhtenäisen tiedoston, jossa kaikkien aineiston artikkelien tekstit ovat samassa tiedostossa. Luomalla korpuksen varmistin samalla, että mahdollisilla muutoksilla alkupe- räisiin artikkeleihin ei ole vaikutusta analyysin tuloksiin.

Korpuksen luomisen jälkeen toteutin terminpoiminnan Sketch Engine -ohjelmalla. Asetin poimittavien termien määrän 1000 termiin, ja ohjelma poimi korpuksesta 1000 yksisanaista termiehdokasta ja 1000 monisanaista termiehdokasta. Rajasin poimittavien termien määrän yhteensä 2000 termiin sopeuttaakseni työtaakkaa sekä resursseihin että

valittuun terminpoimintamenetelmään nähden. Termimäärän rajauksesta on huomiotava se mahdollisuus, että laajemmalla rajauksella olisi termilistaan voinut päätyä enemmän tutkimukselle sopivia termejä. Tällä rajauksella myös siivottavien termien määrä olisi toisaalta samalla kasvanut merkittävästi. Potentiaaliseen työmäärän näkökulmasta manuaalinen terminpoiminta olisi voinut olla mielekkäämpi valinta ajankäytön kannalta laajempaan rajaukseen verrattuna, mutta koska valitsin tutkimuksen terminpoimintataksi puoliautomaattisen terminpoiminnan, päädyin tiukempaan termimäärän rajaukseen.

Puoliautomaattisen terminpoiminnan tuloksena syntyi kaksi termilistaa: yksisanaiset termiehdokkaat ja monisanaiset termiehdokkaat. Latasin tietokoneelleni molemmat termilistat Excel-tiedostoina, jotka säilytin muokkaamattomina, minkä lisäksi loin termilistoista identtiset kopiot, joita lähdin sen jälkeen muokkaamaan valmiiksi, siivotuiksi termilistoiksi. Toteutin termilistoille siivouksen, jossa poistin tutkimukselleni sopimattomat sanat ja epäolennaiset termit. Termilistaa siivotessa käytin termin päättelemiseen oman intuitioni lisäksi Sketch Enginen ominaisuutta, joka näyttää termin sijainnin tekstin kontekstissa, mistä oli mahdollista päätellä, onko kyseessä tekoälytermi, erisnimi, muun erikoisan termi tai muu aiheeseen liittymätön sana. Esimerkkejä pois jättämistäni erisnimistä, muiden erikoisalojen termeistä tai muista aiheeseen liittymättömistä sanoista ovat *Nvidia*, *Macron*, *gigawatti*, *ajatushautomo*, *kotihoidon asiakas* ja *käyttökielto*. Käytin termin päättelemisen tukena lisäksi vapaamuotoista tiedonhakua, mikäli en tunnistanut termin taustalla olevaa käsitettä. Yksisanaisten termien termilistaan valikoitui lopulta 51 termiä ja monisanaisten termien termilistaan 10 termiä eli yhteensä 61 tekoälytermiä, jotka kaikki sijoittuivat vähintään kerran käsiteanalyysissä toteuttamaani satelliittimalliin.

Siivousvaiheessa termeistä karsiutui pois myös sellaisia termejä, jotka olisivat mahdollisesti sopineet osittain jonkin käsitesuhteen kautta myös osaksi muodostamaani käsitejärjestelmää. Tällaiset termit rajasin pois siksi, koska ne sopivat luontevammin jonkin toisen erikoisan alaisuuteen, ja tämä rajaus piti satelliittimallin fokuksen tiukemmin

tekoälyn erikoisalassa. Tällaisista termeistä esimerkkeinä voidaan pitää muun muassa seuraavia käsitteitä: *autonominen droonisodankäynti*, *humanoidirobotti*, *kyberfyysiset järjestelmät* ja *memristori*. Esimerkiksi autonomisen droonisodankäynnin käsitteessä autonomia saavutetaan tekoälyn avulla, mutta käsite sopii kuitenkin paremmin esimerkiksi sodankäynnin pääkäsitteen alle kuin tekoälyn alaisuuteen.

4.2 Satelliittimallin muodostaminen

Koostin valmiiden termilistojen avulla yhtenäisen kokonaiskuvan aineiston käsitteistä koostuvasta kokonaiskäsitejärjestelmästä. Käsitejärjestelmän graafisen esitysmuodon toteuttamiseen käytin Miro-ajatuskarttasovellusta, jolla loin satelliittimallin. Satelliittimallin luomisessa sijoitin ensin keskusnoodiin tekoälyn pääkäsitteen, minkä jälkeen kiinnitin keskusnoodin molemmille puolille apunoodeja, joihin asetin seuraavat käsitesuhteet valmiiksi odottamaan sopivien käsitteiden sijoittelua: ainesosasuhde, alkupe-räsuhde, arvosuhde, edustussuhde, geneerinen käsitesuhde, kausaalisuuhde, kehitys-suhde, koostumussuhde, lisäosasuhde, ominaisuussuhde, riippuvuus- ja korrelaatio-suhde, sijaintisuuhde, temporaalinen käsitesuhde, toimintasuhde ja transmissiosuhde.

Analyysissä otin siten huomioon kaikki ylempänä kuin myös teoriaosuudessa mainitut käsitesuhteet, mutta käsiteanalyysin tuloksena käsitteitä sijoittui vain osaan mahdollisista käsitesuhteista. Tämän takia valmiissa satelliittimallissa esiintyy geneerinen käsite-suhde ja ontologisista käsitesuhteista vain ainesosasuhde, koostumussuhde, ominai-suussuhde, riippuvuus- ja korrelaatio-suhde, sijaintisuuhde ja toimintasuhde.

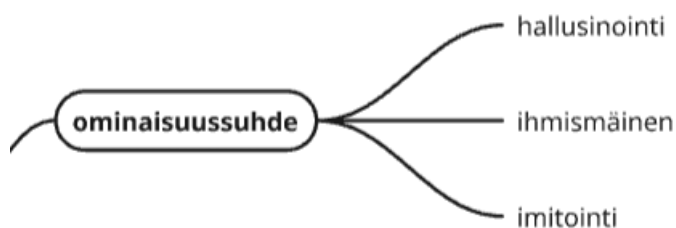
Osa käsitteistä oli luontevaa sijoittaa useamman kuin yhden käsitesuhteen alaisuuteen. Esimerkiksi tekoälyalan käsite oli luontevaa liittää tekoälyn pääkäsitteeseen sijaintisuhteen kautta, mutta se oli mahdollista yhdistää tekoälyn pääkäsitteeseen myös toiminta-käsitesuhteeseen kuuluvan paikkasuhteen kautta. Tällaista moneen eri käsitesuhteeseen sijoittamista olisi voinut toteuttaa useammankin käsitteen kohdalla. Käsitejärjestelmää muodostaessani kuitenkin huomasin, että satelliittimallin selkeyden ja luettavuuden

kannalta oli mielekkäämpää tehdä ratkaisu yksinkertaisemmasta satelliittimallista. Tästä syystä vain osa käsitteistä ilmenee useamman kuin yhden käsitesuhteen yhteydessä, vaikka loogisesti voisi olla perusteltua toistaa muitakin käsitteitä jonkin toisen käsitesuhteen kohdalla esimerkiksi geneerisenä tyyppittelynä.

Satelliittimallin rakenteen pohtimisen kohdalla huomasin myös konkreettisesti, että käsitejärjestelmä on käytännössä vain karkea yksinkertaistus todellisuuden tilasta, sillä mahdollisimman kattavaksi tehtynä satelliittimallista olisi muodostunut iso, risteävä ja itseään toistava kokonaisuus. Koska tutkimukseni tarkoituksena ei ollut tehdä täydellistä kokonaiskuvaa kaikista tekoälyn erikoisalan tekoälykäsitteistä vaan muodostaa yhtenäisen kokonaiskäsitejärjestelmä Ylen yleistajuisissa verkkoartikkeleissa esiintyvistä tekoälykäsitteistä, pidän tätä rajausta tarkoituksenmukaisena tutkielmani kontekstissa.

Käsiteanalyysiä toteuttaessani otin huomioon sen, mitä tietoa termin konteksti artikkeleissa tarjosi käsitteen hierarkiasta tai käsitesuhteista toisiin käsitteisiin. Tämän lisäksi käytin käsitesuhteiden ja hierarkian selvittämisen apuna vapaamuotoista tiedonhakuja sekä omaa tulkintaani. Toisin sanoen tekoälykäsitteiden sijainti luomassani satelliittimallissa pohjautuu aineistossa käytettyihin määritelmiin ja kuvailuihin käsitteiden hierarkiasta tai käsitesuhteista, tiedonhakuun ja omaan pohdintaani perustuvaan tulkintaani.

Satelliittimallin merkintätapojen tulkitsemisen olen pyrkinyt pitämään helposti ymmärrettävänä. Keskusnoodi, eli tekoälyn pääkäsite, sijoittuu satelliittimallin keskelle. Keskusnoodiin on yhdistetty satelliittinoodia, joihin on kirjoitettu käsitesuhteen nimi. Satelliittinoodien lisäksi käytän apunoodia selventämään käsitesuhteita sekä auttamaan käsitteiden jaottelussa. Kuten Kuviossa 1 näkyy, käsitesuhteita sisältävät satelliittinoodit olen merkinnyt lihavoidulla tekstillä sekä noodeja ympäröivällä soikealla kehällä, jotta kyseiset elementit erottuvat käsitteistä. Olen merkinnyt myös apunoodit samaa tapaa käyttäen. Käsitteitä ilmaisen satelliittimallissa pelkällä tyyllittelemättömällä tekstillä, jonka ympärillä ei ole kehää. Käsitteet olen yhdistänyt käsitesuhteisiin tai apunoodeihin toisiinsa yhdistyvillä viivoilla.



Kuvio 1. Esimerkki satelliittimallin merkintätavoista.

Termit on esitetty satelliittimallissa niiden taivuttamattomassa perusmuodossa. Satelliittimallin tulkinnassa on huomioitava se, kun käsite on sijoitettu satelliittimalliin siten, että käsitettä ilmaisee kaksi eri termiä, esimerkiksi *tekoäly-yhtiö / -yritys* tai *kielimalli / laaja kielimalli*. Tässä satelliittimallissa ilmaisuilla tarkoitetaan aineiston artikkelien perusteella käytännössä samaa asiaa, vaikka käsitteitä voidaan pitää todellisuudessa erillisinä, mutta läheisinä käsitteinä. Tällaiset synonyymeillä merkityt käsitteet on laskettu erillisiksi termeiksi terminpoiminnassa. Tästä eteenpäin viitataan luettavuuden kannalta *kielimalli / laaja kielimalli* -termipariin pelkästään kielimalli- tai laaja kielimalli -nimityksillä.

Seuraavaksi esittelen kokonaiskuvan systemaattisella käsiteanalyysillä luomastani käsitejärjestelmästä. Sen jälkeen siirryn esittelemään käsitejärjestelmää yksityiskohtaisemmin edeten geneerisistä käsitesuhteista toimintasuhteisiin ja lopulta muihin käsitesuhteisiin. Tulosten esittelemisen jälkeen kokoan yhteen havainnot, joita tuloksista on tulkittavissa.

4.3 Satelliittimalli verkkoartikkelien tekoälykäsitteistä

Kuvio 2 esittää kokonaiskuvaa aineiston tekoälykäsitteistä muodostamastani kokonaiskäsitejärjestelmästä. Kuvion tarkoituksena on toimia havainnollistavana yleiskatsauksena satelliittimallin koosta ja rakenteesta, eikä kuviota ole tästä syystä tarkasteltavaksi yksityiskohtaisesti. Yksityiskohtia tullaan tarkastelemaan seuraavien alalukujen

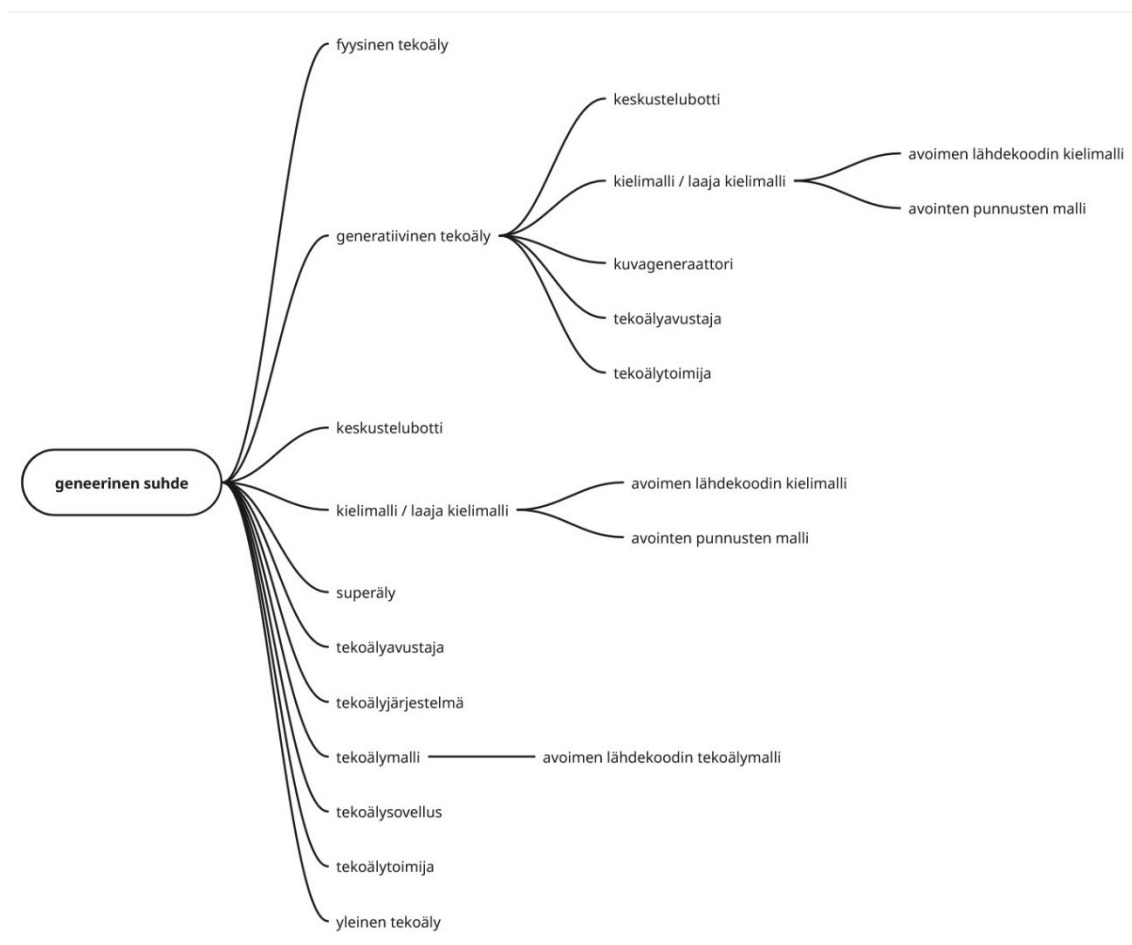
Satelliittimallin yksityiskohtaisemman tarkastelemisen mahdollistamiseksi esittelen käsittejärjestelmän pienemmissä osissa seuraavissa alaluvuissa. Aloitan tarkastelun ensiksi geneerisistä käsitesuhteista, joista tarkempaan tarkasteluun otan laajan kielimallin ja generatiivisen tekoälyn käsitteet käsitesuhteineen. Sen jälkeen esittelen tarkemmin toimintasuhteet. Lopuksi esittelen muut käsitesuhteet, jotka pitävät sisällään ainesosasuhteen, koostumussuhteen, ominaisuussuhteen, riippuvuus- ja korrelaatio-suhteen sekä sijaintisuhteen.

4.4 Satelliittimallin geneeriset käsitesuhteet

Geneeriset käsitesuhteet käsittelevät tekoälyä ilmiönä sen eri tyyppien näkökulmasta. Tarkastelussa ovat siten erilaiset tekoälyteknologiat sekä tekoälyn sovellukset, jotka tekoälyteknologiat mahdollistavat. Tässä luvussa esittelen satelliittimallin geneerisistä käsitesuhteista kokonaiskuvan sekä siitä tehdyn yksinkertaistetun version. Geneeristen käsitesuhteiden käsitteiden yleisen esittämisen jälkeen otan kahdessa alaluvussa tarkasteluun tässä käsitesuhteessa esiintyvän laajan kielimallin sekä generatiivisen tekoälyn käsitteet.

Kuviossa 3 olen kuvannut erilaiset aineistossa ilmenevät tekoälytyypit, minkä lisäksi kuvio sisältää myös käsitteiden välisiä käsitesuhteita. Muita kuin geneerisiä suhteita ilmenee generatiivisen tekoälyn käsitteen ja laajan kielimallin käsitteen kohdalla. Generatiivisen tekoälyn käsitteeseen yhdistyy muita tekoälykäsitteitä koostumussuhteen, ominaisuussuhteen ja toimintasuhteen kautta. Laajan kielimallin käsitteen yhteydessä esittelen tyypittelyn lisäksi koostumussuhteisia ja ominaisuussuhteisia käsitteitä.

Geneeristen suhteiden tarkastelun helpottamiseksi olen luonut yksinkertaistetun version geneerisistä käsitesuhteista. Kuviossa 4 esitän supistetun version, josta puuttuvat kokonaiskäsitejärjestelmässä käsitteiden välillä olevat muut kuin geneeriset käsitesuhteet. Kokonaiskuvassa laajan kielimallin kohdalla on käytetty selkeyttämisen vuoksi geneerisen käsitesuhteen ilmaisemiseksi apunoodia, mutta tässä versiossa selkeyttämiselle ei ole tarvetta, ja alakäsitteet esitän sen kohdalla hierarkiaa esittäville viivoilla.



Kuvio 4. Supistettu versio geneerisistä käsitesuhteista.

Geneerisiin käsitesuhteisiin sijoittamiseen käytin jakoperusteena sitä, onko kyseessä tekoälykäsite, joka edustaa tekoälyteknologiaa tai sen sovellusta, tekoälyn kehittyneisyyden tasoa tai tekoälyteknologian sovellusympäristöä. Systemaattisen käsiteanalyysin perusteella aineistossa korostuvat erityisesti erilaiset tekoälyteknologian ja -sovellusten

tyypittelyt tekoälyn kehittyneisyyden tason ja tekoälytekniikan sovellusympäristön tyypittelyiden saadessa vähemmän painotusta. Erilaisiin tekoälytekniikan tai -sovellusten käsitteisiin kuuluvaksi tulkitsen generatiivisen tekoälyn, keskustelubotin, laajan kielimallin, tekoälyavustajan, tekoälyjärjestelmän, tekoälymallin sekä avoimen lähdekoodin tekoälymallin, tekoälysovelluksen ja tekoälytoimijan. Myös avoimen lähdekoodin kielimalli sekä avointen punnusten malli lasketaan tekoälytekniikan tai -sovellusten tyypeiksi, mutta käsittelen ne tarkemmin seuraavassa alaluvussa. Tekoälyn kehittyneisyyden tasoon kuuluvat superälyn ja yleisen tekoälyn käsitteet, ja tekoälytekniikan sovellusympäristöön lasken fyysisen tekoälyn käsitteen.

Generatiivinen tekoäly sijoittuu aineiston perusteella luontevasti tekoälytekniikaksi tai sen sovellukseksi, vaikka tarkempaa määritelmää käsitteelle ei aineistossa tarjota. Aineiston artikkeleissa generatiivista tekoälyä kuvaillaan kuitenkin esimerkiksi sellaisessa asiayhteydessä, jossa helsinkiläiset koulut ovat keskeyttäneet tekoälyä käyttävien työkalujen ja sovellusten käytön, johon lukeutuvat artikkelin mukaan generatiivisen tekoälyn työkalut, kuten Microsoftin Copilot. Aineistossa pohditaan myös esimerkiksi generatiivisen tekoälyn vaikutusta ihmisten ajattelukyvyille.

Keskustelubotit ovat aineiston artikkelien konteksteissa tekoälysovelluksia, jonka kanssa käyttäjä voi käydä keskustelua, ja jotka voivat artikkelien mukaan esimerkiksi imitoida julkisuuden henkilöitä. Keskustelubotit sijoittuvat artikkelien perusteella myös generatiivisen tekoälyn alakäsitteeksi. Esimerkiksi Open AI:n ChatGPT-sovellukseen viitataan aineistossa keskustelubotti-termillä sellaisessa kontekstissa, jossa käsitellään generatiivisten tekoälyjärjestelmien sääntelyä, miksi asetin keskustelubotin myös generatiivisen tekoälyn tyyppiä.

Laajaa kielimallia kuvataan aineistossa useassa kohdassa sovellukseksi, jolla voidaan esimerkiksi tuottaa ja muokata erilaisia tekstejä, kuten esseitä. Toisaalta osassa artikkeleissa kielimalli- tai laaja kielimalli -termillä viitataan myös erilaisiin generatiivisiin tekoälyinä pidettyihin sovelluksiin, kuten ChatGPT:hen, minkä perusteella laaja kielimalli

käsitteenä sijoittuu satelliittimallissa myös generatiivisen tekoälyn alakäsitteeksi. Lisäksi osassa artikkeleista kielimalleista sanotaan, että ne pohjautuvat generatiiviseen tekoälyyn, kun taas ainakin yhdessä artikkelissa kielimallin sanotaan olevan generatiivisen tekoälyn näyttävin esiintymismuoto. Tämä teki generatiivisen tekoälyn ja laajan kielimallin sijoittamisesta hankalaa, sillä se antaa kahdenlaista, toisistaan poikkeavaa kuvaa ilmiöistä – laajat kielimallit koostuvat generatiivisesta tekoälystä tai laajat kielimallit ovat generatiivisen tekoälyn alakäsite. Päädyin tulkitsemaan ne satelliittimallissa erillisiksi käsitteiksi, koska generatiivinen tekoäly voi tuottaa kielen lisäksi muunkinlaista sisältöä. Samalla sijoitin laajan kielimallin myös generatiivisen tekoälyn alakäsitteeksi, koska tämä on linjassa sekä yllä mainitun aineistoesimerkin että tekoälytutkimuksessa käytetyn jaottelun kanssa (ks. Chen & Chan, 2024, s. 9101, 9103).

Tekoälyavustajalla viitataan artikkeleissa esimerkiksi hyvinvointialueella käytettävään puhelinpalveluun, joka muun muassa kerää potilailta tietoja oireista ja jättää soittopyyntöjä henkilökunnalle. Aineistossa tekoälyavustaja-termillä viitataan myös Claude-sovellukseen, jota voidaan pitää generatiivisena tekoälysovelluksena, mistä syystä tekoälyavustajan käsite sijoittuu satelliittimallissa myös generatiivisen tekoälyn tyypittelyksi.

Tekoälyjärjestelmän käsitteellä viitataan aineistossa laajasti ja yleisesti erilaisiin tekoälyä hyödyntäviin järjestelmiin ja sovelluksiin. Tekoälyjärjestelmä käsitteenä ilmenee esimerkiksi konteksteissa, joissa puhutaan tekoälyjärjestelmien ja -sovellusten käytöstä opettajien työssä tai siinä, kuinka tekoälyjärjestelmien laajalla käyttöönotolla voitaisiin kasvattaa Suomen taloutta. Tekoälyjärjestelmällä ei siten viitata mihinkään tiettyihin tekoälyteknologioihin tai sovelluksiin ja sitä voitaisiinkin pitää kattoterminä laaja-alaiselle ilmiölle.

Tekoälymallilla viitataan aineistossa yleisesti erilaisiin tekoälymalleihin ilman tarkempaa selvitystä siitä, minkälaisesta teknologiasta on kysymys. Tekoälymallin käsite esiintyy esimerkiksi artikkelissa, joka käsittelee tekoälyn etiikkaa, ja artikkelissa peräänkuulutetaan eurooppalaisvalmisteisia tekoälymalleja, jotka edustaisivat eurooppalaisia arvoja.

Tekoälymalli-termiä käytetään joissain artikkeleissa myös viittaamaan esimerkiksi DeepSeek- ja ChatGPT-sovelluksiin, joita voidaan näkökulman mukaan pitää joko generatiivisina tekoälyinä tai laajoina kielimalleina, mikä heijastaa sitä, että tekoälymalli on käsitteeltään laaja.

Tekoälymallin alakäsitteeksi voidaan laskea aineistossa ilmenevä avoimen lähdekoodin tekoälymallin käsite. Avoimen lähdekoodin tekoälymallia käytetään aineiston artikkelissa samassa kontekstissa sekä kiinalaisen kielimallin, avoimen lähdekoodin kielimallien, että yleisesti eurooppalaisten tekoälymallien ja -yritysten yhteydessä, mikä tekee sen sijoittamisesta haasteellista. Osaltaan sitä voitaisiinkin pitää synonyyminä avoimen lähdekoodin kielimallille, mutta samaan aikaan termi sekä käsittelykontekstin laajuus antavat tulkinnan varaa pitää käsitettä laajempänä ilmiönä kuin vain kielimalleihin liittyvänä ilmiönä. Siksi päädyin tulkitsemaan avoimen lähdekoodin tekoälymallin kielimallia laaja-alaisemmaksi erilliseksi tekoälykäsitteeksi.

Tekoälysovelluksen käsite ilmenee aineistossa esimerkiksi hyvinvointialueen lastensuojeluilmoitukseen liittyvän artikkelin yhteydessä. Artikkelissa puhutaan tekoälypohjaisesta sovelluksesta, jonka tarkoituksena on tehdä lastensuojeluilmoitusten käsittelyssä työntekijän muistiinpanoista yhteenvetoja ja koosteita. Eli tekoälysovellus on aineistossa laaja-alainen käsite, joka ei suoranaisesti paljasta sen taustalla olevaa tarkkaa tekoälyteknologiaa, mutta joka on luokiteltavissa eräänlaiseksi tekoälyteknologian sovellukseksi. Tekoälysovellus-termiä käytetään aineistossa myös kuvailemaan erilaisia tunnetumpia tekoälyyn perustuvia ohjelmia, kuten Google Lensiä ja ChatGPT:tä, mikä korostaa käsitteen laajuutta.

Tekoälytoimijan käsitteen kohdalla oli haastavaa tulkita, onko kyseessä tekoälyteknologia tai sovellus vai tekoälyyn jollain tavalla yhteydessä oleva käyttäjä tai kehittäjä. Aineiston artikkelissa tekoälytoimijaa ei määritellä suoraan, mutta artikkelissa sillä viitataan epäsuorasti jollain autonomian tasolla uutisartikkeleita tuottaviin tai niihin liittyvään sisällön parissa toimiviin tekoälyavustajiin. Autonomiaa ei suoranaisesti tuoda esiin, ja

autonomisuus pohjautuukin tulkintaani siitä, mitä tekoälytoimijasta artikkelissa kerrotaan. Tekoälytoimijan kuvaillaan artikkelissa etsivän lähteitä itse ja valitsevan itse, miten tekoälytoimija niitä käsittelee uutissisällön yhteydessä, minkä tulkitsen itsenäiseksi toiminnaksi.

Päädyin sijoittamaan tekoälytoimijan käsitteen yhdeksi tekoälyn tyypiksi, mutta pelkään artikkelin näkökulman perusteella käsitteen voisi sijoittaa myös tekoälyavustajan alakäsitteeksi. Tulkitsen käsitteen kuitenkin satelliittimallissa artikkelin kontekstia laajempänä ilmiönä, eli tekoälytoimija voi olla muunkinlainen tekoälyn sovellus kuin tekoälyavustaja, joka tuottaa tai toimii uutissisällön parissa, miksi päädyin sijoittamaan sen omaksi tyypikseen.

Koska tekoälytoimijan käsitteessä uutisia tuottava niiden parissa toimiva tekoälymalli toimii jollain autonomian tasolla, olen yhdistänyt tekoälytoimijan käsitteen tekoälyn pääkäsitteeseen myös toimintasuhteen kautta. Kuten myöhempanä 4.5 Satelliittimallin toimintasuhteet -alaluvussa on huomattavissa, olen määritellyt tekoälytoimijan käsitteen myös yhdeksi mahdolliseksi tekoälyyn liittyväksi konetoimijaksi.

Tekoälyn kehittyneisyyden tasoon viittaaviin tyypittelyihin lasken yleisen tekoälyn ja superälyn käsitteet. Aineistossa käsitteet viittaavat tekoälyn kehittyneisyyden tasoon eikä suoranaisesti mihinkään tiettyyn tekoälyteknologiaan. Yleisellä tekoälyllä tarkoitetaan aineistossa tekoälyä, jolla on mahdollista suorittaa älykkyyttä vaativat tehtävät ihmisen tasoisesti. Superälyn käsitteellä viitataan samantapaiseen ilmiöön. Superälyllä tarkoitetaan aineistossa vielä edistyneempää tekoälyä, joka on älykkäämpi kuin ihminen ja joka kykenee kehittämään itseään tehokkaasti.

Tekoälyteknologian sovellusympäristöön viittaavaksi tyypiksi lasken fyysisen tekoälyn käsitteen. Fyysisellä tekoälyllä viitataan aineistossa sellaiseen tekoälyteknologian soveltamiseen, jonka avulla erilaiset fyysiset koneet, kuten robotit, kykenevät havaitsemaan ja toimimaan fyysisessä ympäristössä. Aineistossa fyysiselle tekoälylle luetaan

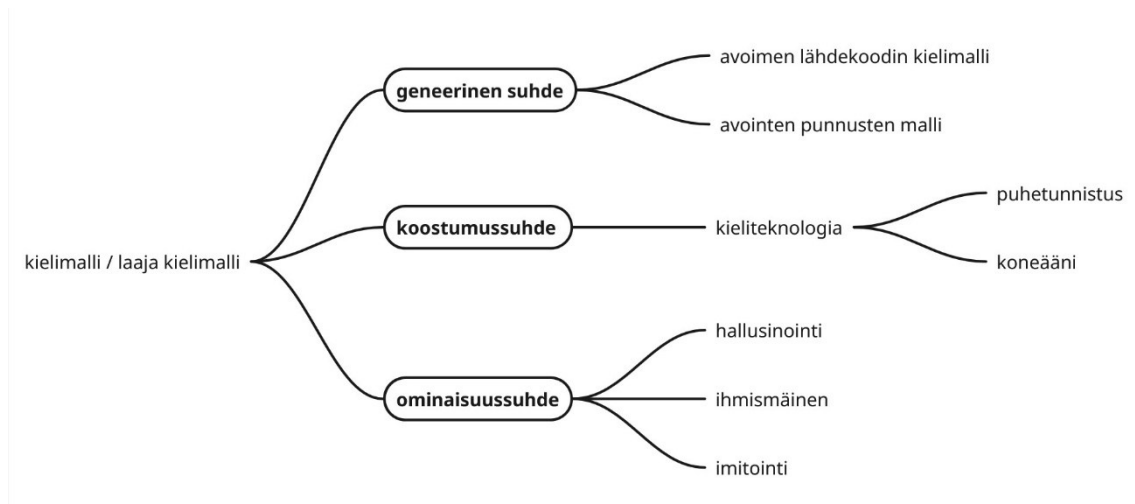
mahdollisiksi käyttöympäristöiksi tai -aloiksi esimerkiksi turvallisuus- ja puolustussektorit, mutta samalla korostetaan, että mahdolliset käyttöympäristöt eivät rajoitu pelkästään niihin.

Käsiteanalyysissä ongelmaksi erilaisten tekoälykäsitteiden tyypittelyssä muodostui se, että monet generisen suhteen eri termeistä tarkoittavat tai voivat tarkoittaa käytännössä samaa käsitettä. Aineiston kontekstissa erilaiset keskustelubotit, tekoälyavustajat, tekoälyjärjestelmät, tekoälymallit ja tekoälysovellukset voivat viitata esimerkiksi siihen teknologiaan eli laajoihin kielimalleihin, joiden avulla kyseiset ohjelmat on luotu. Päädyin tyypittelemään ne erillisiksi tekoälytyypeiksi, koska vaikka ne voivat mahdollisesti pohjautua samaan teknologiaan, tulkitsin niiden käyttötavan tai -tarkoituksen erottavaksi käsitepiirteeksi. Esimerkiksi tekoälymallin käsite on tulkittavissa laaja-alaisemmaksi ilmiöksi kuin esimerkiksi keskustelubotti, jonka funktio on esittää vastauksia ja kysymyksiä ohjelmaa käyttävän ihmiskäyttäjän kanssa, eikä niitä ole siksi mielekäästä pitää täysin synonyymisinä nimityksinä.

Lisäulottuvuuden käsitteiden järjestämisen ongelmallisuudelle tuo myös pohdinta siitä, voidaanko laajaa kielimallia ja generatiivista tekoälyä pitää erillisinä tekoälyn tyypittelyinä, vaikka generatiivisessa tekoälyssä ihmismäistä tekstiä tuottava osuus on laajaan kielimalliin pohjautuvaa. Koska generatiivinen tekoäly on yleisesti ottaen sisältöä tuottava tekoälyohjelma, ja kielimalli keskittyy ihmismäisen kielen prosessointiin, päädyin tyypittelemään molemmat erillisinä käsitteinä. Samanaikaisesti kuitenkin sijoitin laajan kielimallin generatiivisen tekoälyn alalajiksi, koska alalla sitä voidaan pitää myös generatiivisen tekoälyn alalajina (ks. esim. Chen & Chan, 2024, s. 9101, 9103). Toisaalta joissakin aineiston artikkeleissa kielimallien sanotaan pohjautuvan generatiiviseen tekoälyyn. Tämä tyypittely osaltaan reflektoiakin sitä, miten haastavaa tekoälykäsitteistä on muodostaa rakenteeltaan aukoton yhtenäinen kokonaisuus erityisesti valitun aineiston perusteella, joka on asiantuntijatekstin sijaan yleistajuista tekstiä omine reunaehtoineen. Laajemmalla ja/tai toisenlaisella aineistolla kokonaiskuvasta voisikin mahdollisesti muodostua eheämpi kokonaisuus tekoälykäsitteistä.

4.4.1 Laaja kielimalli satelliittimallissa

Kuviossa 5 esitän laajan kielimallin käsitteeseen yhteydessä olevat käsitteet ja käsitesuhteet tarkemmin. Tässä satelliittimallissa huomioitavaa on se, että kielimalli ja laaja kielimalli on sijoitettu samaksi käsitteeksi. Vaikka laajaa kielimallia voidaan pitää kehittyneempänä versiona perinteisemmästä kielimallista, aineistossa kielimallia ja laajaa kielimallia käytetään synonyymien tavoin, minkä takia merkitsin ne yhdeksi käsitteeksi, mutta laskin terminpöiminnassa erillisiksi termeiksi.



Kuvio 5. Kielimallin ja laajan kielimallin käsitesuhteet.

Laajalle kielimallille löytyi aineiston artikkeleista kaksi tyyppiä. Avoimen lähdekoodin kielimallilla määritellään aineistossa kielimalliksi, jonka rakentamiseen käytetty koodi sekä kielimallin ohjelmointiin käytetty data ovat molemmat avoimesti kaikkien saatavilla. Avointen punnusten mallilla kuvaillaan aineiston artikkelissa sen sijaan sellaista kielimallia, jonka ohjelmointiin käytetyt säännöt ja painotukset ovat avoimesti saatavilla, mutta ohjelmointiin käytetty data ei.

Koostumussuhteella laajaa kielimallia tarkastellaan siitä näkökulmasta, minkälaisia eri osia ja komponentteja ilmiö pitää sisällään. Käsiteanalyysissä laajan kielimallin käsitteen

koostumukseen sijoittui kieliteknologian käsite, jonka alakäsitteiksi tyypittelin puhetunnistuksen ja koneäänen. Kieliteknologian käsite esiintyy esimerkiksi artikkelissa, jossa käsitellään eurooppalaista laajan kielimallin kehityshanketta, ja hankkeeseen osallistuu myös suomalainen kieliteknologiaan keskittynyt tutkimusryhmä, minkä perusteella tulkitsin, että kielimalli koostuu kieliteknologiasta.

Puhetunnistus ilmenee yhdessä aineiston artikkelissa, jossa käsitellään Kajaanissa sijaitsevaa tekoälytehdasta ja suomenkielisen kielimallin kehittämistä. Tästä ilmenee puhetunnistuksen yhteys kielimalliin, mutta suoraa yhteyttä kieliteknologiaan ei aineistossa ilmene. Päädyin lopulta sijoittamaan puhetunnistuksen kieliteknologian alakäsitteeksi vapaamuotoisen taustatutkimuksen perusteella. Koneäänen käsite esiintyy artikkelissa, joka käsittelee pohjoissaamenkielisen kielimallin tuottamaa puhetta, ja kielimallin lisäksi artikkelissa esiintyy myös käsitteiden kytkös kieliteknologioiden kehitykseen, joka keskittyy uhanalaisiin kieliin. Käytin puhetunnistuksen sijoittamisen tukena myös taustatutkimusta, minkä perusteella sijoitin käsitteen kieliteknologian alakäsitteeksi.

Kieliteknologian ja sen alakäsitteiden kohdalla pohdin pitkään niiden rajausta pois, sillä ne eivät ole tiukasti tekoälyalan käsitteitä, vaan ne voidaan sijoittaa luontevammin esimerkiksi kieliteknologian pääkäsitteen alaisuuteen kuin myös kielitieteiden alle. Päädyin lopulta sisällyttämään käsitteet käsitejärjestelmään, sillä ne osoittavat, että tekoäly ei ole muusta teknologisesta kehityksestä tai tieteenaloista irrallinen osa vaan käsitesuhteiden kautta se on kytköksissä muun muassa kieliteknologiaan.

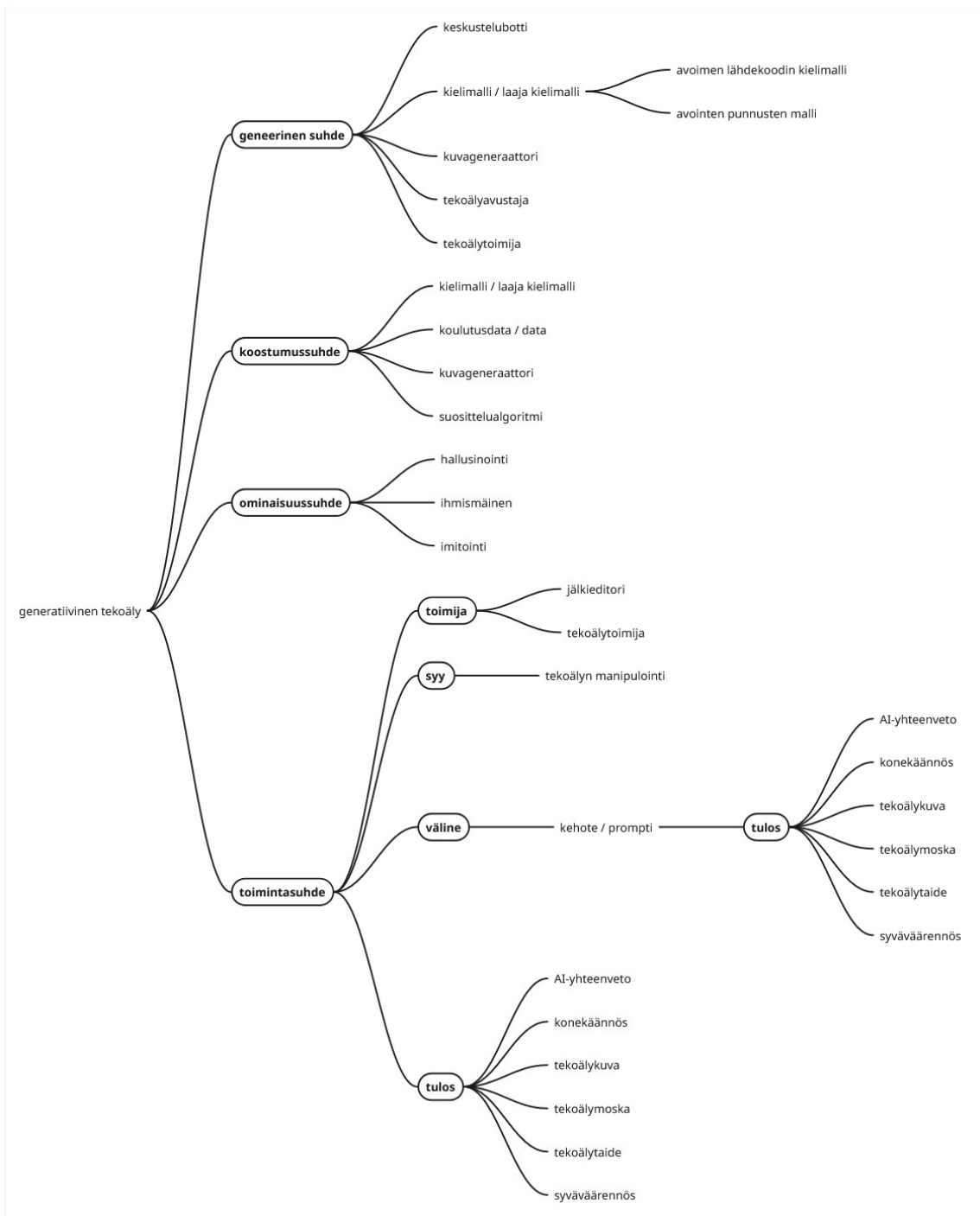
Ominaisuussuhde käsittelee sitä, minkälaisia ominaisuuksia laajalla kielimallilla ilmiönä on havaittavissa. Kielimalliin ominaisuussuhteen kautta kuuluva hallusinointi on tekoälyalalla vakiintunut käsite (ks. esim. Maleki ja muut, 2024), jolla tarkoitetaan yleisesti tekoälyn tekemiä virheitä, jotka vaikuttavat todenmukaisilta. Hallusinaation yhteys laajaan kielimalliin ilmenee aineistossa epäsuorasti. Hallusinointia käsitellään artikkelissa, jossa kerrotaan ihmisten olettavan tekoälyn olevan oikeassa, vaikka totuuden sijaan tekoäly on kehitetty arvaamaan sanoja. Sanojen arvaaminen viittaa siihen, että tekoälyllä

tarkoitetaan artikkelissa tosiasiallisesti kielimallia, mistä syystä sijoitin hallusinaation kielimallin ominaisuudeksi.

Terminpoiminnassa esiin tulleet ihmismäisyyden ja imitoinnin käsitteet eivät sen sijaan ensivaikutelman perusteella heti vaikuttaneet tekoälykäsitteiltä, eikä niitä pelkästään nimitysten perusteella sijoittaisikaan helposti tekoälyalan alaisuuteen. Kuitenkin esiintymiskontekstin tarkastelu aineistossa paljasti, että kyseessä oli kielimalliin pohjautuvan tekoälymallin kykyjä tai ominaisuuksia kuvaavia tekoälykäsitteitä. Ihmismäisyydellä voidaan aineiston tavoin viitata kielimallin tuottaman kielen ihmismäiseen ominaisuuteen, ja imitoinnilla tarkoitetaan aineistossa kielimallin kykyä matkia esimerkiksi julkisuuden henkilöiden puhe- tai kirjoitustapoja. Toisin sanoen nimitysten käyttö niiden kontekstissa muodosti käsitteiden nimityksistä tässä tutkielmassa tarkasteltavia tekoälykäsitteiden termejä, mikä on linjassa esimerkiksi Satokankaan (2021, s. 71) näkemyksen kanssa siitä, että konteksti on yksi tärkeistä termin määräytymiseen vaikuttavista tekijöistä.

4.4.2 Generatiivinen tekoäly satelliittimallissa

Nostan kuviossa 6 lähempään tarkasteluun geneeriseen käsitesuhteeseen kuuluvan generatiivisen tekoälyn sekä siihen geneeristen käsitesuhteiden, koostumussuhteiden, ominaisuussuhteiden ja toimintasuhteiden kautta liittyvät käsitteet. Geneerisen tyyppittelyn kautta generatiivisen tekoälyn alakäsitteiksi olen sijoittanut keskustelubotin, laajan kielimallin, kuvageneraattorin, tekoälyavustajan ja tekoälytoimijan käsitteet. Lisäksi laajan kielimallin käsitteeseen olen sijoittanut alakäsitteiksi avoimen lähdekoodin kielimallin ja avointen punnusten mallin.



Kuvio 6. Generatiivisen tekoälyn käsitesuhteet.

Tässä tyypittelyssä erilaiset alakäsitteet heijastavat generatiivisen tekoälyn määrittelevää käsitepiirrettä, eli sen kykyä tuottaa erilaista sisältöä tai avustaa erilaisten sisältöjen tuottamisessa. Esimerkiksi kuvageneraattori tuottaa aineiston mukaan kehoitteiden avulla kuvasisältöä, kun taas keskustelubotti ja laaja kielimalli edustavat

sisällöntuotannon kielellistä ulottuvuutta. Tekoälytoimijaan sijoittumista myös generatiivisen tekoälyn alakäsitteeksi ei voi muodostaa aineiston käsittelytavan perusteella suoraan, mutta koska käsite edustaa tässä suhteessa tietyn tyyppistä sisältöä tuottavaa tekoälyä, tässä tapauksessa uutisisältöä tuottavaa tekoälyä, tulkitsem sen yhdeksi generatiivisen tekoälyn tyyppiksi. Tästä näkökulmasta katsoen generatiivista tekoälyä voitaisiin osaltaan pitää myös kattoterminä, jonka alle lukeutuvat erilaisia sisältöjä tuottavat tekoälymallit.

Laajan kielimallin ja kuvageneraattorin käsitteet olen sijoittanut myös koostumussuhteeseen, sillä generatiivisen tekoälyn voidaan osaltaan katsoa myös koostuvan näistä komponenteista teknologisesti. Tulkitsem ilmiön koostuvan osaksi myös koulutusdatasta, koska generatiivisen tekoälyn ohjelmointi vaatii dataa, jolla ohjelma ”koulutetaan”. Tätä tulkintaa tukee myös se, että aineistossa koulutusdatan käsite esiintyy kontekstissa, jossa käsitellään tekoäly-yritysten vastuukysymyksiä ja kyseenalaistetaan sitä, onko ChatGPT:n luonut Open AI -yritys tehnyt riittäviä tarkistuksia tietoon, jolla sen tekoälymalli on koulutettu. Lisäksi aineistossa suosittelualgoritmi viittasi sosiaalimedialpalvelu X:n generatiiviseen Grok-tekoälyavustajaan, minkä perusteella tulkitsem generatiivisen tekoälyn voivan koostua myös sisältöjä käyttäjille ehdottavasta suosittelualgoritmista.

Myös generatiivisen tekoälyn ominaisuussuhteissa on havaittavissa se käsiteanalyysiin kuuluva ilmiö, että käsitejärjestelmässä samat käsitteet voivat toistua myös jonkin toisen käsitteen yhteydessä. Ominaisuussuhteessa toistuvat aiemmin laajan kielimallin yhteydessä esiintyneet hallusinoinnin, ihmismäisyyden ja imitoinnin käsitteet. Tämän toiston toteutin osoittaakseni, että jos generatiivisessa tekoälyssä hyödynnetään laajaa kielimallia, ovat sen ominaisuudet silloin myös periydyttävä generatiiviseen tekoälyyn.

Toimintasuhteella pyrin tuomaan esiin generatiiviseen tekoälyyn liittyvään toimintaan kuuluvia toimijoita, toiminnan kohteita, syitä, välineitä, paikkoja ja tuloksia. Generatiivisen tekoälyn toimintasuhteeseen kuuluvat käsitteet olen ryhmitellyt toimija-, syy-, väline- ja tulos-apunoodien avulla. Toimijoihin olen sisällyttänyt käsitteet, jotka tulkitsem

olevan generatiivisen tekoälyn käyttäjiä, eli jälkieditorin ja tekoälytoimijan käsitteet. Jälkieditorilla tarkoitetaan aineistossa uudenlaista ammattinimikettä, joka työskentelee tekoälyn tuottaman sisällön jäljen muokkaamisen parissa. Tekoälytoimijan olen tulkinut kuuluvan geneerisen suhteen lisäksi myös generatiivisen tekoälyn toimijaksi, sillä vaikka se voidaan nähdä generatiivisen tekoälyn tietynlaisena tyyppinä, autonominen tekoälytoimija käyttää samalla generatiivista tekoälyä uutissisällön tuottamiseen. Toimijoihin olisi voinut olla mahdollista sisällyttää myös generatiivisen tekoälyn kehitykseen vaikuttavat toimijat, kuten tekoäly-yrityksen tai tekoälytutkijan käsitteet. Satelliittimallin pitäminen luettavana ja hallittavan kokoisena vaikutti rajauksen ratkaisussa kuitenkin enemmän kuin kaikkien mahdollisten generatiivisen tekoälyn toimijana kytköksissä olevien käsitteiden listaaminen.

Toimintasuhteen syy-apunoodiin olen sisällyttänyt tekoälyn manipuloinnin käsitteen. Tekoälyn manipuloinnilla tarkoitetaan aineistossa ilmiötä, jossa tekoälyn algoritmeihin pyritään vaikuttamaan virheellisen informaation levittämisen mahdollistamiseksi tekoälyn tuottaman sisällön kautta. On huomioitava, että vaikka tekoälyn manipuloinnin, jälkieditorin ja tekoälytoimijan käsitteet ovat käsitejärjestelmässä samalla hierarkian tasolla ja yhteydessä generatiiviseen tekoälyn toimintasuhteen kautta, satelliittimalli ei kuvasta sitä, että kyseiset toimijat harjoittaisivat tekoälyn manipulointia. Yleisesti katsoen generatiivisen tekoälyn käytön yksi syistä voi siis olla tekoälyn algoritmien manipulointi.

Generatiivisen tekoälyn toimintasuhteessa tulokseksi olen sijoittanut tekoälyllä tuotettuja mahdollisia sisältöjä, joita ovat AI-yhteenveto, konekäännös, tekoälykuva, tekoälymoska, tekoälytaide ja syvävääreännös. Näistä käsitteistä AI-yhteenvetoa, konekäännöstä, tekoälykuvaa ja tekoälytaidetta voidaan pitää neutraaleina ilmauksina, kun taas tekoälymoska ja syvävääreännös viittaavat kielteisellä tavalla tekoälyn luoman sisällön laatuun tai eettisyyteen. Tekoälymoskalla viitataan aineistossa ilmiöön, jossa tekoälyllä on tuotettu jollain tapaa merkityksetöntä sisältöä verkossa. Tekoälymoska on aineiston kontekstissa käännös englanninkieliselle *AI slop* -termille. Syvävääreännöksellä viitataan aineistossa ilmiöön, jossa tekoälyllä tuotetussa videosisällössä tapahtuu jotain, mitä ei ole

todellisuudessa tapahtunut, kuten yhdessä artikkelissa viitataan esimerkiksi väärennettyyn videokuvaan, jossa valheellisesti esitetään vaalivilpin tapahtuneen.

Toimintasuhteen väline-suhteessa sijaitsee kehotteen käsite. Kehotetta tulkitseen välineenä, jolla käyttäjä pyrkii tuottamaan generatiivisella tekoälyohjelmalla haluamansa sisällön. Aineistossa kehote esiintyy esimerkiksi artikkelissa, jossa kerrotaan, että kehotteen avulla ChatGPT voidaan saada esimerkiksi juttelemaan vapaamuotoisesti. Tästä syystä olen yhdistänyt yllä mainitut toimintasuhteen tulokseen liittyvät käsitteet myös tulos-apunoodilla kehotteen käsitteeseen. Eli AI-yhteenveto, konekäännös, tekoälykuva, tekoälymoska, tekoälytaide ja syväväärennös ovat tuloksia, joita generatiivinen tekoäly on mahdollista saada tuottamaan kehotteita käyttämällä.

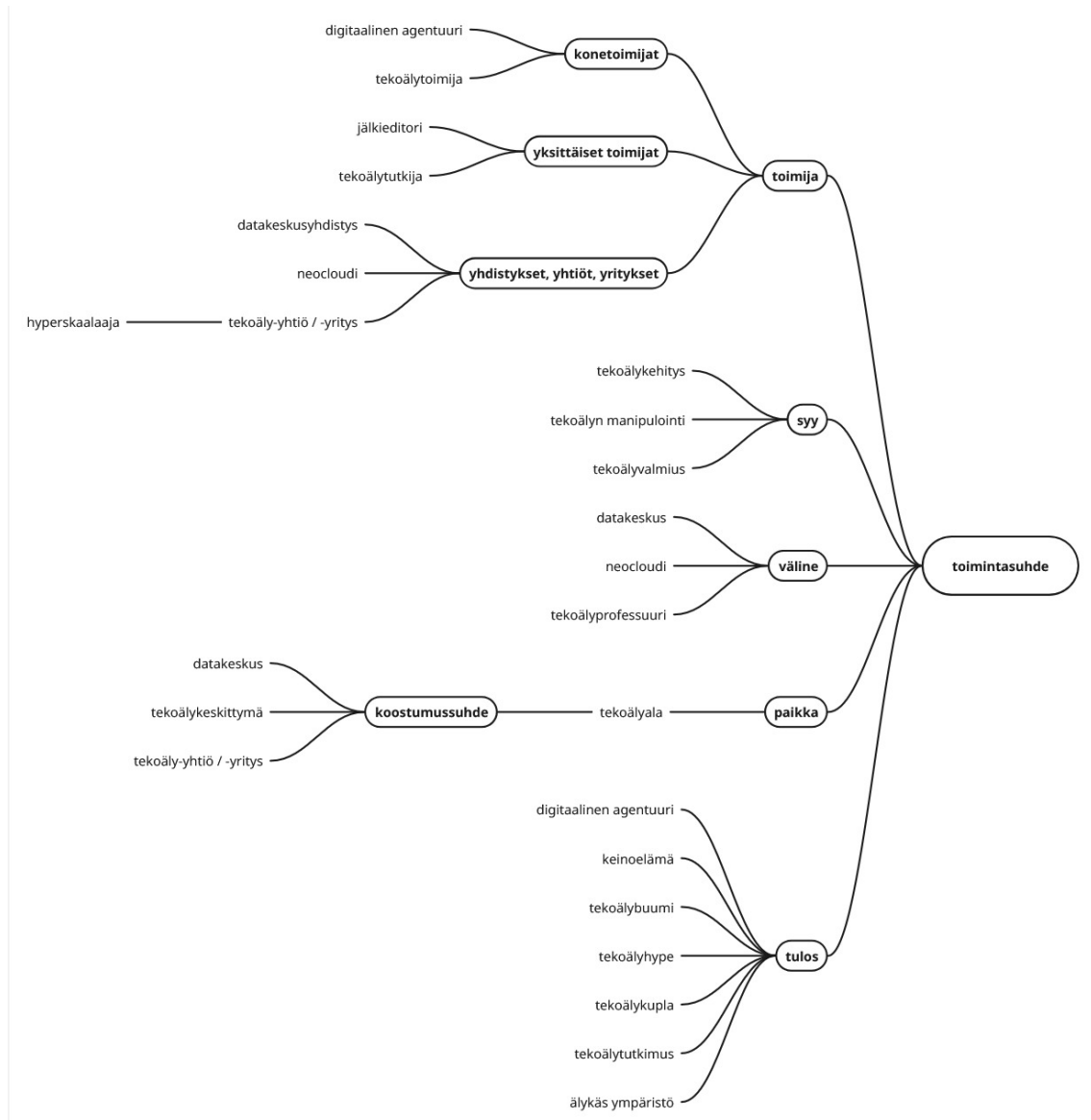
4.5 Satelliittimallin toimintasuhteet

Toimintasuhteen tarkastelun kautta toin esiin tekoälyyn toimintana liittyviä käsitteitä. Toimintasuhteella yhdistin tekoälyn pääkäsitteeseen niitä toimijoita, toiminnan kohteita, toiminnan syitä, toiminnan välineitä, toiminnan paikkoja ja toiminnasta seuraavia tuloksia, jotka aineistosta oli tulkittavissa toimintasuhteisiksi tekoälykäsitteiksi. Kuviossa 7 olen jakanut toimintasuhteen toimija-, syy-, väline-, paikka- ja tulos-apunooideihin.

4.5.1 Toimijat

Toimijasuhteeseen sijoitetut käsitteet olen jakanut kolmen apunoodin avulla omiksi ryhmikseen: konetoimijat, yksittäiset toimijat ja sellaiset toimijat, jotka ovat laskettavissa kollektiivisiksi toimijoiksi, tässä tapauksessa erilaisiksi yhdistyksiksi, yhtiöiksi tai yrityksiksi. Konetoimijaksi olen määritellyt geneerisen käsitesuhteen käsittelyn yhteydessä ilmenneen tekoälytoimijan käsitteen sekä digitaalisen agentuurin. Termin esiintymiskontekstissa digitaalisella agentuurilla tarkoitetaan aineistossa automatisoitua tekoälypohjaista palvelua, joka välittää erilaisia palveluja. Tulkitin esiintymiskontekstin perusteella,

että digitaalinen agenttuuri toimii jollain autonomian tasolla, ja sitä voidaan pitää siten itsenäisenä kone-toimijana. Sekä tekoälytoimijaa että digitaalista agenttuuria voidaan pitää eräänlaisina tekoälyagentteina (ks. esim. Ren ja muut, 2025; Sapkota ja muut, 2026).



Kuvio 7. Tekoälyn toimintasuhteet.

Yksittäisiin toimijoihin kuuluvat käsitejärjestelmässä tekoälyeditorin lisäksi tekoälytutkija. Tekoälytutkijalla viitataan aineistossa tekoälyteknologian tutkimiseen keskittyvään tutkijaan. Tulkitsen käsitettä kuitenkin laajemmin niin, että sillä voidaan yhtä lailla tarkoittaa myös tekoälyä ilmiönä tarkastelevaa tutkijaa.

Monista toimijoista koostuvaa toimijoiden kollektiivista joukkoa toimintasuhteessa edustavat yhdistykset, yhtiöt, yritykset -apunoodin alaisuuteen sijoitetut datakeskusyhdistys, neocloudi sekä tekoäly-yhtiön ja -yrityksen käsite. Yritysmuotoina yritystä ja yhtiötä ei voida pitää täysin samoina käsitteinä, mutta aineiston kontekstissa termejä käytetään synonyymisesti, mistä syystä olen sijoittanut ne samaa käsitettä edustavaan noodiin. En kokenut tarpeelliseksi erottaa eri yritysmuotoja toisistaan, koska se ei ole oleellista tutkielmani kannalta. Mikäli tutkielma käsittelisi tarkemmin esimerkiksi tekoälyyn liittyvän yritystoiminnan käsitteistöä, tekoäly-yritys ja tekoäly-yhtiö olisi siinä tapauksessa aiheellista erottaa omiksi käsitteikseen. Käsitteiden yhdistäminen liittyy myös satelliittimallin esityksen yksinkertaistamiseen. Näiden yritysmuotojen tyyppittely samaksi käsitteeksi helpotti hyperskaalaaja-käsitteen asettamista ainoastaan yhden käsitteen alalajiksi – jos yritysmuodot olisi erotettu, olisi hyperskaalaaja täytynyt asettaa molempien käsitteiden alle.

Kollektiivisista toimijoista erottuvat juuri yllä mainitun hyperskaalaajan käsitteen lisäksi neocloudin käsite, sillä käsitteiden nimityksistä itsessään on vaikea päätellä käsitteiden sisältöä. Hyperskaalaajalla tarkoitetaan aineistossa kooltaan isoja tekoäly-yrityksiä, jotka tarjoavat tekoälysovellusten pyörittämiseen isoksi skaalattuja datakeskusinfrastruktuurejaan. Neocloudilla sen sijaan tarkoitetaan aineiston kontekstissa pilvipalveluita tarjoavia yrityksiä, jotka ovat siirtyneet kryptovaluuttalouhinnasta tarjoamaan datakeskuksiaan tekoälyjärjestelmien käyttöön. Tässä käsitteen määrittelevänä käsitepiirteenä on jokseenkin poikkeuksellisesti nykyisen käyttökohteen, eli tekoälyyn tarkoitettun pilvipalvelun, sijaan käsitteen aikaisempi käyttökohde, eli kryptovaluuttalouhintaan tarkoitettu pilvipalvelu. Toisaalta samanlaista käsitepiirrelogiikkaa käytetään joskus esimerkiksi joidenkin yleiskäsitteiden kohdalla, kuten ex-puolison käsitteen kohdalla – käsite määreytyy aikaisemman suhteen perusteella.

4.5.2 Syyt

Tekoälyyn liittyvän toiminnan syinä voidaan pitää aineistosta löydettyjä tekoälykehityksen, tekoälyn manipuloinnin ja tekoälyvalmiuden käsitteitä, sillä tulkitsemme niitä tekoälyn toimintaan taustalla vaikuttavina motivaatiotekijöinä. Aineiston artikkeleissa ilmenee esimerkiksi se, että tekoälykehitystä ei toteuteta pelkästään taloudellisten syiden takia vaan motivaationa tekoälykehityksen taustalla on lisäksi isojen maiden, kuten Yhdysvaltojen ja Kiinan, tavoite toteuttaa sotateknologisia harppauksia tekoälykehityksen avulla.

Tekoälyvalmiuden käsitteellä viitataan aineistossa ilmiöön, jossa eurooppalaiset tekoälyyritykset sekä tutkimuslaitokset kehittävät yhteistyössä Euroopan teknologista kilpailukykyä tekoälyn alalla. Tässä yhteydessä tekoälyvalmiuden taustalla oleva motivoiva syy on kilpailukyvyn parantaminen. Tulkitsemme käsitteen satelliittimallissa myös yleisenä tekoälyvalmiuden käsitteenä eikä pelkästään Euroopan tekoälykyvykkyyteen liittyvänä ilmiönä. Eli esimerkiksi yritykset voivat kehittää omaa tekoälyvalmiuttaan ottamalla tekoälyteknologioita käyttöön ja kehittämällä omaa tekoälyosaamistaan.

Generatiivisen tekoälyn kohdalla esittelemäni tekoälyn manipuloinnin käsite esiintyy myös tässä yhteydessä yleisesti tekoälyyn liittyvänä mahdollisena toiminnan motivaationa. Tulkitsemme käsitteen laajemmin siten, että tekoälyä on mahdollista manipuloida muiltakin osin kuin generatiivisen tekoälyn algoritmien osalta.

4.5.3 Välineet

Tekoälyn toimintasuhteisiin liittyviin välineisiin laskin neocloudin, datakeskuksen ja tekoälyprofessorin. Neocloudi voidaan pilvipalveluja tarjoavan toimijan lisäksi laskea myös välineeksi, sillä tulkitsemme, että neocloudiyritysten tarjoamien pilvipalvelujen avulla tekoälyyritykset saavat toteutettua tekoälyratkaisujaan. Eli se on samalla väline, jolla tekoälyä toimintana suoritetaan. Samaa logiikkaa käyttäen voisi myös olla mahdollista lisätä hyperskaalaajankin käsite tekoälyn toimintaan liittyväksi välineeksi.

Tulkitsen datakeskusta tässä toimintasuhteen välinesuhteessa konkreettisena työkaluna, jolla mahdollistetaan erilaisia tekoölyyn liittyviä toimintoja. Aineistossa datakeskuksen suhdetta tekoölyyn kuvaillaan siten, että tekoöly tarvitsee datakeskuksia, joiden käyttötarkoituksiksi tarjotaan esimerkiksi tekoölyn kehittäminen. Kuten myöhempanä toimintasuhteen paikka-suhteessa esitän, datakeskus voidaan myös nähdä sijaintina, jossa tekoölyyn liittyvää toimintaa suoritetaan, mikä myös ilmenee yllä mainitussa esimerkissä, jossa datakeskuksia käytetään muun muassa tekoölyn kehittämiseen.

Sijoitin tekoölyprofessorin käsitteen tekoölyn toiminnan välineeksi, sillä tekoölyprofessori on kyseiselle työtehtävälle annettu nimike. Aineistossa tekoölyprofessoriin liittyvä toiminta keskittyy tekoölytutkimuksen tekemiseen, ja teknologian ohella muita näkökulmia tutkimukseen ovat erilaiset tekoölyn yhteiskunnalliset vaikutukset. Mikäli terminpöiminnan seurauksena termilistaan olisi päätynyt tekoölyprofessorin käsite, olisin sijoittanut sen satelliittimallissa tekoölyn toimintaan liittyviin toimijoihin. Tämä olisi korostanut tekoölyprofessorin sijoitusta väline-suhteessa – tekoölyprofessori on väline, jota käyttäen tekoölyprofessori suorittaa tekoölyyn liittyvää tutkimusta.

4.5.4 Paikat

Satelliittimalliin tekoölyn toimintapaikaksi sijoittui tekoölyalan käsite, jota voidaan pitää aineiston tavoin kattoterminä useille käsitteille, jotka liittyvät tekoölyyn toimintana. Yhdistin koostumussuhteen avulla tekoölyalaan niitä komponentteja, jotka tekoölyalan voidaan tulkita pitävän sisällään. Tällaisia käsitteitä satelliittimallissani ovat datakeskus, tekoölykeskittymä ja tekoöly-yhtiö / -yritys.

Tekoölykeskittymällä tarkoitetaan aineistossa ilmiötä, jossa samalle maantieteelliselle alueelle on sijoittunut lukuisia tekoöly-yrityksiä ja erilaista tekoölyosaamista, minkä takia käsite on luontevaa sijoittaa tekoölyalan koostumussuhteen alaisuuteen. Tekoöly-yhtiöiden ja -yritysten käsitteen olisi voinut osaltaan sijoittaa lisäksi tekoölykeskittymän

alakäsitteeksi, mutta rajasin sen pois tarpeettoman toiston välttämiseksi. Samasta syystä hyperskaalaajakaan ei esiinny tässä yhteydessä tekoäly-yhtiöiden ja -yriytysten käsitteen alakäsitteenä, vaikka se siihen luontevasti kuuluisikin.

Vaikka tulkitsem datakeskuksen käsitteen tässä mallissa tekoälytoiminnan välineeksi, sijoitin datakeskuksen myös tekoälyalan koostumussuhteeseen. Tämä johtuu siitä, että datakeskukset ovat samanaikaisesti fyysisiä rakennuksia tai paikkoja, joissa tekoälyalaan liittyvää toimintaa suoritetaan. Tästä syystä datakeskus on luontevaa sijoittaa myös sellaiseksi käsitteeksi, josta tekoälyala koostuu.

4.5.5 Tulokset

Tekoälyyn liittyvän toiminnan tuloksina tulkitsem digitaalisen agentuurin, keinoelämän, tekoälybuumin, tekoälyhypen, tekoälykuplan, tekoälytutkimuksen ja älykkään ympäristön käsitteet. Digitaalinen agentuuri ilmenee käsitejärjestelmässä konetoimijoiden lisäksi myös tekoälyn toiminnan tuloksena. Tämä johtuu siitä, että digitaalisen agentuurin toiminta autonomisena toimijana on saatu aikaan tekoälyyn liittyvän toiminnan tuloksena. Samaa logiikkaa käyttäen esimerkiksi erilaiset tekoälyn tyyppittelyt sekä itse tekoälykehityksen käsite olisi voitu sijoittaa myös tulosten yhteyteen satelliittimallissa.

Keinoelämän käsite viittaa aineiston kontekstissa ilmiöön, jossa tekoäly ja bioteknologia muodostavat yhdessä täysin uusia ihmisen luomia olioita. Aineistossa keinoelämästä esimerkkinä annetaan muun muassa älykkäät hyönteisrobotit ja erilaiset rakenteet, kuten talon seinät, jotka kasvavat tai mukautuvat ympäristöön itsestään. Tästä syystä se onkin luontevaa liittää tekoälyn avulla muodostettujen tulosten joukkoon.

Älykkään ympäristön käsite viittaa aineistossa sellaiseen ympäristöön, jossa virtuaaliset ja fyysiset maailmat toimivat sulavasti yhdessä synteettisten ympäristöjen mahdollistamiseksi. Tulkitsem kontekstin perusteella, että tekoälyllä on tässä prosessissa oma roolinsa, minkä takia sijoitin käsitteen osaksi tekoälyä koskevaa käsitejärjestelmää. Tässä

kontekstissa älykäs ympäristö on siten ainakin osittain tekoälyteknologialla aikaansaatu tulos.

Tekoälykupla, tekoälybuumi ja tekoälyhype voidaan ryhmitellä samantyyppisiksi ilmiöiksi. Tekoälykuplalla tarkoitetaan artikkelien kontekstissa ilmiötä, jossa tekoäly-yrityksiin tai niiden osakkeisiin liittyvä markkina-arvo on kohonnut suuresti investointien myötä, ja arvon epäillään tippuvan myös suuresti, jolloin sillä on merkittäviä vaikutuksia talouteen. Tekoälybuumi sen sijaan on tekoälykuplaan kytköksissä oleva ilmiö, mutta aineistossa sillä viitataan pelkän markkina-arvon kehityksen sijaan yleisemmin tekoälyn suosion kasvuun ja sen ilmiön suosion kasvun vaikutuksiin esimerkiksi rakennusalalla.

Tekoälyhype on tekoälybuumia lähellä oleva ilmiö aineiston kontekstissa. Tekoälyhypellä tarkoitetaan aineistossa tekoälyyn liittyvää innostusta. Tekoälyhypeen liittyvässä artikkelissa mainitaan muun muassa se, että generatiivinen tekoäly on se tekoälyn muoto, joka tällä hetkellä kasvattaa tekoälyinnostusta. Tekoälyhypeen voidaankin katsoa sekä vaikuttavan tekoälyyn että myös vaikuttuvan siitä – tekoälyn suosio voi kasvattaa tekoälyinnostusta ja tekoälyinnostus voi kasvattaa tekoälyn suosiota.

4.6 Satelliittimallin muut käsitesuhteet

Loput käsitesuhteista olen ryhmitellyt muiden käsitesuhteiden alle, sillä yksittäisten käsitesuhteiden sisältö on geneerisiin käsitesuhteisiin ja toimintasuhteisiin verrattuna suhteellisen vähäistä. Kuten kuviosta 8 on nähtävissä, muihin käsitesuhteisiin lukeutuvat ainesosasuhde, koostumussuhde, ominaisuussuhde, riippuvuus- ja korrelaationsuhde sekä sijaintisuhde. Näistä käsitesuhteista koostumussuhde ja ominaisuussuhde ilmenevät satelliittimallissa myös muissa käsitteiden välisissä suhteissa. Aloitan tarkastelun ainesosasuhteisiin ja koostumussuhteisiin liittyvistä käsitteistä.

4.6.1 Ainesosa- ja koostumussuhteet

Tässä luvussa käsittelem ainesosasuhteisiin ja koostumussuhteisiin kuuluvat käsitteet yhdessä niiden käsitteiden määrän vähäisyyden takia sekä sen takia, että molemmissa käsitesuhteissa toistuu osa käsitteistä. Ainesosasuhteella tarkastellaan niitä materiaaleja, joista tekoäly on valmistettu, kun taas koostumussuhteella tarkastellaan tekoälyn eri osatekijöitä ja komponentteja, joista ilmiö koostuu.



Kuvio 8. Muut käsitesuhteet.

Aineosasuhteessa ovat koulutusdata / data -käsite sekä tekoälysiiru. Ilmiönä tekoälysiiru on konkreettisempi kuin tekoälymallien ohjelmointiin käytettävä koulutusdata, joka voi koostua esimerkiksi erilaisista teksti- ja kuvasisällöistä. Tekoälysiirut ovat aineiston kontekstissa tietokonekomponentteja, joita käytetään tekoälymallien laskentatoimintoihin, ja tulkitse sen siitä syystä yhdeksi tekoälyn materiaaliksi, joka mahdollistaa ilmiön olemassaolon.

Koostumussuhteessa toistuvat koulutusdatan ja tekoälysiirun käsitteet. Lisänä koostumussuhteisiin olen sijoittanut laajan kielimallin käsitteen. Koulutusdatan ja tekoälysiirun toistuminen johtuu siitä, että tulkitse käsitteiden sopivan yhtä hyvin sekä ainesosasuhteeseen että koostumussuhteeseen; selkeää eroa siitä, ovatko koulutusdata ja tekoälysiiru tekoälyn materiaaleja vai tekoälyn komponentteja on vaikeaa tehdä. Tutkimuksen kannalta erottamista ei ole toisaalta edes mielekästä tehdä, sillä luokittelu osaltaan osoittaa ilmiön monitulkintaisen luonteen.

Yllä mainituista käsitteistä tulkitse laajan kielimallin käsitteen sen sijaan sopivan paremmin pelkästään koostumussuhteen piiriin, koska se on ilmiönä niin laaja, että sitä on vaikea nähdä yksittäisenä tekoälyn materiaalina. Useissa aineiston artikkeleissa laaja tekoälymalli on teknologiana monen tekoälysovelluksen taustalla, minkä takia tulkitse, että tekoäly ilmiönä koostuu osittain laajasta kielimallista. Mikäli terminpoiminnan tuloksena olisi termilistasta sisältynyt muita tekoälyn taustalla vaikuttavia teknologioita, kuten koneoppiminen tai syväoppiminen, olisin sijoittanut myös ne laajan kielimallin ohella tekoälyn koostumussuhteeseen.

4.6.2 Ominaisuussuhde

Ominaisuussuhteen tarkastelulla tuodaan esille tekoälyn erilaisia ominaispiirteitä. Tekoälyn ominaisuussuhteeseen olen sijoittanut ihmismäisyyden ja neuromorfisuuden käsitteet. Ihmismäisyyden käsite toistuu laajan kielimallin ominaisuussuhteen lisäksi myös tekoälyn ominaisuussuhteessa. Tekoälyn erilaisten määritelmien mukaan tekoälyllä

pyritään jäljittelemään ihmisille ominaisia toimintoja, miksi tulkiten ihmismäisyyden satelliittimallissa myös laajemmin tekoälylle ominaisena käsitepiirteenä.

Neuromorfisella tarkoitetaan aineistossa sellaisia laitteita, joiden toimintaa voidaan kuvailla ihmisaivojen tapaan toimiviaksi. Tekoälyn kontekstissa voitaisiin siis tulkita, että esimerkiksi tietokoneet tai datakeskukset, jotka tekoälyohjelmistoja pyörittävät ovat neuromorfisia, koska tekoäly on teknologiaa, jolla pyritään jäljittelemään tyypillisesti ihmisaivojen mahdollistamia funktioita, kuten tekstin ja puheen tuottamista.

4.6.3 Riippuvuus- ja korrelaationsuhde

Riippuvuus- ja korrelaationsuhteella tarkastellaan tekoälyä siitä näkökulmasta, mitkä tekijät vaikuttavat tekoälyyn ilmiönä ja mistä tekijöistä tekoäly on ilmiönä riippuvainen. Olen käyttänyt käsitesuhteen kohdalla apunoodina mistä ilmiö on riippuvainen -apunoodia, joka jaottelee käsitesuhteen niihin käsitteisiin, joista tekoäly on riippuvainen sekä mikä vaikuttaa ilmiöön -apunoodia, joka jaottelee käsitesuhteen niihin käsitteisiin, joilla voidaan nähdä olevan jonkinlainen vaikutus tekoälyyn ilmiönä.

Käsitteisiin, joista tekoäly on riippuvainen, olen laskenut datakeskuksen, koulutusdatan, tekoälykehityksen ja tekoälytutkimuksen käsitteet. Tekoälytutkimuksella tarkoitetaan aineistossa esimerkiksi yliopistoissa suoritettua tutkimusta tekoälystä. Tästä näkökulmasta katsoen tekoäly ei välttämättä ole riippuvainen tekoälytutkimuksesta. Mutta toisaalta tekoälykehitystä ja tekoälytutkimusta voidaan tulkita osittain samanlaisina käsitteinä – tekoälykehityksessä todennäköisesti tutkitaan tekoälyä osana sen kehitysprosessia. Tästä näkökulmasta katsoen olen sijoittanut tekoälytutkimuksen myös niihin käsitteisiin, joista tekoäly on riippuvainen.

Tekoälyyn vaikuttaviin ilmiöihin olen sisällyttänyt EU-komission käytännestännösten, EU:n tekoälyasetuksen ja säädöksen käsitteet, tekoälybuumin, tekoälyhyphen, tekoälykehityksen, tekoälykuplan sekä tekoälytutkimuksen käsitteet. Tulkiten, että kyseisillä

käsitteillä on omat vaikutuksensa tekoölyyn ilman, että tekoölyn olemassaolo on riippuvainen niistä. Etenkin osassa käsitteistä korostuu se, että vaikutussuhde voi olla molempuolinen. Esimerkiksi tekoölyhyphen kohdalla on aiheellista ajatella, että tekoölyhyphen ilmiönä vaikuttaa tekoölyn kehitykseen tai sen suosion kasvamiseen, mutta samaan aikaan tekoölyn ilmiö kokonaisuutena vaikuttaa tekoölyhyphen syntymiseen sekä sen voimistumiseen tai heikentymiseen.

EU-komission käytäntesäännöstö ja EU:n tekoölyasetus / tekoölysäädös ovat yhteydessä toisiinsa, mutta ne ovat silti erillisiä käsitteitä. Aineistosta ilmenee, että EU:n tekoölyasetus on laki, jolla pyritään parantamaan tekoölyn turvallista käyttöä, ja samaan käsitteeseen viitataan aineistossa myös tekoölysäädös-nimityksellä. Käytäntesäännöstöllä tarkoitetaan aineistossa tekoölyyn liittyvää ohjeistusta, jonka avulla eurooppalaiset yritykset voivat noudattaa läpinäkyvyysvaatimuksia, joita tekoölyasetuksessa on yrityksille asetettu. EU-komission käytäntesäännöstön ja EU:n tekoölyasetuksen ja -säädöksen käsitteiden kohdalla on huomioitava, että vaikka nimitykset viittaavat yksilökäsitteisiin, satelliittimallissa ilmiöillä tarkoitan yleiskäsitteitä. EU-komission ja EU:n nimien esiintymisen tarkoituksena on antaa kontekstia, minkälaiset esimerkkitoimijat näitä tekoölyyn vaikuttavia laillisia ilmiöitä ja ohjeistuksia voivat mahdollisesti muodostaa. Käsitteet tulee kuitenkin tulkita yleiskäsitteinä, eli pelkästään käytäntesäännöstönä, tekoölyasetuksena ja tekoölysäädöksenä, minkä vuoksi ne myös lasketaan tutkielmani kontekstissa yksinäisiksi termeiksi.

4.6.4 Sijaintisuhde

Sijaintisuhteessa tekoölyä tarkastellaan siitä näkökulmasta, missä tekoöly ilmiönä esiintyy. Sijaintisuhteeseen sijoitin samat käsitteet kuin toimintasuhteen paikkasuhteessa, eli tekoölyalan käsitteen, johon yhdistin koostumussuhteen kautta datakeskuksen, tekoölykeskittymän ja tekoöly-yhtiön ja -yrityksen käsitteet. Käsitteet oli mahdollista sijoittaa tulkintani mukaan molempien käsitesuhteiden alaisuuteen.

Toimintasuhteessa paikkajaottelu keskittyy niihin sijainteihin, joissa tekoölyyn liittyvää toimintaa suoritetaan, kun taas sijaintisuhteessa sijaintia ajatellaan tekoölyn näkökulmasta laajemmin. Tämä laajempi näkökulma pitää sisällään siten ne tekoölyyn liittyvät paikat, jossa toimintaa harjoitetaan, mutta myös abstraktimman näkökulman siitä, missä tekoölyn ilmiönä voidaan ajatella sijaitsevan. Sijaintisuhteeseen sopisikin esimerkiksi erilaisten työpaikkojen käsitteiden lisäksi laajemmin vaikkapa koko yhteiskunnan käsite, sillä tekoöly ilmiönä sijaitsee laaja-alaisesti muuallakin kuin vain tekoölyalan piirissä, esimerkiksi yliopistoissa, ihmisten puheissa tai verkkoartikkeleiden aiheena. Tällaisia käsitteitä satelliittimallissa ei kuitenkaan ilmene, sillä rajasin vastaavat käsitteet pois, koska ne eivät ole suoranaisesti tekoölykäsitteitä.

4.7 Yhteenveto käsiteanalyysin tuloksista

Systemaattisella käsiteanalyysillä keskityin vastaamaan ensimmäiseen tutkimuskysymykseen siitä, mitä tekoölykäsitteitä ja -termejä ilmenee Ylen tekoölyä käsittelevissä verkkoartikkeleissa. Käsiteanalyysin tuloksena toteutettu satelliittimalli koostuu niistä tekoölykäsitteistä ja niistä kuvaavista termeistä, joita aineistoon valituista Ylen verkkoartikkeleista on löydettävissä.

Käsiteanalyysin tuloksista on aineiston koon ja rajatun ajankohdan takia mahdotonta tehdä vahvoja yleistyksiä siitä, miten Yle tyypillisesti rakentaa verkkoartikkeillaan tekoölyä koskevaa käsitteistöä. Tästä huolimatta käsitejärjestelmän perusteella on kuitenkin tulkittavissa, että verkkoartikkelien sisällössä on painotusta tietynlaisten käsitteiden suuntaan. Aineistossa ilmenevät tekoölykäsitteet koostuvatkin satelliittimallin perusteella pääasiassa generisen käsitesuhteen ja toimintasuhteen kautta tekoölyn pääkäsitteeseen yhdistettävistä käsitteistä.

Iso osa käsitteistä, joita aineistoon valituissa Ylen yleistajuisesti kirjoitetuissa tekoölyaiheisissa verkkoartikkeleissa ilmenee, edustaa erilaisia yleisluonteisempia ja laaja-alaisemmaksi tulkittavia tekoölyn sovellusten tyypittelyjä. Tällaisia käsitteitä

satelliittimallissa ovat esimerkiksi keskustelubotin, tekoölyavustajan, tekoölyavustajan tai tekoölymallin käsitteet. Käsitteet eivät välttämättä suoranaisesti viittaa mihinkään tiettyyn tekoölyn teknologiaan tai käyttöalaan, vaikkakin esimerkiksi keskustelubotin kohdalla on tulkittavissa, että sen taustalla vaikuttaa esimerkiksi laajaan kielimalliin perustuva teknologia.

Tarkempana tyypittelynä voitaisiin pitää generatiivisen tekoölyn ja laajan kielimallin käsitteitä, sillä generatiivinen tekoöly ja laaja kielimalli viittaavat käsitteinä tiettyntyyppiiseen tekoölyteknologiaan laaja-alaisemman yleistyksen sijaan. Vielä tarkemmin teknologista taustaansa paljastavia käsitteitä ovat laajan kielimallin alakäsitteisiin kuuluvat avoimen lähdekoodin kielimallin ja avointen punnusten mallin käsitteet. Kyseisten käsitteiden teknologiseen taustaan liittyvät käsitepiirteet erottavat ne muusta käsitejärjestelmän tekoölytyypittelyistä yksityiskohtaisemman luonteensa vuoksi, miksi ne oli helppo sijoittaa laajan kielimallin tyypeiksi.

Generatiivisella tekoölyllä tuotetun sisällön tyypittelyssä on havaittavissa, että vaikka osa ilmauksista on tulkittavissa neutraaleiksi, kuten AI-yhteenvedon, konekäännöksen tai tekoölytaiteen käsitteet, osa tekoölysisältöihin liittyvistä ilmauksista viittaa ongelmallisuuksiin laadun tai eettisyyden näkökulmista. Esimerkkejä tällaisista ilmiön kyseenalaiseksi laadun tai eettisyyden näkökulmasta asettavista nimityksistä ovat syväväärengös ja tekoölymoska. Olen muodostanut tämän johtopäätöksen termien ulkoisesta olemuksesta verkkoartikkelien tekstisisällön sijaan. Eli vaikka tulkitsen tekoölytaiteen neutraaliksi nimitykseksi, on mahdollista, että verkkoartikkelien sisällössä tekoölytaidetta ilmiönä tarkastellessa käsitteeseen kytketään kritiikkiä esimerkiksi eettisyyden tai laadun näkökulmista. Tämän selvittämiseen tarvittaisiin kuitenkin erilaista lähestymistapaa kuin mikä tässä tutkielmassa on keskiössä.

Geneerisissä suhteissa yllä mainittuihin käsitteisiin verrattuna laaja-alaisimpia yleiskäsitteitä edustavat fyysisen tekoölyn, superälyn ja yleisen tekoölyn käsitteet. Tämän

tyyppiset käsitteet viittaavat enemmänkin tekoälyn kehittyneisyyden tasoon, kyvykkyyteen tai käyttöympäristöön kuin mihinkään tiettyyn tekoälyteknologiaan.

Geneeristen käsitesuhteiden lisäksi satelliittimallissa korostuvat toimintasuhteeseen liittyvät tekoälykäsitteet ja niiden termit. Yle rakentaa tekoälyn toimintaan liittyvää käsitteistöä erilaisilla toimijoilla, toiminnan syillä, välineillä, paikoilla ja tuloksilla. Toimijoihin tulkitsin kuuluvaksi kolme erityyppistä toimijoiden ryhmää, jotka ovat konetoimijat, yksittäiset toimijat sekä useamman toimijan joukosta muodostuvat erilaiset yhdistykset, yhtiöt ja yritykset. Toimijoihin luettavat käsitteet lähestyvät tekoälyä toimintana useista erilaisista tarpeista ja lähtökohdista. Esimerkiksi jälkieditori tarkistaa generatiivisella tekoälyllä tuotettua sisällön laatua, tekoälytutkija toimii tekoäly-yrityksessä, joka tutkii generatiivisen tekoälyn teknologiaa sen kehityksen kannalta, kun taas neocloudi on se toimija, joka tarjoaa datakeskuspalveluitaan tekoäly-yrityksien käyttöön. Tekoälyä toimintana toteutetaan erilaisten välineiden avulla, kuten datakeskuksen tai neocloudin avulla, ja vaihtelevien syiden takia, kuten tekoälykehityksen tai tekoälyvalmiuden saavuttamiseksi.

Toiminnan välineissä esiintyvä neocloudin käsite toimii esimerkkinä käsiteanalyysissä usein esiin tulleista hankaluuksista käsitesuhteiden tulkinnassa. Neocloudin käsitteen sijoittaminen sekä tekoälyn toimintaa suorittavaksi toimijaksi että toiminnan suorittamisen välineeksi osoittaa, että pelkkä näkökulman vaihtaminen voi tarkoittaa samalla sitä, että sama käsite voi kytkeytyä tekoälyalaan eri tavoilla useiden eri käsitesuhteiden kautta. Yhdestä näkökulmasta neocloudi on yritys, joka toimittaa datakeskuspilvipalveluita tekoäly-yrityksille, mutta samaan aikaan toisesta näkökulmasta se on väline, jota tekoäly-yritykset käyttävät tekoälyohjelmien mahdollistamiseen. Tämä on linjassa esimerkiksi L'Hommen (2020, s. 90) näkemyksen kanssa, jonka mukaan käsitteiden tarkastelu eri näkökulmista vaikuttaa siihen, miten käsitteitä muun muassa kategorisoidaan. Käsiteanalyysin tuloksena monet käsitteet ovatkin tästä syystä satelliittimallissa kytköksissä tekoälyn pääkäsitteeseen useamman kuin yhden käsitesuhteen kautta.

Geneerisiin suhteisiin ja toimintasuhteisiin liittyvien käsitteiden ohella satelliittimallissa nousee esiin myös muita käsitesuhteita, mutta pienemmällä painotuksella. Tällaisia käsitesuhteita ovat ainesosasuhte, koostumussuhte, ominaisuussuhte, riippuvuus- ja korrelaatio-suhte sekä sijaintisuhte. Huomioitavaa on, että useat riippuvuus- ja korrelaatio-suhteen käsitteet sekä sijaintisuhteeseen liitetyt käsitteet ovat silti tulkittavissa niin, että myös ne rakentavat tekoälykäsitteistöä toiminnan kautta. Esimerkiksi tekoälyyn ilmiönä vaikuttavan tekoälyhyphen käsitteen voidaan myös ajatella olevan yhteydessä toimintaan, koska tekoälyalalla suoritettava toiminta ja sen tulokset johtavat tekoälyhyphen syntymiseen, ja samaan aikaan tekoälyhyphen itsessään voi vastavuoroisesti vaikuttaa tekoälyyn liittyvän toiminnan suosion kasvamiseen. Myös sijaintisuhteessa sijaitsevat tekoälyalan käsitteet sekä sen koostumussuhteiset käsitteet liittyvät vielä suuremmin tekoälyyn toimintana, koska ne tuovat esiin niitä konkreettisia paikkoja, joissa tekoäly ilmenee etenkin toiminnan osalta.

Tulosten perusteella siten tulkitsen, että Ylen yleistajuiset verkkoartikkelit rakentavat tekoälyä koskevaa käsitteistöä tukeutuen pääasiassa erilaisten tekoälyteknologioiden- ja sovellusten käsitteiden tyypittelyyn sekä tekoälyyn toimintana liittyviin käsitteisiin. Tekoälyä käsitteenä ei rakenneta niinkään tarkasti tekoälyyn liittyvien teknologisten ominaisuuksien tai tekniikoiden kautta. Tekoälyä rakennetaan artikkeleissa enemmänkin yleisellä tasolla tyypittelemällä sekä viestimällä erilaisista tekoälyn mahdollistamista tuloksista, tekoälyyn ja siihen liittyvään toimintaan kuuluvista toimijoista, paikoista, välineistä sekä ilmiöistä, jotka vaikuttavat tekoälyyn ja joihin tekoäly vastavuoroisesti vaikuttaa. Tämä näkökulma tarvitsee kuitenkin tuekseen jatkossa lisää tutkimusta, jotta johtopäätöksille on mahdollista saada vankempi pohja.

Käsiteanalyysillä on myös Ylen verkkoartikkeleissa ilmenevien tekoälykäsitteiden selvittämisen ohella toinenkin funktio, joka on kytköksissä tässä tutkielmassa suoritettuun termianalyysiin. Tutkielmassani käyttämäni termin määritelmään kuuluvat tekoälyalan erikoisalan yleiskäsitteistä käytetyt nimitykset, jotka täyttävät termin määritelmän perinteisessä terminologisessa mielessä sekä tekoälyalaan kytköksissä olevien yleiskäsitteiden

nimitykset. Tämän käsitteiden kytköksen tekoälyn erikoisalaan todistan käsiteanalyysin tuloksena syntyneellä käsitejärjestelmällä, jota kuvaan satelliittimallilla. Toisin sanoen kaikki satelliittimalliini päätyneet käsitteiden nimitykset ovat potentiaalisia ehdokkaita termianalyysin kohteeksi, koska käsitteillä on selvä kytkös tekoälyn pääkäsitteeseen, mikä tekee nimityksistä tekoälytermejä tämän tutkielman kontekstissa.

5 Termianalyysi tekoälytermeistä

Tässä analyysiluvussa tarkastelen satelliittimallista tarkemman analysoinnin kohteeksi valikoituneita tekoälytermejä termianalyysin avulla. Termianalyysillä pyrin siten vastaamaan toiseen tutkimuskysymykseeni siitä, miten tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit. Kuten luvussa 2.4.2 esitin, määrittelen tutkielmassani termin erikoisalan yleiskäsitteen tai erikoisalaan liittyvän yleiskäsitteen nimitykseksi. Kaikki satelliittimallissa olevat käsitteiden nimitykset, kuten seuraavaksi esittelemäni termit, ovat tutkielmassani termejä, koska luomani satelliittimalli osoittaa käsitteiden kytköksen tekoälyn erikoisalaan. Tämän lisäksi myös puoliautomaattiseen terminpoimintaan käytämäni ohjelma tunnisti nimitykset termeiksi aineiston teksteistä.

Termit valikoituivat tarkempaan analyysiin siten, että tarkastelin valmiin käsitejärjestelmän termejä omakielisyyden ja läpinäkyvyyden näkökulmista. Termin valintaprosessia ohjasi se, vaikuttiko termin ulkoasu tulkintani mukaan muulta kuin suomen kieleltä ja/tai se, oliko itse termistä heti selkeästi ymmärrettävissä, mitä tai minkälaista tekoälykäsitettä termi nimittää. Tämän valintaprosessin perusteella valitsin termianalyysin kohteeksi seuraavat Ylen tekoälyaiheisissa verkkoartikkeleissa ilmenevät tekoälytermit: *AI-yhteenvedo*, *generatiivinen tekoäly*, *tekoälyhype*, *neocloudi* ja *hyperskaalaaja*. Tarkemman termianalyysin kohteeksi voisi olla aiheellista ottaa muitakin käsitejärjestelmän termejä, mutta valitsin termianalyysin kohteeksi yllä luetellut termit, koska tulkitsin ne omakielisyyden ja läpinäkyvyyden kannalta satelliittimallin ongelmallisimmiksi termeiksi.

Määrittelen AI-yhteenvedon tekoälyllä tuotetuksi lyhyeksi tietojen kokoelmaksi yhteenvedon kohteena olevan sisällön, kuten uutisen tai dokumentin, keskeisimmistä asioista. Generatiivisella tekoälyllä tarkoitetaan sellaista tekoälyä, joka tuottaa sisältöä (Sanastokeskus, n.d.a). Helfrich (2024, s. 703) kuvailee tekoälyhypeä tekoälyyn liittyvässä terminologisessa tutkimuksessaan yltiöpositiiviseksi huomioksi mediassa ja uutisissa, joka saa tekoälyn näyttämään erittäin innostavalta ja tärkeältä. Määrittelen tekoälyhypen kuitenkin yleisemmin innostukseksi tekoälystä, joka voi ilmetä niin mediassa kuin myös tavallisten ihmisten puheissa, ajatuksissa tai odotuksissa. Neocloudi on tekoäly-yrityksille

pilvipalveluja tarjoava yritys (ks. Raynovich, 2025) ja hyperskaalaaja on erilaisia pilvipalveluja tarjoava isokokoinen yritys (ks. Malagihal & Momaya, 2025, s. 4156).

Analysoin termianalyysissä yllä esiteltyjä tekoälytermejä niiden omakielisyyden, läpinäkyvyyden ja johdonmukaisuuden näkökulmista sekä termin rakenteen että käytön näkökulmista, joihin sisällytän termin produktiivisuuden, lyhyden, kielellisen sopivuuden sekä moitteettomuuden, tarkoituksenmukaisuuden ja erottuvuuden kriteerit. Termianalyysin tulosten yhteenvedon päätteeksi muodostin tekoälykäsitteille vaihtoehtoiset termit.

5.1 Omakielisyys

Hyvän termin tulee koostua omakielisistä eli tässä kontekstissa suomenkielisistä elementeistä. Käytännössä kaikissa termianalyysin kohteeksi valituissa termeillä on tulkittavissa vieraskielisiä sanoja tai elementtejä niiden kielellisessä ulkoasussa. *AI-yhteenveto*-termin *AI*-osa tarkoittaa tekoälyn englanninkielistä (*artificial intelligence*) lyhennettä. *Tekoälyhype*-termissä yhdyssanan perusosa on englanninkielinen *hype*, jolla tarkoitetaan esimerkiksi ilmiöön liittyvää innostusta, ja sen määriteosa koostuu omakielisestä *tekoäly*-yhdyssanasta. *Neocloudi* on yhdyssana, joka koostuu pilven kautta pilvipalveluun viittaavasta englanninkielisistä *cloudi*-perusosasta ja uuteen viittaavasta *neo*-määriteosasta. *Neocloudia* voidaan vierasperäisenä terminä pitää tyypiltään sitaattilainana, koska se kirjoitetaan ja myös mahdollisesti äännetään alkuperäisen kielen mukaisesti (ks. Nykänen, 1997) vaikkakin termin loppuun on lisätty i-kirjain suomen kielen muotoon sovittamiseksi.

Generatiivinen tekoäly on myös ongelmallinen omakielisyyden näkökulmasta. *Generatiivinen tekoäly* on englanninkielisestä *generative artificial intelligence* -termistä muodostettu sanaliitto, joka on muodoltaan niin sanotun erikoislainan ja käännöslainan välimaastossa. Erikoislaina on vieraskielestä omakieliseen kirjoitukseen ja ääntämiseen muunneltu termi, ja käännöslaina on vieraskielestä peräisin oleva termi, jonka

alkuperäinen rakenne säilyy, mutta sanat käännettään omakielisiksi (ks. Nykänen, 1997). Sanaliittotermin osista *generative* on muunnettu suomenkieliseen kirjoitukseen ja ääntämiseen sopivaksi, kun taas *artificial intelligence* on käännetty omakieliseksi. Termin *generatiivinen*-sanalle on muodostettavissa omakielisempi vastine, esimerkiksi *tuottava*. Myös *hyperskaalaaja*-yhdyssanatermiä voidaan pitää käänöslainana englanninkielisestä *hyperscaler*-termistä.

5.2 Läpinäkyvyys

Termin läpinäkyvyyden kohdalla kiinnitetään huomiota siihen, onko termistä itsestään pääteltävissä käsitteen sisältö esimerkiksi yhden tai useamman määrittelevän käsitepiirteen kautta. Vaikka argumentoin 2.6.1 Hyvän termin kriteerit -luvussa, että vieraskielinen termi ei voi olla läpinäkyvä, tässä osiossa tarkastelen termejä ottamatta kantaa termin omakielisyyteen; termien vieraskielisten osuuksien sisältö ymmärretään vieraskielisyydestä huolimatta.

Termeistä *AI-yhteenveto*, *generatiivinen tekoäly* ja *tekoälyhype* eivät ole ongelmallisia läpinäkyvyyden osalta. *AI-yhteenveto*- ja *tekoälyhype*-yhdyssanatermien kohdalla määriteosa paljastaa minkälaisesta ilmiöstä on kyse. Esimerkiksi *tekoälyhopen* kohdalla on selvää, että innostus liittyy jollain tavalla tekoälyn ilmiön ympärille. *Generatiivinen tekoäly*-sanaliitossa *generatiivinen* viittaa selkeästi siihen, että kyseessä on tekoäly, joka tuottaa jotakin. Toisaalta *generatiivinen tekoäly* voisi olla terminä vielä läpinäkyvämpi, mikäli termistä ilmenisi tarkemmin se, että termin *generatiivinen*-osa viittaa tuottamisen osalta juuri sisällöntuotantoon.

Läpinäkyvyyden kannalta ongelmallisia termejä ovat *neocloudi* ja *hyperskaalaaja*. *Neocloudit* voidaan määritellä tekoälypilvipalveluja tarjoavaksi yrityksiksi, jotka ovat tyypillisesti siirtyneet kryptovaluuttojen louhinnasta tarjoamaan graafista prosessointitehoaan tekoäly-yritysten käyttöön (ks. esim. Raynovich, 2025). Termi paljastaa osittain, että kyseessä on pilvipalveluun liittyvä käsite, mutta olennaiset käsitepiirteet, kuten

pilvipalvelujen kytkös tekoälyyn tai aiempi toiminta kryptovaluuttalouhinnan parissa eivät ilmene termistä. Toisaalta kaikissa neocloudia käsittelevissä lähteissäkään ei aina mainita aikaisempaa kytköstä kryptovaluuttoihin (ks. esim. NEXTDC, 2025). Sen lisäksi ilman tekoälyn kontekstia tarkasteltuna termistä ei välttämättä ilmene edes pilvipalvelun kytkös, sillä termistä itsestään voisi myös yhtä lailla päätellä, että se on nimitys jollekin uudentyyppiselle taivaalla sijaitsevan pilven olomuodolle.

Hyperskaalaajat voidaan määritellä isoiksi pilvipalveluja tarjoaviksi yrityksiksi, jotka tarjoavat yrityksille laajasti erilaisia pilvipalveluja, kuten prosessointia, tallennustilaa ja verkkoyhteyksiä, isossa mittakaavassa (ks. esim. Malagihal & Momaya, 2025, s. 4156). Merkittävät käsitepiirre-erot neocloudeihin keskittyvätkin siis kokoon ja käyttötarkoitukseen – neocloudit ovat pieniä ja keskittyvät tekoälyyn, kun taas hyperskaalaajat ovat isoja ja niiden palvelut ovat tarjolla tekoälyn lisäksi muun tyyppiseen käyttöön. *Hyperskaalaaja*-termissä ongelmallista läpinäkyvyyden osalta on se, että vaikka siitä ilmenee toiminnan tason iso kokoluokka, itse termistä ei selviä lainkaan, että kyseessä on pilvipalveluja tarjoava yritys.

5.3 Johdonmukaisuus

Johdonmukaisuuden tarkastelussa arvioidaan sitä, onko termistä pääteltävissä sen paikka käsitejärjestelmässä ja tietoa käsitteen ylä-, ala- tai vieruskäsitteistä. Johdonmukaisuuden näkökulmasta *AI-yhteenveto*, *tekoällyhype*, *generatiivinen tekoäly* ja *neocloudi* paljastavat elementtejä, joiden avulla niistä on pääteltävissä ainakin alustava paikka käsitejärjestelmässä.

Esimerkiksi *AI-yhteenveto* on tunnistettavissa yhdeksi yhteenvetojen alakäsitteeksi kuluista tyypeistä, ja *generatiivinen tekoäly* paljastaa, että se on tekoälyn tyyppi tietyntyyppisellä funktiolla, joka erottaa sen muista lähellä olevista vieruskäsitteistä. Tekoälyn kontekstissa tarkasteltuna *neocloudi* paljastaa terminä käsitejärjestelmästä sen, että se on

tavallisesta pilvipalvelusta jollain tavalla poikkeava, mutta muista pilvipalveluista tai pilvipalveluyrityksistä erottavaa käsitepiirrettä termistä itsessään on haastavaa tulkita.

Hyperskaalaaja on terminä ongelmallisin johdonmukaisuuden perusteella. Koska termistä ei ilmene, että kyseessä on pilvipalveluja tarjoava yritys, sitä on hankala sijoittaa pelkän termin perusteella mihinkään käsitejärjestelmään. Yhdyssanan *hyper*-määriteosa ja *skaalaaja*-perusosa paljastavat käsitteestä sen, että ilmiö on kooltaan isoa. Tämä osaltaan voisi mahdollisesti auttaa päättämään jotain tietoa ylä-, ala- tai vieruskäsitteistä, mikäli pilvipalvelu ilmiönä kävisi jollain tapaa selväksi itse termistä.

5.4 Rakenne ja käyttö

Omakielisyyden, läpinäkyvyyden ja johdonmukaisuuden lisäksi tarkastelin termejä myös niiden rakenteen ja käyttöön liittyvien hyvän termin kriteerien näkökulmista. Näihin ominaisuuksiin lasken tarkoituksenmukaisuuden, erottuvuuden, lyhyden, produktiivisuuden ja kielellisen sopivuuden sekä kielellisen moitteettomuuden ominaisuudet.

Tarkasteltavat termit eivät ole ongelmallisia tarkoituksenmukaisuuden, erottuvuuden, niiden pituuden tai produktiivisuuden osalta. Termit ovat tarkoituksenmukaisia, sillä ne eivät esimerkiksi herätä haitallisia mielikuvia, ne ovat rakenteeltaan selkeitä ja termit ovat myös suhteellisen muistettavia. Termit eivät myöskään ole helposti sekoitettavissa johonkin toiseen termiin esimerkiksi kirjoitusasun tai ääntämyksen puolesta.

Termit ovat pituuden puolesta ongelmattomia, sillä ne ovat suhteellisen lyhyitä. Kahdesta sanasta koostuva *generatiivinen tekoäly* -sanaliitto erottuu joukosta pisimpänä muiden termien ollessa yhdyssanoja, joista pisimmässä on 14 kirjainta. Toisaalta on huomioitava, että *AI-yhteenveto*-termin kohdalla on turvaututtu *artificial intelligence* -sanan lyhentämiseen. Lyhentämisen tarve saattaa kieliä siitä, että lyhentämättömässä muodossa *artificial intelligence -yhteenveto* voitaisiin mahdollisesti kokea liian pitkäksi termiksi.

Produktiivisuuden tarkastelun tuloksena en havainnut ongelmia termeissä. Termeistä on mahdollista tehdä johtamalla uusia sanoja. Esimerkiksi *neocloudista* voi muodostaa *nen-*päätteellä kuvaamaan sitä, minkälaista pilvipalveluarkkitehtuuria tekoälyn taustalla: *neocloudinen* pilvipalvelu. Myös *hyperskaalaajasta* voi muodostaa johtamalla esimerkiksi verbin: *hyperskaalata*. *Generatiivinen tekoäly* -sanaliitosta voidaan muodostaa esimerkiksi uusia yhdyssanoja välilyönnin ja yhdysmerkin avulla: *generatiivinen tekoäly -ongelma*.

Kielellisen sopivuuden ja kielellisen moitteettomuuden tarkastelu paljastaa joitain ongelmakohtia termeissä. Kielellisesti sopivaa termiä on yksiselitteistä ääntää, kirjoittaa ja taivuttaa. *Neocloudista* on vaikea tulkita, tuleeko termi ääntää kuten se kirjoitetaan, tuleeko se ääntää alkuperäiskielen mukaisesti vai riittääkö pelkästään termin *cloud*-osuuden ääntäminen englannin kielen mukaisesti. Myös *AI-yhteenvedon*, *tekoälyhypen* ja *hyperskaalaajan* kohdalla voidaan pitää epäselvänä sitä, tuleeko *AI-*, *hype-* ja *hyper-*sanat lausua suomen kielen mukaisesti vai lausutaanko ne alkuperäiskielen tavoin.

Tulkitsen kielellisen moitteettomuuden näkökulmasta ongelmallisiksi termeiksi etenkin *neocloudin*, *AI-yhteenvedon* ja *tekoälyhypen*, koska vaikka ne ovat kielelliseltä ulkoasultaan virheettömiä, tulkitsen vieraskielisten sanojen sisällymisen termeihin ongelmallisena. Tämä vieraskielisten sanojen sisällyminen ei ole perusteltua varsinkaan *AI-yhteenvedon* ja *tekoälyhypen* kohdalla, sillä *AI*:lle (*tekoäly*) ja *hypelle* (esim. *innostus*) on jo olemassa luontevat omakieliset vastineet.

5.5 Yhteenveto termianalyysin tuloksista

Termianalyysin tarkoituksena oli vastata toiseen tutkimuskysymykseen: miten tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit? Termianalyysin tuloksena totean, että tarkasteleman Ylen tekoälyaiheisissa verkkoartikkeleissa käytetyt

tekoälytermit täyttävät hyvän termin kriteerit vain osittain. Tämä asettaa mahdollisia haasteita tekoälyilmiöiden ymmärtämiselle.

Vaikka kaikkia kriteereitä on vaikea täyttää (ks. Isohella & Nuopponen, 2016, s. 229) ja termit eivät usein niitä täytäkään (ks. Nykänen, 1997), analysoidut tekoälytermit ovat ongelmallisia tämän tutkielman kontekstissa tärkeimmiksi hyvän termin kriteereiksi määriteltyjen omakielisyyden ja läpinäkyvyyden ominaisuuksien osalta. Termeistä kaikki sisältävät vieraskielisiä elementtejä, ja koska tässä tutkielmassa pidän vieraskielisiä termejä ei-läpinäkyvinä, tulkitsen kaikki tarkasteluun valitut termit ongelmallisiksi omakielisyyden lisäksi myös läpinäkyvyyden näkökulmasta. Termeistä *neocloudin* ja *hyperskaalaajan* termit ovat lisäksi epäselviä läpinäkyvyyden osalta silloinkin, jos vieraskielisyyttä ei oteta huomioon termin ymmärtämisessä, koska termeistä itsestään ei selviä käsitteen sisältö riittävällä tavalla.

Johdonmukaisuuden ongelmia on havaittavissa *hyperskaalaajan* kohdalla ja osittain myös *neocloudin* osalta. *Neocloudin* termi paljastaa tekoälykontekstissa tarkasteltuna, että se eroaa muista pilvipalveluista jollain tapaa, mutta siitä on vaikea päätellä, että yritys tarjoaa nimenomaan tekoälyyn keskittyneitä pilvipalveluja. *Hyperskaalaaja* ei termiinä paljasta riittävästi tietoa sen paikasta käsitejärjestelmässä, eikä sitä voisi pelkästään termin perusteella edes sijoittaa esimerkiksi pilvipalveluyritysten käsitteen alaisuuteen, koska ilmiön yhteys pilvipalvelun käsitteeseen ei ilmene termistä millään tavalla.

Rakenteen ja käytön, eli tarkoituksenmukaisuuden, erottuvuuden, lyhyden, produktiivisuuden ja kielellisen sopivuuden sekä kielellisen moitteettomuuden osalta termit ovat vähemmän ongelmallisia, vaikkakin joitain epäkohtia termeissä on tulkittavissa. Ongelmat tarkastelluissa tekoälytermeissä liittyvät kielelliseen sopivuuteen ja kielelliseen moitteettomuuteen vieraskielisten sanojen sisältymisen ja ääntämisen herättämien kysymysten pohjalta.

Tulkitsen termianalyysin tulosten perusteella, että termien ongelmat, jotka liittyvät erityisesti läpinäkyvyyteen ja omakielisyyteen voivat mahdollisesti vaikeuttaa tekoälyilmiöiden ymmärtämistä, kun tekoälyaiheista kirjoitetaan yleistajuisiksi tarkoitetuissa verkkoartikkeleissa. Tekoälyilmiöiden ymmärtämisen helpottamiseksi ehdotankin tilalle vaihtoehtoisia termejä, jotka voivat helpottaa ymmärtämistä etenkin omakielisyyden ja läpinäkyvyyden takia.

AI-yhteenvedolle ja *tekoälyhypelle* ehdotan käytännössä lähes suoria käännöksiä, sillä termit ovat omakielisyyttä lukuun ottamatta muilta osin osuvia kuvaamaan niiden taustalla olevia käsitteitä. *AI-yhteenvedo*-termille parempi vaihtoehto olisi siten *tekoäly-yhteenvedo* ja *tekoälyhype*-termille *tekoälyinnostus*.

Vaikka *generatiivinen tekoäly* on terminä vakiintunut käytössä, sekin voisi hyötyä termin muokkauksesta. *Generatiivinen tekoäly* olisi terminä kuvaavampi, mikäli *generatiivinen*-sanan tilalle vaihdettaisiin omakielisempi sana ja jos termistä itsestään ilmenisi tarkemmin se, mitä generatiivinen tekoäly käytännössä tuottaa. Tästä syystä ehdotan *generatiiviselle tekoälylle* vaihtoehtoiseksi termiksi *sisällöntuotantotekoälyä*, joka on sekä omakielisempi että läpinäkyvämpi. Termin pituus ei myöskään muutu merkittävästi, sillä muodostamassani vaihtoehtoisessa termissä on vain kaksi kirjainta enemmän.

Neocloudi- ja *hyperskaalaaja*-termeille ehdotan vaihtoehtoisia omakielisiä termejä, jotka tuovat selkeämmin esiin sen, kuinka molemmat käsitteet ovat pilvipalveluyrityksiä. Koska *neocloudi* on yritys, joka keskittyy tarjoamaan pilvipalvelujaan tekoäly-yrityksille, *neocloudin* sijaan sopivampi vaihtoehto termille olisi *tekoälypilviyritys*. Termi olisi mahdollista tehdä vielä läpinäkyvämmäksi, mikäli *pilvi*-sanan sijaan termissä ilmenisi pidempi *pilvipalvelu*-sana, mutta tekoälyn kontekstin ilmeneminen termissä mahdollistaa termin pitämisen lyhyempänä. On myös huomioitava, että vaihtoehtoisesta termistäkään ei ilmene tällaisten yritysten toiminnan kohteen siirtyminen kryptovaluuttojen louhinnasta tekoäly-yrityksiin. Toisaalta koska kyseinen siirtymä ei päde kaikkiin *neocloudi*-yrityksiin,

tulkitsen, että se ei ole käsitepiirteenä välttämätön, enkä siksi sisällytä käsitepiirrettä vaihtoehtoiseen termiin.

Hyperskaalaajat ovat pilvipalveluja tarjoavia suurikokoisia yrityksiä, jotka tarjoavat palvelujaan tekoälyn lisäksi lukuisiin muihin käyttötarkoituksiin. Tästä näkökulmasta katsoen *hyperskaalaaja* ei siten ole pelkästään tekoälytermi, mutta tekoälyn kontekstissa se on silti tulkittavissa tekoälytermiksi. Tästä tekoälytermiksi luokittelusta huolimatta tulkitson, että tekoälyn käsitteen sisällyttäminen jollain tapaa vaihtoehtoiseen termiin ei ole tarkoituksenmukaista, vaan määrittelevinä käsitepiirteinä tulkitson yritysten ison koon sekä sen, että yritykset tarjoavat pilvipalveluja. Tästä syystä ehdotan *hyperskaalaajalle* vaihtoehtoiseksi termiksi *suuri pilvipalveluyritys* -sanaliittotermiä. Muodostamani vaihtoehtoinen termi kuvaa osuvammin käsitesisältöä toisin kuin *hyperskaalaaja*, josta on tulkittavissa mittasuhteen iso koko, mutta ei sitä, että kyseessä on pilvipalveluja tarjoava yritys.

6 Päättäntö

Tutkimukseni tavoitteena on ollut selvittää, millaisia tekoälykäsitteitä ja -termejä ilmenee Ylen yleistajuisissa tekoälyaiheisissa verkkoartikkeleissa. Tavoitteen saavuttamisen tukena olivat kaksi tutkimuskysymystä. Ensimmäinen tutkimuskysymys käsitteli sitä, mitä tekoälykäsitteitä ja -termejä ilmenee Ylen tekoälyä käsittelevissä verkkoartikkeleissa. Toinen tutkimuskysymys selvitti sitä, miten tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit. Tutkimuksessa tarkastelin tekoälyä ilmiönä siitä näkökulmasta, kuinka Yle rakentaa yleistajuisilla verkkoartikkeleilla tekoälykäsitteistöä ja millä tavalla käsitteistä käytetyt termit edistävät tai vaikeuttavat tekoälyilmiön ymmärrettävyyttä.

Tutkimusaineistoni koostui vuonna 2025 julkaistuista Ylen tekoälyaiheisista verkkoartikkeleista, joista loin yhtenäisen tekstikorpuksen Sketch Engine -korpustyökaluohjelmalla, jolla myös suoritin korpukselle puoliautomaattisen terminpoiminnan. Ohjelma muodosti termilistat yksisanaisista termeistä ja monisanaisista termeistä. Karsin termilistoista pois muut kuin sellaiset sanat, jotka tulkitsin tekoälytermeiksi. Tutkimusmenetelminä käytin kahta terminologiseen tutkimukseen kuuluvaa analyysimenetelmää, jotka ovat systemaattinen käsiteanalyysi ja termianalyysi. Käsiteanalyysin tuloksena muodostin termilistojen perusteella aineiston tekoälykäsitteitä kuvaavan kokonaiskäsitejärjestelmän, jonka esitin satelliittimallin muodossa. Käsiteanalyysin jälkeen toteutin termianalyysin hyvän termin kriteerien perusteella viidelle tekoälytermille, jotka valitsin muodostamastani kokonaiskäsitejärjestelmästä omakielisyyden ja läpinäkyvyyden perusteella.

Ensimmäiseen tutkimuskysymykseen vastasin systemaattisella käsiteanalyysillä, jolla käsitesuhteita selvittämällä analysoin aineiston tekoälykäsitteiden suhdetta tekoälyn pääkäsitteeseen sekä muihin tekoälykäsitteisiin. Käsiteanalyysillä kokosin aineiston tekoälykäsitteistä satelliittimallin, joka edustaa aineiston tekoälykäsitteiden kokonaiskäsitejärjestelmää. Muodostamani kokonaiskäsitejärjestelmä paljasti, että Ylen tekoälyaiheisissa verkkoartikkeleissa tekoälykäsitteistöä rakennetaan pääasiassa sekä erilaisina tekoälyn sovellusten ja ilmiöiden käsitteiden tyypittelyinä että tekoälyyn toimintana liittyvien käsitteiden kautta. Tulokset paljastivat myös pienemmällä painotuksella, että

tekoälykäsitteistö rakennetaan lisäksi muunlaisten käsitesuhteiden kuten koostumus-
suhteiden, ominaisuussuhteiden ja riippuvuus- ja korrelaationsuhteiden kautta, mutta
monet muihin kuin generisiin käsitesuhteisiin tai toimintasuhteisiin liittyvistä käsitteistä
on silti yhdistettävissä tekoälyyn toimintana. Käsiteanalyysiä tehdessäni havaitsin konk-
reettisesti sen, miten kontekstilla on merkitystä siihen, mikä voidaan tulkita tekoälyter-
miksi (ks. esim. L’Homme, 2020, s. 66). Esimerkiksi käsitteet, kuten *koulutusdata*, *ihmis-
mäinen* ja *imitointi* olisi vaikeasti tulkittavissa tekoälytermeiksi, mikäli käsitteet irrotet-
tisiin pois tekoälyn asiayhteydestä.

Toiseen tutkimuskysymykseen sain vastauksen termianalyysillä. Termianalyysin tulok-
sena selvitin, että tekoälykäsitteistä käytetyt termit täyttävät hyvän termin kriteerit vain
osittain. Suurimmat ongelmat liittyivät tässä tutkielmassa tärkeimmiksi hyvän termin
ominaisuuksiksi määrittelemiini omakielisyyteen ja läpinäkyvyyteen. Tulosten perus-
teella on tulkittavissa, että tekoälytermeissä on havaittavissa muun muassa vieraskielis-
ten elementtien käyttöä, ei-läpinäkyviä ilmaisuja sekä epäselvyyksiä ääntämisen suhteen.
Tulokset auttavatkin korostamaan sitä, että tekoälykäsitteiden termejä tulisi jatkossa tar-
kastella ja käsitellä kriittisemmin hyvän termin kriteerit huomioon ottaen. Etenkin silloin,
kun tarkoituksena on tuottaa yleistajuista uutisointia monisyisestä ja laajasta ilmiöstä,
joka vaikuttaa kansalaisten elämään erikoisalan ulkopuolellakin. Puutteet hyvän termin
ominaisuuksissa voivat osaltaan mahdollisesti vaikeuttaa tekoälyilmiöiden ymmärrettä-
vyyttä, miksi tekoälystä viestiessä tulisikin kiinnittää huomiota niihin nimityksiin, joita
tekoälyilmiöistä käytetään.

Termianalyysin tulosten yhteenvedon lopuksi muodostin tekoälytermeille vaihtoehtoiset
termit, jotka täyttävät hyvän termin kriteerit paremmin etenkin omakielisyyden ja lä-
pinäkyvyyden osalta. On kuitenkin tiedostettava, että terminologisen tutkimuksen teki-
jöllä on erikoisalan asiantuntijoihin verrattuna huomattavasti pienempi mahdollisuus
vaikuttaa erikoisalan käytäntöihin alalla käytettävien termien suhteen (ks. esim. Nuop-
ponen, 2003, s. 22). Schmitz (2009, s. 4) arvioi termien muuttamista jyrkemmin ja argu-
mentoi, että edes huonoja vakiintuneita termejä ei tulisi muuttaa ilman pakottavaa

tarvetta, kuten kulttuurillisesti herkissä tapauksissa tai jos erikoisalan sisällä samaa termiä käytetään kuvailemaan useampaa kuin yhtä eri käsitettä.

Pidän kuitenkin tekoälytermien muokkaamista aiheellisena etenkin tekoälyn yhteiskunnallisen merkittävyyden takia, koska termien ymmärtäminen ei koske pelkästään erikoisalan asiantuntijoita vaan myös maallikkoja, joita tekoäly ja sen vaikutukset koskettavat. On epätodennäköistä, että muodostamiani tekoälytermivaihtoehtoja otetaan laajempaan käyttöön alkuperäisten termien korvaajiksi, mutta tutkimuksen tuloksia voisi siitä huolimatta hyödyntää eräänlaisena ohjenuorana esimerkiksi journalismissa tekoälykäsitteistä uutisoitaessa. Tulosten pohjalta voisi olla mahdollista ottaa mallia niistä periaatteista, jotka auttavat jatkossa muodostamaan merkityssisällöltään selkeämpiä tekoälytermejä. Tulokset auttavat myös korostamaan sitä, että yleistajuisessa viestinnässä käytetyt tekoälytermit voivat olla ongelmallisia ymmärrettävyyden kannalta.

Tutkimukseeni liittyy rajoituksia ja varauksia. Rajoituksena tutkimukselle on huomioitava, että yksi aineistoon valikoitumisen kriteereistä oli se, että verkkoartikkelissa on käytetty tekoäly-aihetunnistetta. On siis mahdollista, että aineiston ulkopuolelle on jäänyt vuonna 2025 julkaistuja Ylen tekoälyaiheisia verkkoartikkeleita, joita ei ole jostain syystä merkitty tekoäly-aihetunnisteella, vaikka niiden kohdalla olisi ollut perusteltua käyttää kyseistä tunnistetta. Sen lisäksi rajoituksena voidaan pitää myös aineistovalintaa ja sen rajausta. Ottamalla huomioon laajemmin suomalaisen mediakentän, tuloksista olisi voinut mahdollisesti muodostaa kattavamman yleiskuvan siitä, kuinka yleistajuisessa mediassa rakennetaan tekoälykäsitteistöä ja millaisin termein.

Tutkimuksen tuloksiin on suhtauduttava varauksella, sillä käsitteisiin ja termeihin liittyy subjektiivisia tekijöitä. Käsitteen kuvailuun (Sager, 1990, s. 16–17), ymmärtämiseen (Cabré, 1999, s. 43) sekä rajaukseen, kategorisointiin ja määrittelyyn (L'Homme, 2020, s. 90) liittyy aina subjektiivisia ja suhteellisia tekijöitä, joilla on vaikutusta käsiteanalyysiin. Toisin sanoen toisen tutkijan tekemänä tässä tutkielmassa suoritettuna käsiteanalyysin tulos voisi mahdollisesti olla erilainen muun muassa erilaisen näkemyksen, kategorisoinnin,

määrittelyn tai näkökulman takia. Termi on myös suhteellinen käsite (ks. esim. L’Homme, 2020, s. 66) ja termien tunnistaminen on subjektiivista (ks. esim. Pasanen, 2009, s. 45), koska termin tunnistamiseen vaikuttaa erilaisten termin määritelmien ohella muun muassa konteksti ja termin keräämisen käyttötarkoitus (ks. L’Homme, 2020, s. 66). Toisin sanoen toinen tutkija saattaisi kerätä samasta aineistosta erilaisen termilistan, millä olisi vaikutusta sekä systemaattisen käsiteanalyysin että termianalyysin tuloksiin.

Aineistolle suorittamani puoliautomaattinen terminpoimintamenetelmä asettaa myös omat varauksensa tutkimuksen tuloksille. Puoliautomaattinen terminpoiminta sisällyttää termilistoihin paljon tutkimukseen kelpaamattomia termejä, mutta samalla se voi jättää poimimatta termejä, jotka tutkija itse poimisi termilistaansa (ks. esim. Pasanen, 2009, s. 49–50). Puoliautomaattisen terminpoiminnan muodostamien termilistojen siivoamisen jälkeen tarkistin satunnaisesti valitsemistani aineiston artikkeleista, mikäli terminpoimintaohjelmana käyttämäni Sketch Engine jätti poimimatta artikkeleista sellaisia tekoälytermejä, jotka olisin itse manuaalisella terminpoiminnalla hyväksynyt termilistaani. Manuaalisella terminpoiminnalla olisin hyväksynyt termilistaani muun muassa seuraavat aineistosta löytämäni tekoälytermiehdokkaat: *AI-sovellus*, *kehoteinjektio*, *neuroverkkokaarkitehtuuri*, *piiloprompti*, *promptaustaidot*, *superpodit*, *superprompti*, *super-tekoäly*, *tekoälykiihdytin*, *tekoälyn tutkimuskeskus* ja *tekoälyverkko*. Monia edellä mainituista termeistä olisi aiheellista tutkia termianalyysillä etenkin omakielisyyden ja läpinäkyvyyden näkökulmista.

Toisin sanoen pelkästään puoliautomaattista terminpoimintaa käyttämällä ei ollut mahdollista sisällyttää termilistoihin, ja siten muodostamaani käsitejärjestelmään, kaikkia aineistossa ilmeneviä tekoälykäsitteitä. Käsiteanalyysin tuloksena syntynyt kokonaiskuva käsitejärjestelmästä kattaakin tästä syystä ainoastaan osan kaikista mahdollisista aineistossa olevista tekoälykäsitteistä. Muodostamani kokonaiskäsitejärjestelmän perusteella onkin hankalaa muodostaa pitäviä yleistyksiä siitä, millä tavalla Yle rakentaa tekoälykäsitteistöä verkkoartikkeleissaan, mutta se tarjoaa kuitenkin alustavaa suuntaviivaa, jota on mahdollista vahvistaa jatkotutkimuksella.

Jatkotutkimusta ajatellen voisi siten olla tarpeellista tutkia samaa aineistoa samoin metodein yhdistämällä terminpoimintaan myös manuaalisen terminpoiminnan, jolloin aineistosta olisi mahdollista saada analyysiin kattavammin mukaan Ylen verkkoartikkelien tekoälykäsitteitä ja -termejä. Tutkielmani pohjalta jatkotutkimuskohteena voisi myös olla mielekästä tarkastella esimerkiksi vertailevalla otteella sitä, kuinka tekoälykäsitteillä muodostetaan kuvaa tekoälystä ilmiönä eri medialähteissä. Tutkielmassani keskityn yhteen medialähteeseen, jolloin tuloksista ei ole mahdollista tehdä yleistettävää kuvaa koko suomalaisen uutismediakentän tekoälykäsitteiden käytöstä ja niiden termien laadusta. Lisäksi keskityn ajallisesti vuonna 2025 julkaistuihin verkkoartikkeleihin, jolloin aineistovalinta ei julkaisuajankohdan rajauksen puolesta mahdollista ajallista vertailua Ylen verkkoartikkelien kesken. Tämä tarjoaakin jatkotutkimukselle aiheen tarkastella esimerkiksi tekoälykäsitteiden ja termien muutosta jollain tietyllä aikavälillä Ylen tai muun vastaavan medialähteen verkkoartikkeleissa.

Systemaattisen käsiteanalyysin ja termianalyysin ohella tässä tutkielmassa käytettyä tai vastaavaa aineistoa voisi olla tarpeen tutkia myös diskurssianalyysin keinoin siitä näkökulmasta, millä tavalla suomalaiset mediat rakentavat tekoälyä ilmiönä. Lisätutkimusta aiheesta olisi mahdollista tehdä myös tutkimalla sitä, kuinka media kehystää tekoälyä. Tällä voitaisiin selvittää esimerkiksi sitä, kenen äänet ja mitkä näkökulmat pääsevät esiin suomalaisen median tekoälyuutisoinnissa, ja samalla paljastaa kenen äänet ja minkälaiset näkökulmat jäävät mediakeskustelussa lukijoiden ulottumattomiin. Tekoäly on yhteiskunnallisena aiheena merkittävä ja tekoälyä on juuri sen takia merkityksellistä tutkia useista eri viestinnällisistä näkökulmista.

Lähteet

- Ahmad, K. (2001). The Role of Specialist Terminology Management in Artificial Intelligence and Knowledge Acquisition. Teoksessa S. Wright & G. Budin (toim.), *Handbook of Terminology Management* (s. 809–844). John Benjamins Publishing Company.
- Alshtaiwi, M. (2025). Analysis of Artificial Intelligence Terminology: Complexity, Evolution, and Interdisciplinarity. *XLinguae*, *18(1)*, 191–198. <https://doi.org/10.18355/XL.2025.18.01.15>
- Banh, L. & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, *33(1)*, 63, 1–17. <https://doi.org/10.1007/s12525-023-00680-1>
- Bunz, M. & Braghieri, M. (2022). The AI doctor will see you now: Assessing the framing of AI in news coverage. *AI & Society*, *37(1)*, 9–22. <https://doi.org/10.1007/s00146-021-01145-9>
- Cabré, M. (1999). *Terminology: Theory, Methods, and Applications*. John Benjamins Publishing Company.
- Cao, B., Greiner, R., Greenshaw, A. & Sui, J. (2025). AI and Machine Learning Terminology in Medicine, Psychology, and Social Sciences: Tutorial and Practical Recommendations. *Journal of Medical Internet Research*, *27(5)*, e66100, 1–13. <https://doi.org/10.2196/66100>
- Chen, Z. & Chan, J. (2024). Large Language Model in Creative Work: The Role of Collaboration Modality and User Expertise. *Management Science*, *70(12)*, 9101–9117. <https://doi.org/10.1287/mnsc.2023.03014>
- Chen, Y., Zhai, Y. & Sun, S. (2024). The gendered lens of AI: Examining news imagery across digital spaces. *Journal of Computer-Mediated Communication*, *29(1)*, 1–11. <https://doi.org/10.1093/jcmc/zmad047>
- Dandurand, G., McKelvey, F. & Roberge, J. (2023). Freezing out: Legacy media's shaping of AI as a cold controversy. *Big Data & Society*, *10(2)*, 1–13. <https://doi.org/10.1177/20539517231219242>
- Deroy, O. (2023). The Ethics of Terminology: Can We Use Human Terms to Describe AI? *Topoi*, *42(3)*, 881–889. <https://doi.org/10.1007/s11245-023-09934-1>

- Duberry, J. & Hamidi, S. (2021). Contrasted media frames of AI during the COVID-19 pandemic: A content analysis of US and European newspapers. *Online Information Review*, 45(4), 758–776. <https://doi.org/10.1108/OIR-09-2020-0393>
- Euroopan parlamentti. (2023). *Mitä tekoäly on ja mihin sitä käytetään?* Noudettu 5.11.2025 osoitteesta <https://www.europarl.europa.eu/topics/fi/article/20200827STO85804/mita-tekoaly-on-ja-mihin-sita-kaytetaan>
- Euroopan unioni. (2024). *Euroopan parlamentin ja neuvoston asetus (EU) 2024/1689*. Noudettu 5.11.2025 osoitteesta https://eur-lex.europa.eu/legal-content/FI/TXT/HTML/?uri=OJ:L_202401689
- Frank, D.-A., Jacobsen, L., Søndergaard, H. & Otterbring, T. (2023). In companies we trust: consumer adoption of artificial intelligence services and the role of trust in companies and AI autonomy. *Information Technology & People*, 36(8), 155–173. <https://doi.org/10.1108/ITP-09-2022-0721>
- Gotti, M. (2011). *Investigating Specialized Discourse*. Peter Lang.
- Graziani, M., Dutkiewicz, L., Calvaresi, D., Amorim, J., Yordanova, K., Vered, M., Nair, R., Abreu, P., Blanke T., Pulignano, V., Prior, J., Lauwaert, L., Reijers, W., Depeursinge, A., Andrearczyk, V. & Müller, H. (2023). A global taxonomy of interpretable AI: Unifying the terminology for the technical and social sciences. *The Artificial Intelligence Review*, 56(4), 3473–3504. <https://doi.org/10.1007/s10462-022-10256-8>
- Guzman, A. & Lewis, S. (2020). Artificial intelligence and communication: A Human-Machine Communication research agenda. *New Media & Society*, 22(1), 70–86. <https://doi.org/10.1177/1461444819858691>
- Haenlein, M. & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Haleem, A., Javaid, M., Asim Qadri, M., Pratap Singh, R. & Suman, R. (2022). Artificial intelligence (AI) applications for marketing: A literature-based study. *International Journal of Intelligent Networks*, 3, 119–132. <https://doi.org/10.1016/j.ijin.2022.08.005>

- Haskell, C. & Clark, S. J. (2025). Leadership in AI Terminology Governance: From Anomia to Agency. *Journal of Leadership Studies*, 18(4), 55–66. <https://doi.org/10.1002/jls.70002>
- Hauptman, A., Schelble, B., Duan, W., Flathmann, C. & McNeese, N. (2024). Understanding the influence of AI autonomy on AI explainability levels in human-AI teams using a mixed methods approach. *Cognition, Technology & Work*, 26(3), 435–455. <https://doi.org/10.1007/s10111-024-00765-7>
- Helfrich, G. (2024). The harms of terminology: why we should reject so-called “frontier AI”. *AI and Ethics*, 4(3), 699–705. <https://doi.org/10.1007/s43681-024-00438-1>
- Hirvonen, A. (2018). *Yle vastaa: Yle edistää elävää kieltä, vastuullisesti*. Yle.fi. Noudettu 2.11.2025 osoitteesta <https://yle.fi/aihe/a/20-278405>
- Hollanek, T., Peters, D., Drage, E. & Hernandez, R. (2025). AI, journalism, and critical AI literacy: Exploring journalists’ perspectives on AI and responsible reporting. *AI & Society*, 40(8), 6393–6405. <https://doi.org/10.1007/s00146-025-02407-6>
- Hongisto, P. (2023). *Tekstitaidot: Kielenhuolto. Oikeinkirjoitus: yhdyssanat*. Kielibuusti. Noudettu 23.3.2026 osoitteesta <https://www.kielibuusti.fi/sites/default/files/2023-07/Yhdyssanat.pdf>
- Huang, M.-H. & Rust, R. T. (2021). Engaged to a Robot? The Role of AI in Service. *Journal of Service Research*, 24(1), 30–41. <https://doi.org/10.1177/1094670520902266>
- Huang, M.-H. & Rust, R. T. (2024). The Caring Machine: Feeling AI for Customer Care. *Journal of Marketing*, 88(5), 1–23. <https://doi.org/10.1177/00222429231224748>
- Isohella, S. & Nuopponen, A. (2016). Terminologia kohtaa käytettävyyden. Terminologisen käytettävyyden ydintä rakentamassa. Teoksessa P. Hirvonen, D. Rellstab & N. Siponkoski (toim.), *Teksti ja tekstuaalisuus*. VAKKI-symposiumi XXXVI (s. 226–237). VAKKI. Noudettu 2.10.2025 osoitteesta https://www.vakki.net/publications/2016/VAKKI2016_Isohella&Nuopponen.pdf
- Kauppinen, L. (2025). Puhuvat ”tekoälystä” eivätkä ATK:sta. *Aamulehti*. Noudettu 10.11.2025 <https://www.aamulehti.fi/kolumnit/art-2000011605502.html>

- Kozak, M. (2026). From AI peer review to AI ethics: addressing terminological confusion in science. *Journal of Documentation*, 82(1), 111–119. <https://doi.org/10.1108/JD-07-2025-0201>
- Laatu, E. (2025). *Generatiivisen tekoälyn hallusinaation käsitteen ja vaihtoehtoisten termien tarkastelu luonnollisen kielen tuottamisen viitekehyksestä* [pro gradu -tutkielma, Jyväskylän yliopisto]. JYX. <https://urn.fi/URN:NBN:fi:ju-202501141184>
- Lewandowska-Tomaszczyk, B. & Pawlowski, G. (2025). Testing ChatGPT on Terminology Generation, Definitions: Translation, and Ontology Creation in German, English and Polish. *Research in Language*, 23, 320–340. <https://doi.org/10.18778/1731-7533.23.20>
- L'Homme, M-C. (2020). *Lexical semantics for terminology: An introduction*. Amsterdam: John Benjamins.
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I. & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790, 1–13. <https://doi.org/10.1016/j.ijme.2023.100790>
- Malagihal, S. S. & Momaya, K. S. (2025). Thinking beyond core activities: Evidence from telecom service providers competing in the digital platform economy. *Technology Analysis & Strategic Management*, 37(13), 4154–4167. <https://doi.org/10.1080/09537325.2024.2441806>
- Maleki, N., Padmanabhan, B. & Dutta, K. (2024). AI Hallucinations: A Misnomer Worth Clarifying. *IEEE*, 133–138. <https://doi.org/10.1109/CAI59869.2024.00033>
- Nedobity, W. (1985). Terminology and Artificial Intelligence. *Knowledge Organization*, 12(1), 17–19. <https://doi.org/10.5771/0943-7444-1985-1-17>
- NEXTDC. (2025). *What is a Neocloud?* Noudettu 20.4.2026 osoitteesta <https://www.nextdc.com/blog/what-is-a-neo-cloud>
- Nguyen, D. (2023). How news media frame data risks in their coverage of big data and AI. *Internet Policy Review*, 12(2), 1–30. <https://doi.org/10.14763/2023.2.1708>

- Nguyen, D. & Hekman, E. (2024). The news framing of artificial intelligence: A critical exploration of how media discourses make sense of automation. *AI & Society*, 39(2), 437–451. <https://doi.org/10.1007/s00146-022-01511-1>
- Nuopponen, A. (1999). Mihin terminologian teoriaa ja menetelmiä voidaan hyödyntää? Teoksessa K. Kuhmonen & Tekniikan sanastokeskus (toim.), *Toimikunnista termi-talkoisiin: 25 vuotta sanastotyön asiantuntemusta* (s. 91–98). Tekniikan sanastokeskus.
- Nuopponen, A. (2003). Käsiteanalyysi asiantuntijan työvälineenä. Teoksessa M. Koskela & N. Pilke (toim.), *Kieli ja asiantuntijuus. AFinLAN vuosikirja, 61*, (s. 13–24). Suomen soveltavan kielitieteen yhdistys.
- Nuopponen, A. (2009). Käsiteanalyysia käsiteanalyysistä – kohti systemaattista käsiteanalyysia. Teoksessa M. Enell-Nilsson & N. Nissilä (toim.), *Kieli ja valta. VAKKI-symposiumi XXIX XXIV* (s. 308–319). Vaasan yliopisto.
- Nuopponen, A. (2011). Methods of concept analysis – tools for systematic concept analysis. *LSP Journal*, 2(1), (s. 4–15). <https://urn.fi/URN:NBN:fi-fe202103197869>
- Nuopponen, A. (2016). Satellite system as a visualization tool for concept analysis. Teoksessa H. E. Thomsen, A. Pareja-Lora & B. N. Madsen (toim.), *Term Bases and Linguistic Linked Open Data: Proceedings of TKE 2016 the 12th International conference on Terminology and Knowledge Engineering* (s. 190–200). Copenhagen Business School.
- Nuopponen, A. (2020). Systemaattinen käsiteanalyysi tutkijan työssä. Teoksessa H. Katajamäki (toim.), *Tieteellinen kirjoittaminen tiedeyhteisössä* (s. 94–122). VAKKI ry. <http://www.urn.fi/URN:NBN:fi-fe2020101383810>
- Nuopponen, A. (2022). Conceptual relations: From the General Theory of Terminology to knowledge bases. Teoksessa P. Faber & M-C. L'Homme (toim.), *Theoretical Perspectives on Terminology: Explaining terms, concepts and specialized knowledge. Terminology and Lexicography Research and Practice*, 23, 63–68. <https://urn.fi/URN:NBN:fi-fe2023022128129>
- Nuopponen, A. & Pilke, N. (2016). *Ordning och reda: Terminologilära i teori och praktik*. <https://urn.fi/URN:NBN:fi-fe2024061250675>

- Nykänen, O. (1997). Mitä TSK sanoo hyvistä ja huonoista termeistä — ja milloin? *Terminfo* 1/97. Noudettu 22.9.2025 osoitteesta https://sanastokeskus.fi/tsk/fi/mit%C3%A4_tsk_sanoo_hyvist%C3%A4_ja_huonoista_termeist%C3%A4_%E2%80%94_ja_milloin-409.html
- Pasanen, P. (2009). *Merenkulun turvallisuuden koetinkiviä: Terminologisen tiedon poiminta tekstistä* [väitöskirja, Helsingin yliopisto]. Helda. <http://urn.fi/URN:ISBN:978-952-10-5673-4>
- Pasanen, P. (2015). Tavallisten kielenkäyttäjien käsityksiä termeistä. *VAKKI Publications*, 4(1), 178–187. <https://doi.org/10.70484/vakki.149178>
- Plechko, A., Chukhno, T., Nikolaieva, T., Apolonova, L. & Leleka, T. (2022). Anglicisms in the Field of Information Technology: Analysis of Linguistic Features. *International Journal of Computer Science and Network Security*, 22(4), 183–192. <https://doi.org/10.22937/IJCSNS.2022.22.4.23>
- Raevaara, T. (2016). *Tajuaako kukaan?: Opas tieteen yleistajuistajalle*. Vastapaino.
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Raynovich, R. S. (2025). Inside The Neocloud Economy: What’s Next For GPU-As-A-Service. *Forbes*. Noudettu 20.4.2026 osoitteesta <https://www.forbes.com/sites/rscottraynovich/2025/11/06/inside-the-neocloud-economy-whats-next-for-gpu-as-a-service/>
- Rehak, R. (2023). *The Language Labyrinth: Constructive Critique on the Terminology Used in the AI Discourse*. arXiv.org. <https://doi.org/10.48550/arxiv.2307.10292>
- Ren, Y., Liu, Y., Ji, T. & Xu, X. (2025). AI Agents and Agentic AI—navigating a plethora of concepts for future manufacturing. *Journal of Manufacturing Systems*, 83, 126–133. <https://doi.org/10.1016/j.jmsy.2025.08.017>
- Roe, J. & Perkins, M. (2023). ‘What they’re not telling you about ChatGPT’: Exploring the discourse of AI in UK news media headlines. *Humanities & Social Sciences Communications*, 10(1), 753, 1–9. <https://doi.org/10.1057/s41599-023-02282-w>

- Sager, J. C. (1990). *A Practical Course in Terminology Processing*. John Benjamins Publishing Company.
- Sanastokeskus. (n.d.a). *Generatiivinen tekoäly*. TEPA-termipankki. Noudettu 13.5.2026 osoitteesta <https://termipankki.fi/tepa/fi/haku/generatiivinen%20teko%C3%A4ly>
- Sanastokeskus. (n.d.b). *Tekoäly*. TEPA-termipankki. Noudettu 30.10.2025 osoitteesta <https://termipankki.fi/tepa/fi/haku/teko%C3%A4ly>
- Sanastokeskus TSK. (2006). *Terminologian sanasto, TSK 36*. Sanastokeskus.fi. Nordterm, Helsinki. Noudettu 15.2.2026 osoitteesta <https://sanastokeskus.fi/tiedostot/pdf/TerminologianSanasto.pdf>
- Sanastokeskus. (2020). *Käsitä käsitteet, tartu termiin – perehdytys sanastotyön perusteisiin*. Sanastokeskus ry. Noudettu 20.2.2026 osoitteesta https://sanastokeskus.fi/tiedostot/kurssit/perehdytys_sanastotyohon.pdf
- Sapkota, R., Roumeliotis, K. & Karkee, M. (2026). AI Agents vs. Agentic AI: A Conceptual taxonomy, applications and challenges. *Information Fusion*, 126, 103599, 1–30. <https://doi.org/10.1016/j.inffus.2025.103599>
- Satokangas, H. (2021). *Termien selittäminen tietokirjoissa* [väitöskirja, Helsingin yliopisto]. Helda. <http://urn.fi/URN:ISBN:978-951-51-7738-4>
- Schmitz, K.-D. (2009). Terminological recommendations for software localization. *Language at Work – Bridging Theory and Practice*, 3(5), 1–16. <https://doi.org/10.7146/law.v3i5.6191>
- Suonuuti, H. (2024). *Sanastotyön opas*. Sanastokeskus ry. Noudettu 12.2.2026 osoitteesta https://sanastokeskus.fi/tiedostot/pdf/Sanastotyon_opas.pdf
- Tammes-Peters, H. (2025). *Ohjeistus generatiivisen tekoälyn hyödyntämisestä työn tukena ja apuvälineenä julkisessa hallinnossa*. Valtiovarainministeriö. Julkaisuarkisto Valto. <https://urn.fi/URN:ISBN:978-952-367-846-0>
- Tieteen termipankki. (n.d.a). *Abstrahointi*. Noudettu 13.2.2026 osoitteesta <https://tieteentermipankki.fi/wiki/Filosofia:abstrahointi>
- Tieteen termipankki. (n.d.b). *Antropomorfismi*. Noudettu 10.11.2025 osoitteesta <https://tieteentermipankki.fi/wiki/Nimitys%3Aantropomorfismi>

Vilkka, H. (2025). *Tutki ja kehitä* (6. painos.). Santalahti-kustannus.

Xian, L., Li, L., Xu, Y., Zhang, B. Z. & Hemphill, L. (2024). *Landscape of Generative AI in Global News: Topics, Sentiments, and Spatiotemporal Analysis*. 1–11.

<https://doi.org/10.48550/arxiv.2401.08899>

Xing, X., Zhang, Z. & He, W. (2026). AI technology, AI narrative, and firm value. *Technovation*, 149, 103349, 1–15. <https://doi.org/10.1016/j.technovation.2025.103349>

Yle. (n.d.). *Mitä on Ylen julkinen palvelu?* Yle.fi. Noudettu 2.11.2025 osoitteesta

<https://yle.fi/aihe/a/20-10002653>