

Roope Paukku

# GENERATIIVINEN TEKOÄLY LIIKETOIMINTATIEDON HALLINNAN TUKENA

Kandidaatintutkielma  
Johtamisen ja talouden tiedekunta  
Tarkastaja: Krista Sorri  
Toukokuu 2024

# TIIVISTELMÄ

Roope Paukku: Generatiivinen tekoäly liiketoimintatiedon hallinnan tukena  
Kandidaatintutkielma  
Tampereen yliopisto  
Tietojohtaminen  
Toukokuu 2024

---

Generatiivisen tekoälyn kehitys ja siihen liittyvien sovellusten määrä on kasvanut merkittävästi vuoden 2022 jälkeen. Teknologian helppokäyttöisyyden ja saatavuuden ansiosta organisaatioiden liiketoimintatietoa tullaan väistämättä käsittelemään generatiivisen tekoälyn kautta. Aiheeseen liittyvä tutkimus on kuitenkin vähäistä, ja organisaatioilla voi olla haasteita hahmottaa generatiiviseen tekoölyyn liittyviä sovelluksia ja haasteita. Tämän tutkimuksen tavoitteena on antaa kokonaisvaltainen kuva generatiiviseen tekoölyyn ja liiketoimintatiedon hallintaan liittyvistä sovelluksista sekä niihin liittyvistä haasteista.

Tutkimus toteutettiin kirjallisuuskatsauksena. Tiedonhaussa hyödynnettiin kahta eri tietokantaa aihepiirin kattavan aineiston kokoamiseksi. Aineisto koostui vertaisarvioituista sekä esijulkaisuista tieteellisistä artikkeleista. Suoraan tutkimuskysymyksiin vastaavaa aineistoa ei löytynyt. Tämän takia aineistoa haettiin niin, että sen sisältö liittyi generatiiviseen tekoölyyn ja vähintään yhteen tiedonhallinnan näkökulmaan. Tutkimuksessa käytetty aineisto rajattiin ajallisesti vuoden 2023 ja siitä uudempiin julkaisuihin. Lisäksi sisällön tuli käsitellä tiedon siirtymistä tai jalostumista organisaation sisällä.

Tutkimus tunnisti generatiivisen tekoälyn sovelluksia jokaiseen liiketoimintatiedon hallinnan prosessin vaiheeseen liittyen. Tutkimusaineistosta tunnistettiin myös useita haasteita, jotka vaikuttavat generatiivisen tekoälyn hyödyntämiseen liiketoimintatiedon hallinnassa. Lisäksi tutkimusaineistosta tunnistettiin generatiivisen tekoälyn pääkyvykkyudet liiketoimintatiedon hallinnan kannalta. Tunnistetut kyvykkyudet ja haasteet yhdistettiin liiketoimintatiedon hallinnan prosessin vaiheisiin kokonaiskuvan saamiseksi. Tutkimustulosten perusteella voidaan päätellä, että generatiivisella tekoölyllä on potentiaalia tehostaa liiketoimintatiedon hallintaa, mikäli organisaatiot ottavat huomioon siihen liittyvät rajoitteet ja haasteet.

Avainsanat: generatiivinen tekoäly, liiketoimintatiedon hallinta, tiedon jalostuminen

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin Originality Check -ohjelmalla.

# ABSTRACT

Roope Paukku: Supporting business intelligence management with generative artificial intelligence  
Bachelor's thesis  
Tampere University  
Degree Programme of Information and Knowledge Management  
May 2024

---

The development of generative artificial intelligence and the number of related applications have increased significantly since 2022. Due to the ease of use and availability of generative artificial intelligence, business intelligence of organizations will inevitably be processed through it. However, research on the subject is scarce, and organizations may have challenges in understanding the applications and challenges related to generative artificial intelligence in this context. The aim of this study is to provide a comprehensive picture of the applications related to generative artificial intelligence and business intelligence management, as well as the challenges associated with them.

The study was conducted as a literature review. Two different databases were used to gather a comprehensive set of material on the topic. The material consisted of peer-reviewed and pre-published scientific articles. No material directly answering the research questions was found. Because of this, material was searched so that its content was related to generative artificial intelligence and some aspect of organizational information management. The material used in the study was limited to publications from 2023 onwards. In addition, the content had to deal with the transfer or expansion of information within the organization.

The study identified applications of generative artificial intelligence related to each stage of the business intelligence management process. Several challenges were also identified from the research material, which affect the utilization of generative artificial intelligence in business intelligence management. In addition, the main capabilities of generative artificial intelligence related to business intelligence management were identified from the research material. The identified capabilities and challenges were combined with the process stages of business intelligence management to get an overall picture on the subject. Based on the results, it can be concluded that generative artificial intelligence has the potential to enhance business information management, provided that organizations take into account the constraints and challenges associated with it.

Keywords: generative artificial intelligence, business intelligence management, information expansion

The originality of this thesis has been checked using the Turnitin Originality Check service.

# SISÄLLYSLUETTELO

1. JOHDANTO.....	1
1.1 Tutkimusongelma ja rajaukset.....	2
1.2 Tutkimuksen rakenne.....	3
2. TUTKIMUKSEN KUVAUS.....	4
2.1 Tutkimusmenetelmä.....	4
2.2 Tutkimusmateriaali.....	4
3. LIIKETOIMINTATIETO JA SEN JALOSTUMINEN ORGANISAATIOISSA.....	7
3.1 Tiedon tasot ja sen jalostuminen.....	7
3.2 Liiketoimintatieto ja sen hallinta.....	9
4. TEKOÄLY.....	11
4.1 Tekoälystä yleisesti.....	11
4.2 Generatiivinen tekoäly.....	12
4.3 GPT-malliarkkitehtuuri.....	14
4.4 Toimialakohtaisen tiedon hyödyntäminen generatiivisissa tekoälymalleissa.....	17
5. LIIKETOIMINTATIEDON HALLINTA JA GENERATIIVINEN TEKOÄLY.....	20
5.1 Tiedon jalostuminen generatiivisen tekoälyn avulla.....	20
5.2 Liiketoimintatiedon hallinnan sovellukset.....	21
5.3 Generatiivisen tekoälyn haasteet liiketoimintatiedon hallinnassa.....	25
6. YHTEENVETO.....	29
6.1 Tulokset ja päätelmät.....	29
6.2 Arviointi ja jatkotutkimusehdotukset.....	33
LÄHTEET.....	35

## LYHENTEET JA MERKINNÄT

DIKW	Engl. Data-Information-Knowledge-Wisdom, tiedon tasoja kuvaava lähestymistapa
GPT	Engl. generative pre-trained transformer, yksi generatiivisen tekoälyn mallityypeistä, pohjautuu transformer-arkkitehtuuriin
LLM	Engl. Large language model, suuri kielimalli, esimerkiksi GPT-4
RAG	Engl. Retrieval-augmented generation, tapa täydentää generatiivisten tekoälymallien tietämystä

# 1. JOHDANTO

Generatiivinen tekoäly ja suuret kielimallit ovat aiheina varsin ajankohtaisia. OpenAI:n kehittämä ChatGPT-sovellus teki historiaa saavuttamalla 100 miljoonaa käyttäjää vain kahdessa kuukaudessa (Hu 2023). GPT-3- ja GPT-4-malleihin pohjautuva chatbot-tyylinen sovellus pystyy käymään keskustelua käyttäjiensä kanssa käyttäen luonnollista kieltä ja antamaan suuren koulutusdatamääränsä ansiosta vastauksia lähes mihin tahansa käyttäjiä askarruttavaan kysymykseen.

Generatiivinen tekoäly soveltuu monipuolisuudessaan vuoksi useisiin organisaatioiden tiedonhallinnan tehtäviin (Sumbal & Amber 2024). Generatiiviset tekoälymallit kykenevät esimerkiksi luomaan näkemyksiä, jakamaan tietoa sekä avustamaan päätöksenteossa. Siinä missä organisaatioiden tiedon elinkaarta on kuvattu datan jalostumisella informaatioksi ja siitä eteenpäin tietämykseksi, generatiivinen tekoäly kykenee luomaan tietämystä suoraan datasta algoritmien ja kontekstualisoinnin avulla. (Alavi et al. 2024)

Generatiivisella tekoälyllä on potentiaalia siirtää tiedon jalostamisen painopistettä ihmisten luomasta tiedosta tekoälyn jalostamiin näkemyksiin (Alavi et al. 2024). Generatiivisten tekoälysovellusten yleistyessä myös niiden vastuu päätöksenteossa, mukaan lukien moraalisisissa päätöksissä, kasvaa (Krügel et al. 2023, Feuerriegel et al. 2023 mukaan). Generatiivinen tekoäly voi myös tehostaa ja jopa muuttaa perinteisiä organisaation tiedonhallinnan prosesseja, kuten tiedon hankintaa, jakamista ja hyödyntämistä (Alavi et al. 2024). Voidaankin olettaa, että generatiivisen tekoälyllä ja sen jalostamalla tiedolla tulee olemaan suuri merkitys organisaatioiden päätöksenteossa. Vaikkakin generatiivisen tekoälyn potentiaali organisaatioiden tiedonhallinnassa on tunnistettu, teknologian kehittyvän luonteen ja viimeaikaisten läpimurtojen vuoksi tutkimus aiheesta on kuitenkin vähäistä (Sumbal & Amber 2024).

Generatiivisen tekoälyn kasvavan tiedonhallinnallisen merkityksen ja vähäisen aiheeseen liittyvän tutkimuksen vuoksi tutkielmassa paneudutaan generatiivisen tekoälyn rooliin organisaatioiden liiketoimintatiedon hallinnassa. Aihepiiri on tieteellisesti ajankohtainen olemassa olevan tutkimuksen vähäisyyden takia. Lisäksi uusia, aihepiiriin liittyviä tutkimuksia julkaistaan kiihtyvällä tahdilla. Aihe on mielenkiintoinen myös organisaatioiden näkökulmasta, sillä organisaatioilla voi olla haasteita käsittää generatiivisen tekoälyn mahdollistamien liiketoimintatiedon hallinnan ratkaisujen käyttöön liittyvät haasteet ja toisaalta myös niiden tarjoamat mahdollisuudet. Lisäksi

generatiiviseen tekoälyyn pohjautuvien helppokäyttöisten sovellusten yleistyminen, esimerkiksi ChatGPT tai Microsoft Copilot, johtaa todennäköisesti joka tapauksessa liiketoimintatiedon käsittelyyn työntekijöiden toimesta. Tästä syystä organisaatioiden tulisi ymmärtää generatiivisella tekoälyllä liiketoimintatiedon käsittelemisen seuraukset ja niihin liittyvät mahdolliset riskit. Riskien ja seurausten ymmärtäminen mahdollistaa organisaatioille sopivien sovellusten ja käyttötarkoitusten tunnistamisen sekä näiden pohjalta mahdollisten linjauksien, käytäntöjen ja pitkän tähtäimen suunnitelmien muodostamisen.

## 1.1 Tutkimusongelma ja rajaukset

Tutkielman ensisijaisena tavoitteena on vastata kysymykseen ”Miten generatiivista tekoälyä kannattaa hyödyntää liiketoimintatiedon hallinnan tukena?”. Päättökysymyksen tavoitteena on selvittää generatiivisen tekoälyn optimaalisimmat käyttökohteet liiketoimintatiedon hallinnassa ottaen kokonaisvaltaisesti huomioon sen kyvykkyydet ja sen käyttöön liittyvät haasteet.

Päättökysymystä tukevat alatutkimuskysymykset ovat seuraavat:

- Miten generatiivinen tekoäly voi tukea ihmislähtöistä tiedon jalostamista?
- Mitä sovelluksia generatiiviselle tekoälylle on liiketoimintatiedon hallinnan näkökulmasta?
- Mitä haasteita ja rajoitteita generatiivisen tekoälyn käyttöön liiketoimintatiedon hallinnassa liittyy?

Alatutkimuskysymykset voidaan nähdä suoraan päättökysymyksen tarkempaa jaotteluna. Ensimmäinen alatutkimuskysymys on oleellinen teoreettinen kysymys tutkimuksen aiheen kannalta. Kysymyksessä tarkastellaan ihmislähtöistä tiedon jalostamista nimenomaan generatiivisen tekoälyn avulla, ei generatiivisen tekoälyn toimimista tietoa käyttävänä toimijana. Toinen ja kolmas alatutkimuskysymys ovat käytännönläheisempiä kysymyksiä, jotka vastaavat tarkemmin tutkimuksen pääongelmaan ja antaen kokonaisvaltaisen kuvan.

Liiketoimintatiedon hallintaa käsitteleviltä osin työssä rajataan tarkastelu Laihonon et al. (2013) mukaisen liiketoimintatiedon hallinnan prosessimallin ympärille. Lisäksi työssä tarkastellaan tiedon tasoja ja sen jalostumista, sillä ne ovat oleellinen osa liiketoimintatiedon hallintaa.

Tiedon jalostumista ja tiedon tasoja käsitteleviltä osin työssä rajataan tarkastelu datan ja tietämyksen välille, sillä tämä on tietojohdamisen alalla yleinen tiedon jäsentelytapa

(Laihonen et al. 2013). Tietämyksen osalta työssä keskitytään siihen, tapahtuuko generatiivisen tekoälyn ansiosta sellaisia toimenpiteitä, että koulutusdataa tai syötteitä käsitellessä syntyy sivu- tai päätuotteena jossain muodossa olevaa tietämystä. Näihin rajauksiin liittyvät käsitteet on määritelty alaluvussa 3.1. Lisäksi tiedon jalostumista käsitteleviltä osilta on päätetty rajoittaa tarkastelu erityisesti liiketoimintatietoon.

Työn teknistä osaa käsittelevien osuuksien kohdalta on päätetty rajoittaa tarkastelu tekoälymallien sisäiseen toimintaan, eli siihen miten se jäsentelee koulutusdataa ja jalostaa siitä mahdollista uutta tietoa. Lisäksi työssä tarkastellaan mallien tietämyksen täydentämistä toimialakohtaisella tiedolla. Syötteiden muokkaaminen (*engl. prompt engineering*) ei liity tiedon jalostumiseen tekoälymallin sisällä, joten työssä ei oteta kantaa siihen, kuinka esimerkiksi erilaisia syötteitä antamalla voi saada subjektiivisesti hyödyllisempiä vastauksia. Työssä siis keskitytään tiedon matkaan aina koulutusdatasta syntetisoiduksi vastaukseksi.

Työn teknisessä osassa tarkastelu on myös rajoitettu tekstiä käsitteleviin generatiivisiin tekoälymalleihin, sillä suuri osa organisaatioiden eksplisiittisestä tiedosta on tekstimuodossa, kuten dokumentteina, sähköposteina tai laskentataulukkoina ja esimerkiksi kuvia syntetisoivien mallien tekninen toimintaperiaate eroaa huomattavasti tekstiä syntetisoivista. Tekstiä syntetisoivista generatiivisista tekoälymalleista keskitymme GPT-arkkitehtuuriin pohjautuviin malleihin, sillä nämä ovat generatiivisista tekoälymalleista tunnetuimpia ja niiden pohjalle on kehitetty liiketoimintatiedon hallintaan liittyviä ratkaisuja. GPT-malliarkkitehtuuri on esitelty alaluvussa 4.3.

## 1.2 Tutkimuksen rakenne

Ensimmäisessä luvussa keskitytään työhön johdatteluun ja kuvataan tutkimusongelma sekä tutkimuksen rakenne. Lisäksi ensimmäisessä luvussa esitetään työhön valitut rajaukset. Toisessa luvussa esitetään käytetty tutkimusmenetelmä sekä tutkimusaineisto ja siihen tehdyt rajaukset. Kolmas ja neljäs luku keskittyvät työn kahden teoreettisen näkökulman tarkasteluun: liiketoimintatiedon tasoihin, jalostumiseen ja sen hallintaan sekä generatiiviseen tekoälyyn.

Viidennessä luvussa käsitellään työn tutkimusongelmaa. Luvussa yhdistetään lukujen 3 ja 4 teoriapohjat sekä kerätty tutkimusaineisto alatutkimuskysymyksiin vastaamiseksi. Viimeisessä, eli kuudennessa luvussa pyritään vastaamaan päättökysymykseen vetämällä yhteen tutkimuksen tulokset ja esittämällä mahdolliset päätelmät. Lisäksi luvussa arvioidaan tehtyä tutkimusta sekä esitetään mahdolliset tulevaisuuden jatkotutkimusehdotukset työn aihepiiristä.

## 2. TUTKIMUKSEN KUVAUS

Tässä luvussa esitetään tutkimuksessa käytetty tutkimusmenetelmä. Lisäksi luvussa käsitellään tutkimuksen aineistoa, sen keruuta ja tehtyjä rajoituksia.

### 2.1 Tutkimusmenetelmä

Työn tutkimusmenetelmäksi valittiin kirjallisuuskatsaus. Tutkimuksessa käytetään seitsenvaiheista mallia, joka sisältää järjestyksessä seuraavat kohdat (Fink 2014):

- tutkimuskysymyksen asettaminen
- tietokantojen ja kirjallisuuden valinta
- hakusanojen ja -lauseiden valinta
- käytännön hakukriteerien valinta
- metodologinen rajaus
- katsauksen tekeminen
- tulosten syntetisointi.

Ennen varsinaisen tutkimuksen aloittamista suoritettiin aiheanalyysi, jossa aihepiiriä tarkasteltiin ja rajattiin käsiteltävissä olevaksi. Analyysissä käsiteltiin alustavasti tutkimusongelmaa, tutkimuksen rajausta, tutkimusaineistoa, tutkimukselle keskeisiä käsitteitä sekä mahdollisia tuloksia. Aiheanalyysin lisäksi tutkimusongelmaa, tutkimuskysymyksiä, tutkimusmenetelmää sekä tutkimuksen tavoitteita, rakennetta ja aineistoa tarkasteltiin tarkemmin tutkimussuunnitelmassa.

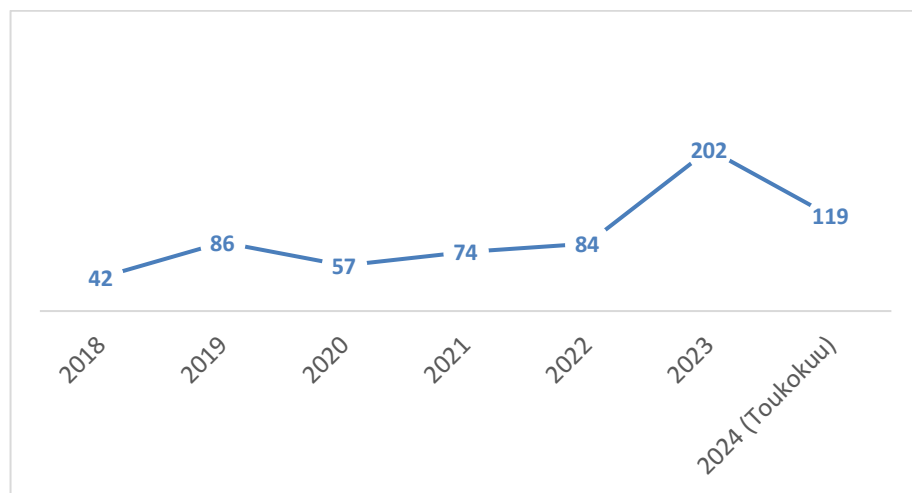
### 2.2 Tutkimusmateriaali

Tässä kappaleessa esitetään tutkimuksen aineisto, sen keruu ja siihen tehdyt rajoitukset. Tutkimuksen aineistoa etsittiin Andor- ja Google Scholar-tietokannoista käyttäen taulukossa 1 esiteltyjä hakulausekkeita.

**Taulukko 1.** Haun tuottamat tulokset hakusanoittain

Hakulauseke	Andor	Google Scholar
("generative ai" OR "generative artificial intelligence") AND "knowledge management"	432	2640
("generative ai" OR "generative artificial intelligence") and "business intelligence"	131	1430
("generative ai" OR "generative artificial intelligence") and "knowledge management" and "business intelligence"	10	702

Taulukosta 1 nähdään, miten hakulausekkeita muokkaamalla tutkimuksen aineistoa rajattiin. Aineistoa löytyi hakutulosten määrän perusteella paljon, mutta tarkemmalla tarkastelulla suuri osa aineistosta on ei-tieteellistä sisältöä, kuten uutiskirjeitä tai kuluttajalehtiartikkeleita. Esimerkiksi taulukossa 1 viimeisenä esitetyllä hakulausekkeella Andorista haetut tulokset olivat kaikki uutiskirjeitä, toisin sanoen ei-tieteellistä sisältöä. Ei-tieteelliset tulokset poistettiin tutkimusaineistosta, jolloin ainoastaan Google Scholarista saatu aineisto päätyi jatkotarkasteluun. Kuvassa 1 esitetään Google Scholarin antamien hakutuloksien vuosittaista määrää annetulle hakulausekkeelle.



**Kuva 1:** Google Scholarin antamat hakutulokset vuosittain hakulausekkeella “("generative ai" OR "generative artificial intelligence") and "knowledge management" and "business intelligence"”

Kuvasta 1 voidaan nähdä, että aihepiiriin liittyvä tutkimus on kasvanut huomattavasti vuoden 2023 jälkeen. Tutkimusaineisto rajattiin alkamaan vuodesta 2023, sillä generatiiviset tekoälyteknologiat ja niiden sovellukset kehittyvät erittäin nopeasti.

Tutkimusaineistoon sisällytettiin vain vertaisarvioituja tai esijulkaistuja tieteellisiä artikkeleita.

Hakutuloksien tarkemmat rajauskriteerit on esitelty taulukossa 2. Rajauskriteerien tarkoituksena on kasata aiheen kannalta relevantti ja käsiteltävän kokoinen tutkimusaineisto. Tutkimusaineiston halutaan käsittelevän generatiivista tekoälyä jonkin tiedonhallinnan linssin kautta, sillä suoraan tutkimuskysymykseen vastaavia artikkeleita ei löytynyt. Lisäksi tutkimusaineiston haluttiin käsittelevän tiedon elinkaarta tai sen siirtymistä organisaation sisällä.

**Taulukko 2.** Tutkimusaineiston rajauskriteerit

Rajauskriteeri	Tiedonhallinta	Generatiivinen tekoäly	Tiedon jalostuminen	Vuosi
<b>Sisällyttäminen tutkimukseen</b>	Sisältää tietoa tiedonhallinnan, tietojohdantamisen tai liiketoimintatiedon hallinnan menetelmistä, prosesseista tai elinkaarista	Liittyy generatiiviseen tekoälyyn, suuriin kielimalleihin ja näihin liittyviin tekniikoihin. Voi liittyä mihin tahansa yleiseen toimialaan.	Sisältää tietoa tiedon elinkaaresta tai liikkumisesta organisaation sisällä	2023 tai uudempi
<b>Tutkimuksesta jääväminen</b>	Liittyy tiedonhallintaan, mutta ei käsittele generatiivista tekoälyä	Liittyy generatiiviseen tekoälyyn, mutta ei käsittele tiedonhallintaa tai tiedon elinkaarta	-	2022 tai vanhempi

Artikkeleista luettiin tiivistelmät ja niiden teoriapohjaa, esimerkiksi tiedonhallintaan liittyviä näkökulmia, tarkasteltiin niiden sopivuuden määrittelemiseksi. Lopullisen tutkimusaineiston muodostavat liitteessä 1 esitellyt 8 artikkelia. Artikkelien määrä on rajallinen, mutta toisaalta rajauskriteerien ansiosta artikkelit sisältävät laajat teoreettiset linssit, joiden avulla tutkimuskysymyksiin vastaamiseksi on saatu riittävästi aineistoa. Tutkimusaineisto on ristiintaulukoitu liiketoimintahallinnan prosessin vaiheiden ja generatiiviseen tekoälyyn liittyvien haasteiden mukaisesti liitteessä 2. Yhdistettynä luvuissa 3 ja 4 kappaleissa esitettyihin teoriapohjiin tutkimusaineisto riittää vastaamaan tutkimuskysymyksiin tutkielman edellyttämällä tasolla.

### 3. LIIKETOIMINTATIETO JA SEN JALOSTUMI- NEN ORGANISAATIOISSA

Tässä luvussa käsitellään tietoa, sen tasoja sekä sen jalostumista. Lisäksi luvussa tarkastellaan liiketoimintatietoa ja sen jalostumiseen liittyviä prosesseja organisaatioissa. Tutkimuksessa hyödynnetään DIKW-pyramidia tiedon mallintamistyökaluna. DIKW-pyramidi on yleisesti käytetty tapa tiedon jaotteluun eri tasoihin, joita ovat data, informaatio (*engl. information*), tietämys (*engl. knowledge*) ja viisaus (*engl. wisdom*) (Bellinger et al. 2004; Baskarada & Koronios 2013). Tutkimus keskittyy pyramidin kolmeen alimpaan tasoon, eli dataan, informaatioon ja tietämykseen, sillä se on tietojohdamisen alalla yleinen tiedon jäsentelytapa (Laihonen et al. 2013, s. 17-18). Lisäksi työssä hyödynnetään organisaation tietoperustaista näkemystä (*engl. knowledge-based theory of the firm*), sillä tiedon jalostuminen organisaatioissa on oleellinen näkemys tutkimuksen kannalta (Nonaka 2008).

Tieto on resurssina ainutlaatuinen, sillä sen arvo ei vähene sitä käytettäessä eikä jaettaessa useille eri yksilöille (Laihonen et al. 2013, s. 52; Nonaka 2008, s. 6-7) Tieto on paikatonta ja ajatonta, joten toisin sanoen se on rajaton resurssi. Tietoa voidaan kuluttaa ja yhdistää uuteen tietoon. Juuri tiedon yhdistäminen sekä kuluttaminen synnyttää uutta tietoa ja luo tiedolle arvoa kontekstin kautta. Merkittävä tiedon ominaisuus on se, että tietoa luodaan juuri ihmisten välisissä vuorovaikutuksissa. Nämä ominaisuudet erottavat tiedon muista organisaatioiden resursseista, kuten raaka-aineista. (Nonaka 2008, s. 6-7)

#### 3.1 Tiedon tasot ja sen jalostuminen

DIKW-pyramidin alimmalla tasolla on data, joka kuvaa itsessään merkityksettömiä ja rakenteettomia faktatietoja ja objektiivisia havaintoja (Bellinger et al. 2004; Baskarada & Koronios 2013; Laihonen et al. 2013, s. 17-18). Data ei itsessään ole merkityksellistä, mutta sitä voidaan jalostaa merkityksellisemmäksi tiedoksi.

Tiedon seuraava taso, informaatiota, voidaan kuvata tulkituksi dataksi (Baskarada & Koronios 2013). Informaatiolle on luotu merkitys esimerkiksi kontekstin kautta ja täten sitä voidaan hyödyntää esimerkiksi päätöksenteossa (Bellinger et al. 2004; Baskarada & Koronios 2013). Informaation tasolla on ymmärrystä alemman tason tiedon keskinäisistä suhteista. Informaation avulla voidaan esimerkiksi vastata kysymyksiin ”kuka”, ”mitä”, ”missä” ja ”milloin”. (Bellinger et al. 2004)

Tutkimuksen kannalta tiedon korkein taso, tietämys, on yhdistelmä prosessoitua, jäsenneilyä ja tulkittua dataa ja informaatiota sekä aiempaa tietoa, osaamista ja kokemuksia, joka mahdollistaa syvemmän ymmärryksen. Tietämystä pidetään hiljaisen tiedon muotona. (Baskarada & Koronios 2013; Laihonen et al. 2013, s. 17-18) Tietämyksen avulla voidaan kuvata toistuvuuksia ja loogisia syy-seuraussuhteita asioiden välillä, mahdollistaen esimerkiksi johtopäätösten tekemisen ja tulevaisuuden ennakoimisen. Tietämystä voidaankin hyödyntää sellaisenaan päätöksentekoon. (Bellinger et al. 2004; Baskarada & Koronios 2013) Tietämyksen avulla voidaan vastata kysymykseen ”miksi” (Bellinger et al. 2004). Tutkielman kannalta relevantit tiedon tasot on koottu taulukkoon 3.

**Taulukko 3.** Tiedon tasot (Bellinger et al. 2004; Baskarada & Koronios 2013; Laihonen et al. 2013, s. 17-18)

<b>Tietämys</b>	Yhdistelmä dataa, informaatiota ja kokemuspohjaista tietoa, jota voi hyödyntää sellaisenaan esimerkiksi päätöksentekoon.
<b>Informaatio</b>	Tulkittua ja jäsenneilyä dataa. Dataa, jolle on annettu merkitys.
<b>Data</b>	Merkityksetöntä ja rakenteetonta faktatietoa ja havaintoja.

Tietoa siis määritellään sen erityisen ja laajan luonteen vuoksi esimerkiksi jaottelemalla tieto eri tasoihin tai erottamalla hiljainen ja eksplisiittinen tieto toisistaan. Tietoa voidaan myös kuvata jatkuvana prosessina, sillä käyttäjän kuluttaessa tietoa syntyy hänelle uutta tietoa. Yhdisteltäessä uutta tietoa olemassa olevaan tietoon ja kokemuksiin syntyy myös uutta tietoa. (Nonaka 2008, s. 6-15) Tieto ei siis kulu, vaan se kehittyy, moninkertaistuu ja jalostuu käyttäjiensä vaikutuksesta.

Tiedon jalostumisella kuvataan tiedon kehittymistä yksinkertaisista tosiasioista, faktoista ja kokemuksista arvokkaaksi tiedoksi ja ymmärrykseksi. Jalostettu, korkeamman tason tieto on käyttäjälleen arvokkaampaa kuin jalostamaton, matalan tason tieto. Tiedon jalostumisella kuvataan siis prosessia, jossa tiedon arvo ja laatu kasvavat. (Baskarada & Koronios 2013; Laihonen et al. 2013). Tiedon jalostumista pidetään sosiaalisissa vuorovaikutuksissa sekä yksilöllisissä kognitiivisissa prosesseissa tapahtuvana prosessina (Alavi et al. 2024).

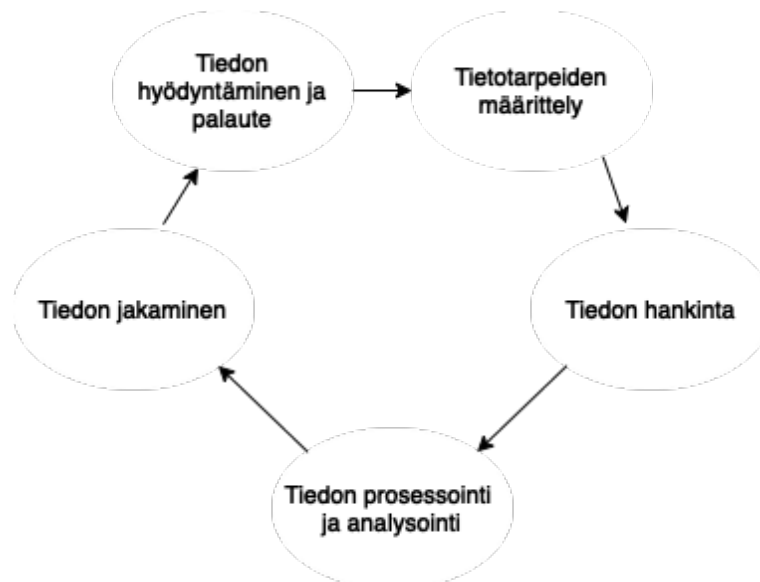
Nonakan (2008, s. 18-26) esittelemä SECI-malli kuvaa tiedon muuttumista ja jalostumista organisaatiossa. SECI-mallin vaihteita ovat sosialisointi, ulkoistaminen, yhdisteleminen, sekä sisäistäminen. Sosialisointiossa hiljaista tietoa siirretään toisille henkilöille, ulkoistamisessa hiljaista tietoa muunnetaan eksplisiittiseksi tiedoksi jolloin se muuttuu hyödynnettäväksi muille, yhdistelemisessä eksplisiittisiä tiedonpalasia yhdistetään uudeksi tiedoksi ja lopulta sisäistämässä eksplisiittistä tietoa muunnetaan

hiljaiseksi tiedoksi ymmärtämisen ja kokemuksen kautta. (Nonaka 2008, s. 18-26; Laihonen et al. 2013, s. 56-58)

### 3.2 Liiketoimintatieto ja sen hallinta

Liiketoimintatiedoksi (*engl. business intelligence, BI*) voidaan käsittää kaikki organisaation tuottama ja hyödyntämä sisäinen ja ulkoinen tieto. Liiketoimintatiedolle tyypillistä on, että sen tarkoitus on tukea organisaation päättäjiä päätöksenteossa ja suunnittelussa. (Foley & Guillemette 2010; Laihonen et al. 2013, s. 45; Negash 2004)

Liiketoimintatieto ei ole automaattisesti arvokasta organisaatiolle ja sitä tuleekin johtaa, jotta siitä olisi hyötyä (Laihonen et al. 2013, s. 45). Tätä toimintaa kutsutaan liiketoimintatiedon hallinnaksi. Liiketoimintatiedon hallinnan avulla tietoa jalostetaan päätöksenteon tueksi. (Laihonen et al. 2013, s. 45) Onnistunut liiketoimintatiedon hallinta tyydyttää organisaation tietotarpeet osuvasti ja oikea-aikaisesti. Liiketoimintatiedon hallinnan tehtävänä on yhdistää näennällisesti irrallisia tiedonpalasia niiden kontekstin ja merkityksen ymmärtämiseksi (Laihonen et al., s. 45). Parhaimillaan liiketoimintatiedon hallinta parantaa organisaation kilpailukykyä, tuottavuutta ja kannattavuutta tehokkaamman, tietoperustaisen päätöksenteon kautta (Laihonen et al. 2013, s. 50). Liiketoimintatiedon hallinnan prosessimalli on esitelty kuvassa 2.



**Kuva 2:** Liiketoimintatiedon hallinnan prosessimalli (mukaillen Laihonen et al. 2013, s. 46)

Kuten kuva 2 osoittaa, liiketoimintatiedon hallinta on jatkuva prosessi. Prosessin aloituspisteenä voidaan kuitenkin pitää tietotarpeiden määrittelyä. Tässä vaiheessa

tavoitteena on selvittää mitä tietoa päätöksenteon tueksi tarvitaan, milloin ja missä muodossa (Laihonen et al. 2013, s. 47). Teknologinen kehitys, nopeat tiedonsiirtoyhteydet ja globalisaatio ovat lisänneet saatavilla olevan tiedon määrää, ja liiallinen tieto voi pahimmassa tapauksessa hankaloittaa päätöksentekoa. Tietotarpeiden määrittelyllä koitetaan siis kohdentaa tiedon hankintaa paremmin ja täten vähentää turhan tiedon keräämistä. (Laihonen et al. 2013, s. 44, 47)

Prosessin seuraava vaihe, tiedon hankinta, keskittyy ensimmäisessä vaiheessa määriteltujen tietotarpeiden täyttämiseen. Tiedon hankinnassa tärkeää on tiedon laatu ja käyttökelpoisuus, ei pelkästään sen määrä. Tiedon luotettavuus ja oikeellisuus eivät myöskään ole itsestäänselvyksiä, ja tiedolle voidaan tarvittaessa vaatia useampi lähde. Tietolähteet vaihtelevat organisaatioista, toimialasta ja tilanteista riippuen aina haastatteluista ja kirjallisuudesta organisaation tietojärjestelmiin sekä internettiin. Tietolähteet ovat usein eksplisiittisiä lähteitä, kuten tietokantoja, mutta hyödyllisimmät tietolähteet voivat olla myös inhimillisiä hiljaisen tiedon lähteitä. (Laihonen et al. 2013, s. 47-48)

Hankittu tieto ei välttämättä ole päätöksenteon kannalta arvokasta sellaisenaan. Prosessin kolmas vaihe, tiedon prosessointi ja analysointi, keskittyy hankitun tiedon karsimiseen, arvioimiseen ja luokitteluun organisaation tarpeiden mukaisesti (Laihonen et al. 2013, s. 48). Uutta hankittua tietoa voidaan yhdistää olemassa olevaan organisaation tietoon, jolloin sille annetaan organisaationaalinen konteksti. Tietoa voidaan myös analysoida eri menetelmin, esimerkiksi tilastollisesti tai visuaalisesti. (Laihonen et al. 2013, s. 48) Tässä vaiheessa hankittua tietoa siis jalostetaan hyödyllisemmäksi, korkeamman tason tiedoksi. Laihosen et al. (2013, s. 48) mukaan tiedon analysoinnissa inhimillinen panos on oleellista tiedon merkityksen arvioimiseksi ja johtopäätösten tekemiseksi.

Prosessin neljännellä vaiheella, tiedon jakamisella, viitataan jalostettujen tietotuotteiden siirtämiseen päätöksentekoa varten. Tietoa voidaan jakaa niin sanottuina tietotuotteina, joita jalostetaan prosessin kolmannessa vaiheessa. Tietotuotteita on esimerkiksi raportit, analyysit ja uutiskoosteet. Tietoa voidaan myös siirtää muilla keinoin, esimerkiksi sähköpostitse, tietojärjestelmän kautta tai suorassa vuorovaikutuksessa. (Laihonen et al. 2013, s. 48-49)

Lopuksi, jotta jalostetusta tiedosta olisi hyötyä päätöksenteossa, toisin sanoen jotta tieto olisi arvokasta, tulee sitä hyödyntää jonkin tavoitteen saavuttamiseksi. Tieto voi esimerkiksi vahvistaa käsityksiä tai tuoda uusia näkökulmia. (Laihonen et al. 2013, s. 49) Tieto voi myös nostattaa uusia kysymyksiä, jolloin prosessi alkaa alusta.

## 4. TEKOÄLY

Tässä luvussa käsitellään tekoälyä ja sen alalajeja yleisellä tasolla sekä tekoälyn yleisiä käyttökohteita. Luvussa tarkastellaan tarkemmin generatiivista tekoälyä sekä sen taustalla olevia teknologioita ja toimintaperiaatteita. Lisäksi luvussa tarkastellaan toimialakohtaisen tiedon hyödyntämistä generatiivisen tekoälyn yhteydessä. Tavoitteena on luoda lukijalle ymmärrys generatiivisen tekoälyn toiminnasta sekä siitä, miten generatiivinen tekoäly voi hyödyntää organisaation liiketoimintatietoa.

### 4.1 Tekoälystä yleisesti

McCulloch ja Pitts esittivät vuonna 1943 mallin keinotekoisista neuroneista, joissa neuronit aktivoituivat niitä ympäröivien neuroneiden aktivoitumisesta. Tätä mallia pidetään ensimmäisenä tekoälyä esittävänä työnä. McCulloch ja Pitts osoittivat myös, että tällaisia neuroverkkoja voisi hyödyntää minkä tahansa laskentafunktion suorittamiseen. Lisäksi he esittivät, että neuroverkot voisivat myös oppia. Termi tekoäly (artificial intelligence) esiintyi ensimmäisen kerran 1950-luvulla. (Russell & Norvig 2016, s. 16)

Nykyään tekoälysovelluksia esiintyy lähes kaikkialla. Kohdennettu mainonta, sähköpostin roskapostisuodattimet ja navigointisovellukset ovat esimerkkejä erilaisista tekoälyistä. Tekoälyteknologiat ovat kehittyneet erittäin hienostuneiksi ja oleellisiksi päivittäisen elämän kannalta. (Marr 2019) Tekoäly tieteenalana on kasvanut myös huomattavasti. Tekoälyartikkelien määrä maailmanlaajuisesti on yli kaksinkertaistunut pelkästään vuosien 2015 ja 2021 välillä (Our World in Data 2024).

Tekoäly käsitteenä viittaa ympäristöönsä älykkäästi reagoiviin koneisiin (Stephenson 2018, kpl. 2). Tekoäly voidaan jakaa myös rationaalisesti ja inhimillisesti käyttäytyviin koneisiin (Russell & Norvig 2016, s. 2). Tekoäly kattokäsitteenä sisältää monta eri osa-aluetta. Tarkastellaan seuraavaksi tämän tutkielman kannalta merkittävimpiä tekoälyn osa-alueita.

Koneoppiminen (*engl. machine learning, ML*) on tekoälyn osa-alue, jossa koneoppimismallit oppivat säännönmukaisuuksia esimerkeistä ja havainnoista (Janiesch et al. 2021). Koneoppiminen kuvaa joukkoa erilaisia menetelmiä, joilla voidaan kouluttaa tietokone ratkaisemaan erilaisia todellisen maailman ongelmia ilman, että suoraa ratkaisua ohjelmoidaan (Koza et al. 1996, Kühl et al. 2022 mukaan). Koneoppiminen voidaan jakaa ohjattuun (*engl. supervised*) sekä ohjaamattomaan (*engl. unsupervised*)

oppimiseen. Kolmas koneoppimisen osa-alue on vahvistusoppiminen (*engl. reinforcement learning, RL*). Ohjatussa oppimisessa koulutusdata sisältää mitattavan tiedon sekä siihen liittyvän selitteen. Ohjaamattomassa oppimisessä käytössä on vain mitattava tieto. Vahvistusoppimisessa tekoälytoimijat (*engl. AI Agent*) oppivat palkintojen ja rangaistusten avulla. (Russell & Norvig 2016, s. 693–695; Kühl et al. 2022; Kämäräinen 2023, s. 94–97, 152)

Keinotekoiset neuroverkot (*engl. artificial neural networks, ANN*) ovat yksi koneoppimisen lähestymistavoista (Kämäräinen 2023, s. 116–121). Keinotekoiset neuroverkot koostuvat yksittäisistä päätöksistä tekevistä päättelimestä (*engl. perceptron*), joita on yhdistelty rinnakkain ja peräkkäin ihmisaivojen hermosolujen tapaan (Banh & Strobel 2023; Kämäräinen 2023, s. 116–121). Keinotekoisia neuroverkkoja saatetaan myös kuvata termillä syväoppiminen (*engl. deep learning, DL*), jolloin usein viitataan monikerroksisiin neuroverkkoihin (Banh & Strobel 2023; Kämäräinen 2023, s. 116–121). Neuroverkkojen avulla on mahdollista havaita toistuvuuksia ja hierarkioita monimutkaisesta ja suuresta datamäärästä (Banh & Strobel 2023), ja riittävän koulutusdatan avulla neuroverkot voivat approksimoida lähes mitä tahansa funktiota (Kämäräinen 2023, s. 116–121). Aiemmin mainittujen ominaisuuksiensa ansiosta neuroverkkoja voidaan soveltaa moneen eri koneoppimisen ongelmaan, muun muassa kuvan, kielen ja äänen prosessointiin (Banh & Strobel 2023; Kämäräinen 2023, s. 116–121).

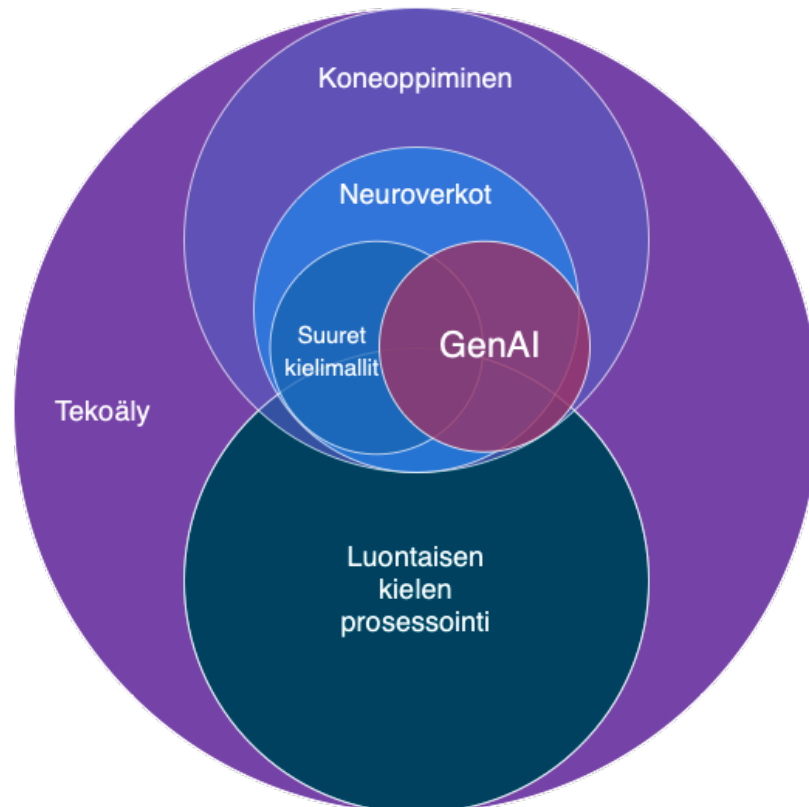
Luontaisen kielen prosessointi (*engl. natural language processing, NLP*) on tekoälyn osa-alue, joka keskittyy luontaisen kielen käsittelyyn (Russell & Norvig 2016, s. 860). Luontaisen kielen prosessointiin liittyvät tehtävät voivat olla esimerkiksi tekstin luokittelua, käännöstehtäviä ja puheentunnistusta (Russell & Norvig 2016, s. 860, 907; Deng & Liu 2018, s. 1-2).

Suuret kielimallit (*engl. large language models, LLM*) ovat suurella, luokittelemattomalla datamäärällä koulutettuja kielimalleja, jotka kykenevät suorittamaan useita erilaisia luontaisen kielen prosessointiin liittyviä tehtäviä. Tehtävät voivat sisältää esimerkiksi käännöstehtäviä, tekstin tiivistämistä, tekstin generoimista ja loogista päättelyä. ChatGPT-sovelluksen takana olevat GPT-3.5- ja GPT-4-mallit ovat tunnettuja esimerkkejä suurista kielimalleista. (Liu et al. 2023)

## 4.2 Generatiivinen tekoäly

Generatiivinen tekoäly kykenee luomaan uutta sisältöä, kuten tekstiä, kuvia, ääntä tai synteettistä dataa (Mandapuram et al. 2018; Feuerriegel et al. 2023; Google Cloud Tech

2023; Linkon et al 2024). Generatiiviset tekoälymallit muodostavat yksinkertaistetun esityksen koulutusdatasta ja luovat tämän avulla uutta sisältöä, joka muistuttaa alkuperäistä dataa mutta ei kuitenkaan ole kopio siitä (Google Cloud Tech 2023; Martineu 2023). Kuvaan 3 on mallinnettu generatiivisen tekoälyn ja muiden tekoälyyn liittyvien termien keskinäisiä suhteita lähteiden mukaisesti (Banh & Strobel 2023; Google Cloud Tech 2023; Iorliam & Ingio 2024; Janiesch et al. 2021; Linkon et al 2024; Liu et al. 2023).



**Kuva 3:** Generatiivinen tekoäly tekoälyn osa-alueena

Kuvan 3 mukaisesti generatiiviset tekoälymallit hyödyntävät menetelmiä koneoppimisesta sekä luontaisen kielen prosessoinnista. Vaikkakin generatiivinen tekoäly hyödyntää koneoppimisen metodeja, sen oppimistavoite ja koulutustapa eroavat kuitenkin perinteisestä koneoppimisesta (Feuerriegel et al. 2023).

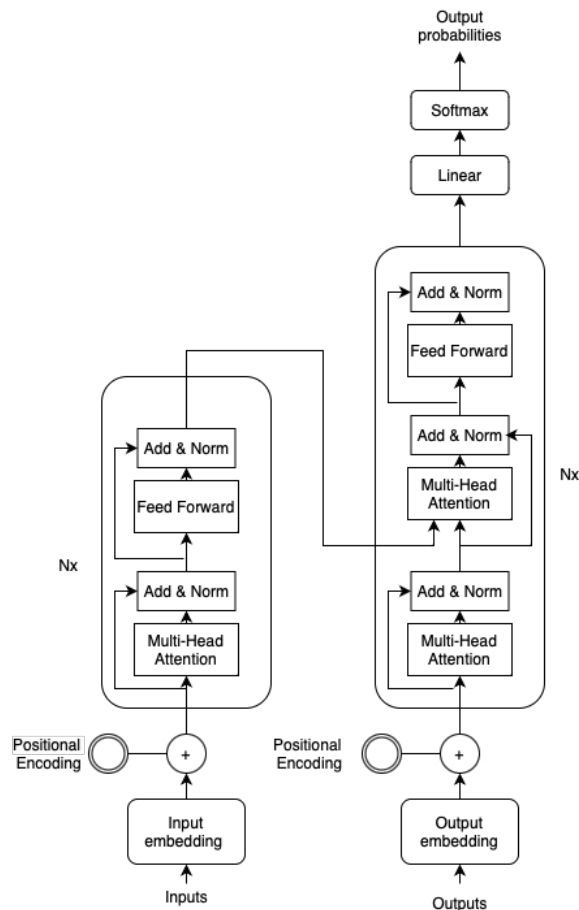
Perinteisiä koneoppimismalleja on käytetty esimerkiksi tunnistamaan, luokittamaan tai ennustamaan muuttujia tai luokkia datan perusteella. Generatiiviset tekoälymallit koulutetaan oppimaan säännönmukaisuuksia, toistuvuuksia ja konteksteja rakenteettomassa datassa (Feuerriegel et al. 2023; Google Cloud Tech 2023; Linkon et al 2024). Matemaattisin termein voidaan puhua todennäköisyysjakauman  $P(X, Y)$  oppimisesta annettujen havaintojen  $X$  ja selitteiden  $Y$  välillä. Tämän pohjalta mallit voivat esimerkiksi syntetisoida uusia havaintoja annettujen selitteiden perusteella. (Bishop

2006, Feuerriegel et al. 2023 mukaan) Generatiiviset tekoälymallit siis hyödyntävät oppittuja säännönmukaisuuksia ja konteksteja muodostaakseen uutta dataa (Feuerriegel et al. 2023; Google Cloud Tech 2023).

### 4.3 GPT-malliarkkitehtuuri

Erilaisia generatiivisen tekoälyn malliarkkitehtuureita on olemassa useisiin eri käyttötarkoituksiin. Esimerkiksi *Stable Diffusion*, jota käytetään kuvien luomiseen tekstisyötteistä, *MusicLM*, jota käytetään musiikin syntetisointiin tai *AlphaCode* jota käytetään ohjelmakoodin luomiseen (Feuerriegel et al. 2023). Eri käyttötarkoituksiin luotujen mallien koulutusmenetelmät myös eroavat toisistaan (Feuerriegel et al. 2023). GPT-mallien (*engl. Generative Pre-trained Transformer*) taustalla oleva transformer-arkkitehtuuri on tutkielman näkökulmasta keskeisin, sillä GPT-mallit keskittyvät tekstin syntetisointiin, ja suurin osa organisaatioiden liiketoimintatiedosta on tekstimuodossa. Transformer-arkkitehtuuri on yleisimmin käytetty generatiivisen tekoälyn arkkitehtuuri (Linkon et al. 2024).

GPT englanninkielisen nimensä mukaisesti pohjautuu alalla kuuluisassa Vaswani et al. (2017) julkaisemassa *Attention is all you need* -artikkelissa esitellylle transformer-arkkitehtuurille. OpenAI:n lisäksi myös Google on kehittänyt transformer-arkkitehtuurille pohjautuvia generatiivisia tekoälymalleja, joista viimeisin, Ultra, ja siihen pohjautuva chatbot Gemini, ovat osoittautuneet GPT-4:sta paremmaksi esimerkiksi vastausten tiedon oikeellisuudessa ja ongelmanratkaisutilanteissa (Rane, Choudhary & Rane 2024). Transformer-arkkitehtuuri on esitelty kuvassa 4.



**Kuva 4:** Transformer-arkkitehtuuri (mukailien Vaswani et al. 2017)

Transformer-mallin voi jakaa kuvassa 4 vasemmalla olevaan enkooderiin ja oikealla olevaan dekooderiin (Vaswani et al. 2017). Enkooderin tarkoitus on esittää syötetty teksti moniulotteisessa vektoriavaruudessa. Dekooderin tarkoitus on muuntaa tämä vektorimuotoinen esitys takaisin alkuperäiseen dataformaatiin. (Alto 2023)

Ennen malliin syöttämistä alkuperäinen teksti jaetaan osiin, tokeneihin, kuten sanoihin ja sanojen osiin. Nämä tokenit muutetaan vektorimuotoiseksi esitykseksi "Input Embedding"-vaiheessa. (Yenduri et al. 2023) Näiden vektoreiden matemaattinen etäisyys toisistaan kuvaa niiden samankaltaisuutta. Suuri etäisyys kuvaa heikkoa samankaltaisuutta ja pieni etäisyys kuvaa vahvaa samankaltaisuutta. (Alto 2023) "Positional Encoding"-komponentti lisää vektoreihin tietoa niiden suhteellisesta tai absoluuttisesta sijainnista tekstissä (Vaswani et al. 2017; Alto 2023; Yenduri et al. 2023). Tässä vaiheessa luodaan siis vektorimuotoinen esitys sanoista ja tokeneista, joka sisältää tietoa niiden semanttisesta merkityksestä sekä niiden kontekstista.

Seuraavaa komponenttia, "Multi-Head Attention":ia, käytetään tekstin aiempien sanojen merkityksen selvittämiseen (Alto 2023). Kyseinen vaihe auttaa selvittämään mihin tekstin osioihin tulisi keskittyä ja mitkä sanat merkitsevät tekstissä eniten tulevien sanojen

kannalta (Alto 2023; Yenduri et al. 2023). Yksittäinen attention-komponentti laskee matriisien Q (query), K (key) ja V (value) avulla pisteytyksen annetun tekstin sanoille käsiteltävän sanan kannalta. Korkea pisteytys kuvaa suurta merkitsevyyttä ja matala pisteytys pientä merkitystä sanojen välillä. Malli oppii edellämainittujen matriisien arvot koulutusprosessin aikana. "Multi-Head" viittaa siihen, että attention-prosessia suoritetaan rinnakkaisesti eri tekstinosille. "Add & Norm"-vaiheessa näiden prosessien tulokset yhdistetään ja normalisoidaan. (Alto 2023)

Viimeinen komponentti, "Feed Forward", muuntaa edellisessä vaiheessa muodostetun esityksen lopulliseen tulokseen sopivaksi. Komponentti koostuu neuroverkosta (*engl. fully connected layer*) sekä aktiivointifunktiosta, joka muuntaa dekooderille lähtevän ulostulon ei-lineaariseksi. (Alto 2023)

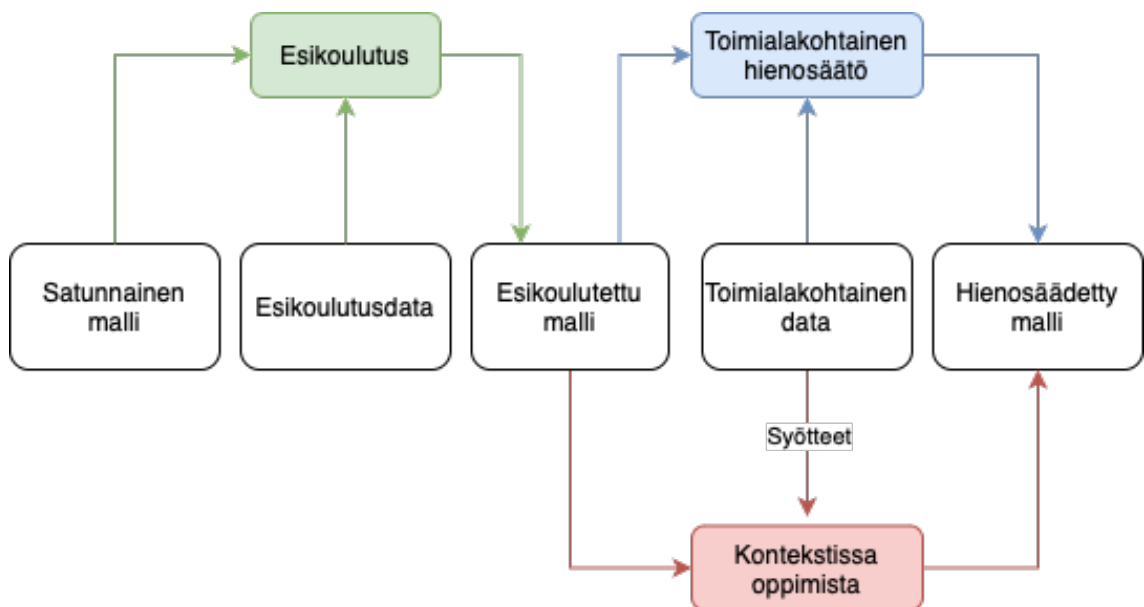
Dekooderi muistuttaa muilta ydinosiltaan enkooderia, mutta sen täytyy lisäksi muuntaa sille syötetty vektorimuotoinen esitys alkuperäiseen dataformaattiin (Alto 2023), tutkimuksen tapauksessa tekstiksi. Kuvan 3 Linear- ja Softmax-komponentit vastaavat tästä muunnoksesta. Linear-komponentti hyödyntää neuroverkkoa, joka muuntaa sisääntulevan esityksen ulottuvuuksien määrän vastaamaan ulostulon ulottuvuuksien määrää (Alto 2023). Softmax-komponentti muuntaa tämän esityksen todennäköisyyksiä sisältäväksi vektoriksi. Tästä vektorista valitaan se sana, jolla on suurin todennäköisyys. (Alto 2023; Yenduri et al. 2023) GPT-mallit hyödyntävät pelkästään dekooderia, eli ne syöttävät annetun tekstisyötteen suoraan dekooderiin. (Alto 2023)

GPT-mallin koulutus sisältää kaksi päävaihetta: esikoulutus (*engl. pre-training*) ja hienosäätö (*engl. fine-tuning*) (Yenduri et al. 2023). Esikoulutuksessa malli koulutetaan ohjaamattomasti suurella määrällä dataa (Yenduri et al. 2023), esimerkiksi GPT-3:n tapauksessa tekstikorpuksella, joka sisälsi 374 miljardia sanaa (Alto 2023). Hienosäädössä esikoulutettua mallia voidaan mukauttaa eri käyttötarkoituksiin (Alto 2023; Yenduri et al. 2023). Käytännössä tällä tarkoitetaan mallin parametrien säätämistä tietyn tehtävän optimoimiseksi käyttäen esimerkiksi vahvistusoppimista (Yenduri et al. 2023). ChatGPT:n tapauksessa esikoulutettua GPT-3-mallia hienosäädettiin ihmisohtajan vahvistusoppimisen (*engl. Reinforcement Learning from Human Feedback, RLHF*) avulla sen dialogin parantamiseksi (Alto 2023). Esikoulutettua mallia voi myös hienosäätää muihinkin käyttötarkoituksiin, kuten tekstin luokitteluun (Yenduri et al. 2023).

#### 4.4 Toimialakohtaisen tiedon hyödyntäminen generatiivisissa tekoälymalleissa

Yleisesti saatavilla olevat GPT-mallit ovat koulutettu ymmärtämään luonnollista tekstiä ja luomaan sisältöä tehokkaasti. Tämä voi kuitenkin johtaa siihen, että erityisillä toimialoilla mallien suorituskyky ei välttämättä ole yhtä tehokasta. Yksi tärkeimmistä haasteista on alakohtaisen datan saatavuus. (Yenduri et al. 2023) Mallien koulutusdata voi myös itsessään olla vanhentunutta ja se saattaa sisältää epäluotettavia lähteitä (Liu et al. 2023). Eri alojen erikoistunut termistö, sääntelykehykset ja sidosryhmien välinen dynamiikka voi tuottaa GPT-malleille haasteita vastausten oikeellisuuden sekä kontekstin ymmärtämisen suhteen (Kishore 2024). Omien suurten kielimallien luominen taas vaatisi merkittävästi resursseja (Xia et al. 2024). Organisaatioilla voi siis olla tarvetta täydentää olemassaolevien mallien tietämystä ulkoisella ja sisäisellä liiketoimintatiedolla.

Yksi ratkaisu täydentää mallin tietämystä on mallin hienosäätö (Liga & Robaldo 2023; Yenduri et al. 2023; Kishore 2024). OpenAI:n (2024) mukaan hienosäätö voi parantaa GPT-mallin tuloksia, vähentää sen käyttämiä tokeneita ja täten myös nopeuttaa mallin toimintaa. Hienosäädön tavoitteena on optimoida mallin käytöstä uutta data-aineistoa vastaavaksi (Kishore 2024). Hienosäädön prosessi on esitelty kuvassa 5.



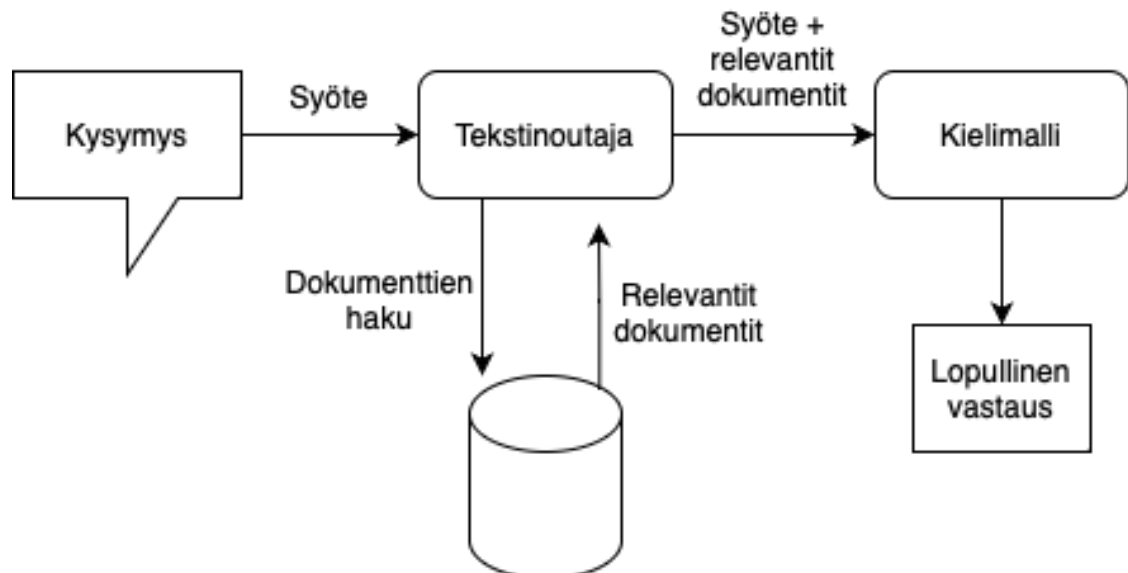
**Kuva 5:** Tyypillinen suurten kielimallien esikoulutus- ja hienosäätöprosessi (mukaillen Linkon et al. 2024)

Hienosäätö on parantanut mallien vastauksien oikeellisuutta, asianmukaisuutta ja tarkkuutta toimialakohtaisissa tilanteissa (Liga & Robaldo 2023; Akhondi-Asl et al. 2024). Toimialakohtaisilla data-aineistoilla, kuten lakitekstillä tai lääketieteellisillä dokumenteilla koulutetut mallit kykenevät tuottamaan parempaa toimialakohtaista tekstiä (Liga &

Robaldo 2023; Yenduri et al. 2023). Olemassa olevan pohjamallin (*engl. foundation model*) hienosäätö on myös edullisempaa kuin kielimallin kouluttaminen alusta alkaen (Linkon et al. 2024). Mallin oppimisnopeutta ja koulutusjaksojen määrää muuttamalla käyttäjällä on mahdollisuus vaikuttaa siihen, kuinka merkittävästi hienosäätöprosessi vaikuttaa esikoulutetun mallin käyttöön (Kishore 2024).

Organisaatioiden tietoresurssien integroiminen generatiivisiin tekoälymalleihin voi tuottaa haasteita. Tietosuojahaasteet, esimerkiksi aran potilasdatan tai yrityksen sisäisen liiketoimintatiedon käyttö voivat rajoittaa saatavissa olevan tiedon määrää (Kishore 2024). Arkaluontoista tietoa, kuten potilastietoja, ei välttämättä haluta sisällyttää malleihin itsessään, sillä tieto olisi tällöin kaikkien mallin käyttäjien saatavissa. Eri organisaatioiden data-aineistojen muodot ja rakenteet sekä tietokannat ja -järjestelmät voivat olla keskenään hyvin erilaisia, mikä voi osaltaan vaikeuttaa toimialakohtaisen tiedon integroitavuutta (Kishore 2024). Datan laatu voi myös olla heikkoa (Kishore 2024).

Toinen ratkaisu täydentää mallin tietämystä onkin hyödyntää RAG:ia (*engl. Retrieval Augmented Generation*) (Li et al. 2024). RAG:in tarkoitus hyödyntää mallin kontekstissa oppimisen kyvykkyyttä (*engl. in-context learning*). RAG:n toimintaperiaate pohjautuu tekstinoutajaan, joka etsii relevantin sisällön annetusta toimialakohtaisesta tekstikorpuksesta. Noudettu sisältö voi olla esimerkiksi organisaation dokumentteja. Noudettu sisältö lisätään kielimallin syötteeseen parantaakseen sen ymmärrystä käsiteltävästä aiheesta. (Li et al. 2024) RAG:in toimintaperiaate on esitelty kuvassa 6.



**Kuva 6:** RAG-prosessi (mukaillen Li et al. 2024)

RAG:ien tekstinoutajat pohjautuvat kuitenkin täydellisiin osumiin tai semanttiseen samankaltaisuuteen. Niillä ei ole niin sanotusti päättelykykyä, joka voi johtaa siihen, että

ne eivät välttämättä hae dokumentteja, jotka täysin vastaavat haluttua kyselyä. Noudetut dokumentit voivat sisältää epärelevanttia tai harhaanjohtavaa tietoa. (Li et al. 2024)

Tutkimukset esittävät myös toimialakohtaisten kielimallien kouluttamista (Wu et al. 2023; Li et al. 2024). Pienet toimialakohtaiset kielimallit voivat toimia itsenäisesti, mutta niistä puuttuu esimerkiksi suurten kielimallien päättelykyvykkyudet (Li et al. 2024). Suurten toimialakohtaisten kouluttaminen on mahdollista (Wu et al. 2023), mutta kuluttaa myös paljon resursseja (Linkon et al. 2024). Li et al. (2024) esittää yhtenä vaihtoehtona pienten toimialakohtaisten kielimallien hyödyntämisen esikoulutettujen suurten kielimallien rinnalla. Kyseisessä BLADE-kehityksessä (*Black-box LARge language models with small Domain-specific models*) pientä toimialakohtaista kielimallia käytetään luomaan tietoa liittyen käyttäjän syötteeseen. Suuri kielimalli taas syntetisoi tämän tiedon pohjalta kattavan vastauksen. BLADE:n todettiin parantavan pohjamallien suoritusta merkittävästi toimialakohtaisissa kysymyksissä. (Li et al. 2024)

Organisaatioille on myös olemassa valmiita, tuotteistettuja ratkaisuja liiketoimintatiedon hyödyntämiseen generatiivisten tekoälymallien kanssa. Esimerkiksi Microsoftin Copilot for Microsoft 365 voi hyödyntää Office 365-ympäristöön lisättyjä tiedostoja, kuten Word-dokumentteja ja Excel-taulukoita, sähköposteja tai esimerkiksi Teams-palaverien nauhoitteita. Copilot for Microsoft 365 kykenee esimerkiksi tiivistämään dokumenttien sisältöä, etsimään ja kokoamaan tietoa sekä luomaan uusia dokumentteja. (Microsoft 2024) Googlen Gemini for Google Workspace tarjoaa vastaavia kyvykkyksiä Googlen omassa pilviympäristössä (Google 2024). Googlen (2024) mukaan Gemini voi auttaa havaitsemaan trendejä, koostamaan tietoja ja tunnistamaan liiketoimintamahdollisuuksia. IBM:n WatsonX on organisaatioille tarkoitettu tekoäly- ja data-alusta. Alusta mahdollistaa uusien mallien luomisen ja kouluttamisen sekä olemassa olevien mallien hienosäädön ja käyttöön ottamisen. (IBM 2024a) WatsonX tarjoaa myös datavaraston, jonne organisaatiot voivat lisätä omia, vapaamuotoisia dokumentteja. Datavarastoon lisättyjä dokumentteja voidaan myöhemmin hyödyntää esimerkiksi RAG-prosessissa. (IBM 2024b)

## 5. LIIKETOIMINTATIEDON HALLINTA JA GENERATIIVINEN TEKOÄLY

Tässä luvussa yhdistetään tutkimusaineiston avulla liiketoimintatieto, sen jalostuminen ja elinkaari edellisessä luvussa esiteltyyn generatiiviseen tekoälyyn. Luvun tavoitteena on tunnistaa miten generatiivinen tekoäly tukee ihmislähtöistä tiedon jalostamista sekä tunnistaa generatiivisen tekoälyn sovellukset liiketoimintatiedon hallinnassa ja niihin liittyvät haasteet, vastaten taulukossa 4 esiteltyihin alatutkimuskysymyksiin.

**Taulukko 4.** Tutkimuksen pää- ja alatutkimuskysymykset

<b>Päätutkimuskysymys</b>	Miten generatiivista tekoälyä kannattaa hyödyntää liiketoimintatiedon hallinnan tukena?
<b>Alatutkimuskysymys 1</b>	Miten generatiivinen tekoäly voi tukea ihmislähtöistä tiedon jalostamista?
<b>Alatutkimuskysymys 2</b>	Mitä sovelluksia generatiiviselle tekoälylle on liiketoimintatiedon hallinnan näkökulmasta?
<b>Alatutkimuskysymys 3</b>	Mitä haasteita ja rajoitteita generatiivisen tekoälyn käyttöön liiketoimintatiedon hallinnassa liittyy?

### 5.1 Tiedon jalostuminen generatiivisen tekoälyn avulla

Luvun 3 mukaisesti liiketoimintatiedon hallinnan tehtävänä on jalostaa tietoa päätöksen tueksi. Tiedon jalostumista pidetään sosiaalisissa vuorovaikutuksissa sekä yksilöllisissä kognitiivisissa prosesseissa tapahtuvana prosessina (Alavi et al. 2024). Generatiivista tekoälyä ei voidakaan perinteisen määritelmän mukaan pitää tietoa hyödyntävänä toimijana (Alavi et al. 2024), mutta sen kykyä tukea ihmislähtöistä tiedon jalostamista voidaan tarkastella tarkemmin.

Verrattuna perinteisiin määritelmään tiedon jaottelusta dataksi, informaatioksi ja tietämykseksi, generatiivinen tekoäly ikään kuin tiivistää informaatiotason ja luo tietämystä suoraan suuren datamäärän prosessoinnin ansiosta (Alavi et al. 2024). Verrattuna perinteiseen liiketoiminta-analytiikkaan, generatiivinen tekoäly pystyy hyödyntämään datasta löydettyjä säännönmukaisuuksia uuden tiedon luomiseksi (Nazeer et al. 2023).

Sumbal & Amber (2024) mukaan ChatGPT:tä voi käyttää tiedon luomiseen SECI-mallia mukaillen. ChatGPT:llä on erinomainen kyvykyys yhdistää ja sisäistää uutta tietoa vanhan tiedon kanssa. GPT-mallien luonteen vuoksi ChatGPT pystyy sisäistämään valtavan määrän tietoa ja päivittämään sen perusteella omaa tietokantaansa. (Sumbal &

Amber 2024) Myös Alavi et al. (2024) mukaan generatiivinen tekoäly tukee erityisesti tiedon yhdistämisen- ja sisäistämisen-vaiheita SECI-mallissa.

Generatiivisen tekoälyn kyky löytää säännönmukaisuuksia, prosessoida valtavia määriä dataa, tarjota näkemyksiä ja ennustaa lopputuloksia mahdollistaa tietämyksen tarjoamisen organisaatioille suoraan matalamman tason informaation sijasta (Alavi et al. 2024). Voidaan siis todeta, että generatiivinen tekoäly kykenee sisäistämään suuria määriä tietoa ja tarjota näistä jo tietylle asteelle jalostettuja näkemyksiä, täten tukien tiedon jalostamista organisaatioiden kontekstissa.

## **5.2 Liiketoimintatiedon hallinnan sovellukset**

Generatiivisella tekoälyllä on potentiaalia parantaa ja muuttaa tiedonhallinnan prosesseja, kuten tiedon luomista, varastoimista, jakamista ja hyödyntämistä (Alavi et al. 2024). Generatiivinen tekoäly voi auttaa organisaatioita esimerkiksi automatisoimalla korkean tason tiedon luomista laajasta määrästä rakenteetonta dataa, tehostamalla tiedon jakamista organisaatiossa luomalla, tiivistämällä ja levittämällä sisältöä ja tarjoamalla tietoa henkilökohtaisesti yksilöiden tietotarpeiden ja mieltymysten mukaisesti (Feuerriegel et al. 2023). Tarkastellaan seuraavaksi liiketoimintatiedon hallinnan prosessimallin eri vaiheita yksitellen generatiivisen tekoälyn näkökulmasta.

### **Tietotarpeiden määrittely**

Generatiivinen tekoäly voi tukea ideoiden luomista liiketoiminnan ja organisaation kontekstissa (Feuerriegel et al. 2023; Korzynski et al. 2023; Nazeer et al. 2023). Generatiivinen tekoäly voi tunnistaa ihmisille vaikeasti hahmotettavia näkemyksiä, jotka voivat tukea työntekijöiden ideointia (Benbya et al. 2024). Yksi esimerkki on uusien tuoteyhdistelmien tunnistaminen laajasta määrästä rakenteetonta dataa (Feuerriegel et al. 2023). Voidaan todeta, että generatiivinen tekoäly voi auttaa muun muassa määrittelemään tietotarpeita organisaation kontekstissa.

Generatiivinen tekoäly voi myös suositella käyttäjälleen sopivaa tietoa perustuen käyttäjän aiempaan osaamiseen. Tämä voi olla esimerkiksi suosituksia eri koulutusmateriaaleista. (Feuerriegel et al. 2023). Tällöin generatiivinen tekoäly voi auttaa määrittelemään tietotarpeita myös käyttäjän kontekstin mukaisesti.

### **Tiedon hankinta**

Generatiivista tekoälyä voidaan hyödyntää organisaationaallisen tiedon jaottelamiseen, suodattamiseen ja analysointiin suurissa volyymeissa (Alavi et al. 2024; Vadari & Malladi 2024). Generatiivinen tekoäly kykenee muuttamaan rakenteetonta tekstiä

rakenteelliseen dataformaattiin (Wei et al. 2023), jolloin se on helpommin organisaation hyödynnettävissä. Generatiivinen tekoäly kykenee tallentamaan tietoa niin sisäisistä kuin ulkoisistakin lähteistä (Alavi et al. 2024; Vadari & Malladi 2024). Käyttäjän ei välttämättä tarvitse tietää mistä tietoa tulisi hankkia, sillä tieto on sisällytetty malliin itsessään (Vadari & Malladi 2024). Luonnollisen kielen hyödyntäminen auttaa etenkin teknisesti orientoitumattomia löytämään tietoa ja muodostamaan siitä näkemyksiä (Rane 2023).

Syötteiden kautta käyttäjät pääsevät helposti ja nopeasti käsiksi mallin sisältämään tietoon. Tällöin tieto on aina käyttäjien saatavissa. (Alavi et al. 2024) Generatiivisen tekoälyn kyvyt ymmärtää käyttäjien syötteitä lähenevät ihmiskuuntelijan kyvykkyyksiä, jolloin se kykenee tuottamaan aiempia algoritmisia ratkaisuja parempia hakutuloksia (Benbya et al. 2024). Generatiivisen tekoälyn intuitiivinen käyttöliittymä mahdollistaa entistä paremmin oikean tiedon saamisen oikeaan aikaan (Vadari & Malladi 2024).

Organisaation linjausten ja työkalujen salliessa generatiiviset tekoälymallit voivat oppia käyttäjien syötteistä uutta tietoa, joka ei välttämättä sisälly niiden koulutusdataan (Alavi et al. 2024; Benbya et al. 2024; Sumbal & Amber 2024). Voidaan puhua organisaation hiljaisen tiedon eksplikoinnista tai SECI-mallin mukaisesti tiedon ulkoistamisesta (Sumbal & Amber 2024). Generatiiviset tekoälymallit ovat luonteensa vuoksi erinomaisia kodifioimaan hiljaista ja rakenteetonta tietoa (Magnier-Watanabe & Senoo 2009, Sumbal & Amber 2024 mukaan). Käyttäjiltä voidaan oppia esimerkiksi kokemuksia, parhaita käytäntöjä sekä virheistä saatuja oppeja ja näitä voidaan taas edelleen hyödyntää muiden käyttäjien syötteisiin vastattaessa (Sumbal & Amber 2024). Tämä mahdollistaa organisaation tietopohjan laajentamisen ja hyödyntämisen täysin uusilla tavoilla (Benbya et al. 2024). Generatiivinen tekoäly voi myös auttaa organisaatioita kartoittamaan osaamisverkostoja, jotka voivat auttaa tiedon löytämisessä (Bendya et al. 2024).

Lisäksi generatiivista tekoälyä voi hyödyntää keinotekoisien datan luomiseen (*engl. data augmentation*). Generatiivinen tekoäly pystyy esimerkiksi parantamaan datan laatua ja määrää tilanteissa, joissa alkuperäisen datan määrä tai sen laatu on heikkoa. (Liu et al. 2023) Generatiivista tekoälyä voi myös käyttää täysin synteettisten datasettien luomiseen, esimerkiksi muiden tekoälymallien kouluttamista varten (Alavi et al. 2024).

Voidaan todeta, että generatiivinen tekoäly kykenee tukemaan tiedon hankintaa usein eri tavoin. Näitä tapoja ovat esimerkiksi organisaation sisäisen ja ulkoisen tiedon sisäistäminen, organisaation hiljaisen tiedon kerääminen, luonnollisen kielen kautta suoritettut tietohaut sekä datan täydentäminen ja sen syntetisointi.

## **Tiedon prosessointi ja analysointi**

Generatiivinen tekoäly kykenee käsittelemään ja sisäistämään suuria määriä eri muodoissa olevaa organisaation sisäistä ja ulkoista dataa (Alavi et al. 2024). Generatiivinen tekoäly voi auttaa organisaatioita perinteisissä data-analytiikan tehtävissä (Rane 2023). Verrattuna perinteiseen liiketoiminta-analytiikkaan, generatiivinen tekoäly pystyy hyödyntämään datasta löydettyjä säännönmukaisuuksia uuden tiedon luomiseksi (Nazeer et al. 2023). Alaluvun 5.1 mukaisesti generatiivinen tekoäly kykenee yhdistelemään tiedonpalasia sekä etsimään niille asiayhteyksiä ja merkityksiä organisaation kontekstissa, luoden merkityksellistä, korkeamman tason tietoa. Lisäksi suuren datamääränsä ja prosessointikykyjensä ansiosta generatiivinen tekoäly kykenee luomaan ihmisille vaikeasti hahmotettavissa olevia näkemyksiä (Benbya et al. 2024). Generatiivinen tekoäly voi myös luoda organisaation dataan pohjautuvia tietotuotteita automaattisesti (Feuerriegel et al. 2023).

Verrattuna perinteiseen organisaation tietohallintoon, joka kykenee tarjoamaan suhteellisen rajatun määrän eksplisiittistä tietoa, generatiivinen tekoäly voi aiemmin mainittujen ominaisuuksiensa ansiosta toimia tiedon yhteisluojana (Alavi et al. 2024). Generatiivisen tekoälyn roolista tiedon yhteisluojana on näyttöä esimerkiksi lääke- ja materiaaliteollisuudessa (Alavi et al. 2024; Benbya et al. 2024). Generatiivista tekoälyä voidaan hyödyntää esimerkiksi tieteellisen kirjallisuuden, patenttien ja tietokantojen analysointiin ja tämän pohjalta uusien kemiallisten yhtisteiden tai materiaalien ehdottamiseen (Benbya et al. 2024).

Voidaan todeta, että generatiivinen tekoäly voi tukea tiedon prosessointia ja analysointia. Generatiivinen tekoäly voi tukea tätä prosessia niin perinteisin menetelmin, kuten datan analysoinnilla, kuin myös uusilla tavoilla, kuten korkean tason näkemysten luomisella ja tiedon yhteisluojana toimimisella. Generatiivinen tekoäly mahdollistaa myös osaltaan tiedon prosessoinnin ja analysoinnin automatisoinnin.

## **Tiedon jakaminen**

Generatiivinen tekoäly kykenee tarjoamaan käyttäjille tietoa kattavasti ja nopeasti (Alavi et al. 2024). Generatiivinen tekoäly voidaan kouluttaa organisaation datalla ja tällöin käyttäjät voivat hyödyntää tekoälyä chatbottimaisesti vastaamaan syötteisiin organisaation tietoon pohjautuen (Sumbal & Amber 2024). Generatiivinen tekoäly voi auttaa tiedon löytämistä esimerkiksi luomalla organisaation dataan pohjautuvia monikielisiä tietotuotteita automaattisesti (Feuerriegel et al. 2023). Generatiivinen tekoäly mahdollistaa tiedon esittämisen eri kielillä, jolloin se myös mahdollistaa tiedon jakamisen laajemmalle ihmisryhmälle (Vadari & Malladi 2024).

Generatiivinen tekoäly kykenee tarjoamaan tietoa henkilökohtaisesti räätälöidyillä tavoilla (Feuerriegel et al. 2023; Alavi et al. 2024; Vadari & Malladi 2024). Generatiivinen tekoäly voi suositella käyttäjän tilanteeseen sopivaa tietoa perustuen käyttäjän aiempaan työkokemukseen ja tietämykseen (Feuerriegel et al. 2023). Personoitu tiedon tarjoaminen voi tehostaa käyttäjien oppimista (Bendya et al. 2024). Tiedon hankkiminen generatiivisilta tekoälymalleilta voi madaltaa kommunikaatioon liittyviä muureja, verrattuna esimerkiksi asioiden kysymiseen esihenkilöiltä tai kollegoilta. Tämä voi helpottaa tiedon löytämistä ja tehostaa oppimista. (Alavi et al. 2024)

Generatiivinen tekoäly voi myös tukea hiljaisen tiedon jakamista siirtämällä esimerkiksi käyttäjien syötteistä ja organisaation dokumenteista opittua tietoa toisille käyttäjille (Alavi et al. 2024; Benbya et al. 2024). Korzynski et al. (2023) mukaan generatiivinen tekoäly voi toimia alustana, jolla organisaation tiimien jäsenet voivat jakaa ja vaihtaa hiljaista tietoa keskenään. Generatiivinen tekoäly voi myös tukea työntekijöiden välistä tiedon jakamista kartoittamalla organisaation sisäisiä osaamisverkostoja (Bendya et al. 2024). Voidaan todeta, että generatiivinen tekoäly voi tukea tiedon jakamista. Keinoja tiedon jakamiseen on esimerkiksi automaattinen tietotuotteiden luominen, tiedon kääntäminen eri kielille, henkilökohtaiset tavat toimittaa tietoa sekä hiljaisen tiedon siirtämisen tehostaminen.

### **Tiedon hyödyntäminen ja palaute**

Generatiivinen tekoäly tarjoaa pääsyn laajaan tietomäärään intuitiivisesti luontaisen kielien kautta (Alavi et al. 2024; Vadari & Malladi 2024). Generatiiviset tekoälymallit kykenevät ymmärtämään käyttäjiensä konteksteja ja tarjoamaan niihin sopivia vastauksia nopeasti (Alavi et al. 2024; Bendya et al. 2024; Vadari & Malladi 2024). Tämä helpottaa tiedon saamista päätöksentekoon etenkin teknisesti orientoitumattomille (Vadari & Malladi 2024). Lisäksi generatiivisen tekoälyn on todettu tehostavan tietotyötä (Alavi et al. 2024; Bendya et al. 2024).

Generatiivinen tekoäly voi auttaa tiedon tulkitsemisessa. Generatiivista tekoälyä on esimerkiksi käytetty kuvaamaan liiketoimintaprosesseja abstraktilla tasolla (Fill et al. 2023, Feuerriegel et al. 2023 mukaan). Nazeer et al. (2023) mukaan generatiivinen tekoäly kykenee esittämään suuria määriä dataa rakenteellisesti ja selkeästi, mahdollistaen paremman ymmärryksen antamisen tiedon käyttäjälle. Generatiivinen tekoäly auttaa erityisesti vähäisen työkokemuksen omaavia käsittelemään tietoa. Esimerkiksi rahoitusallalla generatiivinen tekoäly voi auttaa aloittelevia analyytikoita tekemään tarkkoja ennusteita ja antamaan parempia suosituksia. (Bendya et al. 2024) Lisäksi generatiivinen tekoäly

voi auttaa tiedon esittämisessä käyttäjän mieltymysten ja kokemuksen mukaan (Feuerriegel et al. 2023; Nazeer et al. 2023; Alavi et al. 2024), jolloin tietoa voi olla mielekkäämpää ja helpompaa hyödyntää.

Lisäksi generatiivinen tekoäly ei anna pelkästään vastauksia käyttäjien kysymyksiin, vaan se voi myös antaa käyttäjilleen palautetta ja kehoitteita. Tähän näkemykseen pohjautuen generatiivista tekoälyä voidaan pitää tiedon ja arvon yhteisluojana, joka voi omaksua erilaisia rooleja yhteistyössä. (Feurriegel et al. 2023) Nazeer et al. (2023) korostavat myös generatiivisen tekoälyn roolia tiedon yhteisluojana.

Voidaan todeta, että generatiivinen tekoäly voi tukea tiedon hyödyntämistä ja palautetta. Generatiiviset tekoälymallit mahdollistavat pääsyn laajaan tietomäärään luontaisen kielien kautta ja ne pystyvät esittämään tietoa helposti ymmärrettävässä muodossa henkilökohtaisin tavoin. Generatiivinen tekoäly kykenee myös antamaan palautetta ja kehoitteita, omaksuen samalla yhteistyöllisen roolin päätöksenteossa.

### **5.3 Generatiivisen tekoälyn haasteet liiketoimintatiedon hallinnassa**

Edellisissä alaluvuissa tarkasteltiin tiedon jalostumista generatiivisen tekoälyn avulla sekä sen mahdollisia liiketoimintatiedon hallinnan sovelluksia. Kokonais kuvan saamiseksi tässä alaluvussa tarkastellaan generatiiviseen liiketoimintatiedon hallintaan liittyviä haasteita.

#### **Tiedon laatu**

Generatiiviset tekoälymallit voivat tuottaa harhaanjohtavia ja virheellisiä tuloksia. Mallit tuottavat todennäköisimmän tuloksen, joka ei välttämättä ole faktuaalisesti oikea. Generoitu teksti on kielipillisesti oikeaa, mutta se voi samaan aikaan olla järjetöntä ja asiavirheellistä. Toisin sanoen generatiivinen tekoäly tuottaa sisältöä, joka ei ole faktoihin perustuvaa vaan enemmänkin mallin omiin oletuksiin ja ennakoasenteisiin perustuvaa. (Feuerriegel et al. 2023) Generatiivisten tekoälymallien laskentatehokkuuden parantamiseksi rajoitettu konteksti myös rajoittaa niiden kykyä käsitellä monimutkaisia ilmiöitä ja kielellisiä yksityiskohtia, joka voi osaltaan johtaa harhaanjohtaviin tai epärelevantteihin vastauksiin (Rane 2023). Alavi et al. (2024) tiivistää asian seuraavasti: generatiivinen tekoäly kykenee tarjoamaan tietoa kattavasti ja nopeasti, mutta sen tarjoaman tiedon laatu voi olla kyseenalaista.

Generatiivisen tekoälyn vastausten oikeellisuus riippuu suuresti koulutusdatan laadusta ja koulutusprosessista (Feuerriegel et al. 2023; Vadari & Malladi 2024). Suurimpien kaupallisten kielimallien käyttämään koulutusdataan ja niiden sisältämään tietoon on

vaikea paneutua syvällisesti (Feuerriegel et al. 2023). Malleilta voi kysyä perusteluja ja lähteitä väitteisiin, mutta nämäkin pohjautuvat mallien oletuksiin ja tilastoihin oikeiden faktojen sijasta ja täten voivat sisältää asiavirheitä (Feuerriegel et al. 2023; Liu et al. 2023). Lisäksi mallien koulutusdata rajoittuu ajallisesti tiettyyn pisteeseen asti, jolloin malleilla ei välttämättä ole uusinta tietoa eri aiheista (Liu et al. 2023). Organisaation tietoa hyödyntäessä on olemassa riski vanhentuneen tiedon, esimerkiksi vanhentuneiden käytänteiden noudattamisesta sekä innovaatiota rajoittavasta kehäpäättelystä (Benbya et al. 2024).

On myös huomattava, että yksilöt valitsevat usein helposti saatavan mutta heikomman laatuksen tietolähteen verrattuna vaikeammin saatavaan hyvälaatuiseen tietolähteeseen (Simon 1956, Alavi et al. 2024 mukaan). Generatiivinen tekoäly tarjoaa tietoa nopeasti ja kattavasti, mutta sen tarjoaman tiedon laatu voi olla kyseenalaista (Alavi et al. 2024). Sumbal & Amber (2024) havaitsivat, että käyttäjät pitävät ChatGPT:tä luotettavana tietolähteenä, ikään kuin henkilökohtaisena assistenttina. Bendya et al. (2024) mukaan onkin olemassa riski siitä, että työntekijät voivat tulla liian luottavaisiksi generatiivista tekoälyä kohtaan tietolähteenä, sivuuttaen muiden ihmisten tietämyksen ja osaamisen, joka on erittäin arvokasta monimutkaisissa ongelmissa. Alavi et al. (2024) mukaan generatiivisen tekoälyn tarjoaman tiedon helppo saatavuus saattaa vähentää sosiaalisia vuorovaikutuksia, joista tietoa on perinteisesti haettu. Nazeer et al. (2023) nostavat myös esille riskin liiasta luottavaisuudesta mallien vastauksiin. Ilman tiedon oikeellisuuden tarkistamista generatiivisen tekoälyn antamaa tietoa voidaan hyödyntää väärin (Alavi et al. 2024).

Feuerriegel et al. (2023) mukaan päätöksenteon kannalta tarvitaan algoritmisia ratkaisuja hallusinoinnin havaitsemiseksi ja lieventämiseksi. Myös käyttäjäkeskeisiä ratkaisuja tarvitaan tämän ongelman ratkaisemiseksi. Yksi esimerkki olisi kehittää uusia keinoja, joilla käyttäjä voi tarkistaa syntetisoidun sisällön oikeellisuuden. (Feuerriegel et al. 2023)

### **Tietoturva**

Generatiivisten tekoälysovellusten hyödyntäminen voi johtaa organisaatioiden tiedon ja liikesalaisuuksien vuotamiseen (Alavi et al. 2024). Esimerkiksi Samsungin työntekijät syöttivät ChatGPT:lle liikesalaisuuksia hyödyntäessään sitä työtehtävissä. OpenAI hyödyntää käyttäjien syötteitä koulutusdatana, jolloin näitä liikesalaisuuksia voidaan käyttää uusien mallien koulutuksessa. (Nazeer et al. 2023)

Generatiivinen tekoäly vähentää organisaatioiden mahdollisuuksia hallita työntekijöille saatavissa olevaa tietoa. Vaikkakin mallien hienosäätö on mahdollista, mallien

tarjoaman tiedon täydellinen hallinta on haastavaa. (Alavi et al. 2024) Generatiivista tekoälyä voidaan esimerkiksi huijata hakemaan tietoa organisaation sisällä rajoitetuista tietolähteistä (Renaud et al. 2023, Bendya et al. 2024 mukaan). Jopa käytettäessä organisaatiokohtaisia generatiivisia tekoälymalleja, joissa tieto pysyy organisaation sisällä, voi rajoitettu tieto päätyä muiden työntekijöiden saataville palautesilmukoiden kautta. Generatiivinen tekoäly asettaa haasteita esimerkiksi HR- tai R&D-yksiköille, jotka käsittelevät arkaluontaista dataa kuten henkilötietoja. (Bendya et al. 2024) Italian asettama kiello ChatGPT:lle viestii generatiivisten tekoälymallien tietoturvaan liittyvistä epävarmuuksista (Hacker et al. 2023, Bendya et al. 2024 mukaan).

### **Toimialakohtaisuuden haasteet**

Suurilla kielimalleilla ei välttämättä ole suoraan tarpeeksi kattavaa toimialakohtaista tietoa (Rane 2023). Tämä on oleellinen ongelma, sillä alakohtaisesti on tarvetta erilaisille korkean tason kontekstualisoinneille (Feuerriegel et al. 2023). Vaikka generatiivisen tekoälyn mukauttaminen eri käyttötarkoituksiin on mahdollista, on se edelleen haastavaa (Feuerriegel et al. 2023). Esimerkiksi mallien integraatio olemassa oleviin tietojärjestelmiin ja tietorakenteisiin voi tuottaa haasteita (Vadari & Malladi 2024). Accenturen mukaan organisaatioiden datakyvykkyydet eivät vielä vastaa generatiivisen tekoälyn vaatimuksia, joka jarruttaa sen käyttöönottoa organisaatioissa (Foley 2023). Malleja tulisi myös kouluttaa jatkuvasti, jotta niistä on organisaatioille eniten hyötyä. Tämä tuo organisaatioille omat haasteensa mukautettujen mallien hyödyntämisessä. (Rane 2023)

### **Eettiset haasteet**

Generatiivisen tekoälyn hyödyntäminen nostaa esille eettisiä haasteita, kuten misinformaation jakamisen, yksityisyydensuojan, tasapuolisuuden, läpinäkyvyyden, tekijänoikeuksien noudattamisen ja ennakkoasenteet eri ihmisryhmiä kohtaan (Vadari & Malladi 2024). Generatiivisella tekoälyllä on havaittu olevan ennakkoasenteita esimerkiksi eri ihmisryhmiin liittyen (Nazeer et al. 2023). ChatGPT ei ole myöskään poliittisesti täysin neutraali (Hartmann et al. 2023, Liu et al. 2023 mukaan). Generatiivisen tekoälyn näkemykset eri ihmisryhmistä, kulttuureista, uskonnoista heijastuvat sen koulutusdatasta (Feuerriegel et al. 2023; Liu et al. 2023; Rane 2023; Vadari & Malladi 2024). Myös suurten kielimallien kouluttamiseen usein käytetty ihmisohjattu vahvistusoppiminen voi luoda malleille ennakkoasenteita (Feuerriegel et al. 2023).

ChatGPT on antanut moraalisesti eriäviä vastauksia toistuviin kysymyksiin vaunuongelmista (*engl. trolley problem*). ChatGPT:n epäohdonmukaisuus moraalisisa

ongelmissa voi vaikuttaa ihmisten moraaliin päätöksenteossa. (Liu et al. 2023) ChatGPT on osoittautunut epäjohdonmukaiseksi myös järkeilyssä, tiedon oikeellisuudessa, matematiikassa, koodaamisessa ja ennakoasenteissa (Borji 2023, Liu et al. 2023 mukaan). Mallien vastausten vaihtelevuus voi heikentää niiden luotettavuutta päätöksenteossa. Tämä nostaa kysymyksen generatiivisen tekoälyn moraalista vastuusta päätöksenteossa (Krügel et al. 2023, Feuerriegel et al. 2023 mukaan).

Suurimmat generatiiviset tekoälymallit toimivat ikään kuin mustina laatikkoina, jolloin niiden niin sanottu ajatusprosessi jää epäselväksi (Nazeer et al. 2023; Alavi et al. 2024). Tällöin esimerkiksi mallin vastausten oikeellisuuden ja järkevyyden tarkastaminen on haastavampaa ja jää käyttäjän vastuulle (Nazeer et al. 2023). Tämä läpinäkyvyyden puute voi koitua haasteeksi päätöksenteon siirtyessä entistä enemmän generatiivisille tekoälymalleille (Alavi et al. 2024).

Generatiiviset tekoälymallit voivat luoda uutta sisältöä joka muistuttaa tai on kopio jostain tekijänoikeudella suojatusta sisällöstä tai jopa patentista, ilman korvausta alkuperäisen tekijänoikeuden omistajalle (Feuerriegel et al. 2023; Alavi et al. 2024). Tästä syystä koulutusdatan tulisi olla tekijänoikeuksista vapaata (Feuerriegel et al. 2023). Aiemmin mainittu läpinäkyvyyden haaste suurimissa kielimalleissa haastaa myös koulutusdatan tarkistamista tekijänoikeuksien kannalta. Esimerkiksi tiukasti patenteilla kilpailevassa teknologiateollisuudessa tekijänoikeusloukkaukset voivat johtaa suuriin oikeudenkäynteihin ja korvaustenhakuihin (Feuerriegel et al. 2023), jolloin tekijänoikeuksien ja patenttien tarkistaminen on oleellinen haaste generatiivisten tekoälymallien hyödyntämisessä..

Useat aiemmin mainituista haasteista liittyvät erityisesti mallien käyttämään koulutusdataan ja sen laatuun. Tämän vuoksi koulutusdatan laadu ja sen käsittelyprosessien tulee olla hyvällä tasolla (Feuerriegel et al. 2023). Mallien neutraalisuuden varten on tärkeää luoda selkeät ohjelmointisäännöt sekä tarkastaa mallin laatua koulutuksen eri vaiheissa. Lisäksi syötteitä muokkaamalla ”konepellin alla” on mahdollista saada tasapuolisempia vastauksia. (Feuerriegel et al. 2023)

## 6. YHTEENVETO

Tutkimuksen tavoitteena oli luoda ymmärrystä generatiivisen tekoälyn roolista liiketoimintatiedon hallinnassa. Tutkimuksessa analysoitiin mitä sovelluksia generatiiviselle tekoälylle on liiketoimintatiedon hallinnassa ja mitä haasteita niihin liittyy. Lisäksi tutkimuksessa tarkasteltiin sitä, miten generatiivinen tekoäly nivoutuu tiedon elinkaareen liiketoimintatiedon hallinnan prosessimallin mukaisesti. Tässä luvussa esitetään tiivistetysti tutkimuksen keskeisimmät tulokset ja päätelmät. Lisäksi luvussa arvioidaan tutkimuksen onnistumista. Lopuksi esitetään mahdolliset jatkotutkimusehdotukset, jotka voivat edistää generatiivisen tekoälyn ja liiketoimintatiedon hallinnan tutkimusta entisestään.

### 6.1 Tulokset ja päätelmät

Tutkimuksessa tarkasteltiin tiedon jalostumista generatiivisen tekoälyn avulla sekä generatiivisen tekoälyn sovelluksia liiketoimintatiedon hallinnan prosessin eri vaiheissa. Ensimmäinen alatutkimuskysymys käsitteli sitä, miten generatiivinen tekoäly voi tukea ihmislähtöistä tiedon jalostamista. Tutkimuksen mukaan generatiivinen tekoäly haastaa tiedon perinteisen tiedon jaottelun dataksi, informaatioksi ja tietämykseksi, sillä generatiivinen tekoäly ikään kuin tiivistää informaatiotason ja luo tietämystä suoraan suuren datamäärän prosessoinnin ansiosta. Generatiivinen tekoäly on erinomainen havaitsemaan säännönmukaisuuksia ja kykenee jalostamaan tietoa SECI-mallin mukaisesti, erityisesti sitä yhdistelemällä ja sisäistämällä. Generatiivisen tekoälyn kyky löytää säännönmukaisuuksia, prosessoida valtavia määriä dataa, tarjota näkemyksiä ja ennustaa lopputuloksia mahdollistaakin tietämyksen tarjoamisen organisaatioille suoraan matalamman tason informaation sijasta.

Tutkimuksen toisen, liiketoimintatiedon hallinnan sovelluksiin liittyvän alatutkimuskysymyksen osalta generatiiviselle tekoälylle löydettiin useita eri käyttökohteita jokaiselle liiketoimintatiedon hallinnan prosessin vaiheelle. Generatiivisen tekoälyn liiketoimintatiedon hallintaan liittyvät sovellukset on koottu taulukkoon 5.

**Taulukko 5.** Generatiivisen tekoälyn liiketoimintatiedon hallintaan liittyvät sovellukset

Liiketoimintatiedon hallinnan prosessin vaihe	Generatiivisen tekoälyn sovellukset
<b>Tietotarpeiden määrittely</b>	Generatiivinen tekoäly voi tukea tietotarpeiden ideoimista ja määrittelyä esimerkiksi tunnistamalla ihmisille vaikeasti hahmotettavasti olevia näkemyksiä. Kontekstuaalisen ymmärtämisen kautta generatiivinen tekoäly kykenee tunnistamaan ja määrittelemään erilaisiin ja yksilöllisiin tilanteisiin mukautettuja tietotarpeita.
<b>Tiedon hankinta</b>	Generatiivinen tekoäly voi prosessoida suuren määrän rakenteetonta tietoa ja muuttaa sen helposti hyödynnettävään muotoon. Käyttäjän ei tarvitse tietää mistä hakea tietoa, sillä tieto on sisällytetty malliin itsessään. Luonnollisen kielen käyttöliittymä voi erityisesti auttaa teknisesti orientoitumattomien tiedon hankintaa. Kontekstuaalinen ymmärtäminen voi tuottaa parempia hakutuloksia kuin perinteisillä algoritmeilla. Lisäksi generatiivista tekoälyä voidaan hyödyntää organisaation hiljaisen tiedon hankintaan, osaamisverkostojen karitoitukseen sekä keinotekoisien datan syntetisointiin.
<b>Tiedon prosessointi ja analysointi</b>	Generatiivinen tekoäly voi toteuttaa perinteisiä data-analytiikan prosesseja. Suuren prosessointikykyensä ansiosta se voi myös luoda näkemyksiä, jotka voivat olla ihmisille vaikeasti hahmotettavissa. Generatiivinen tekoäly kykenee muodostamaan valmiita tietotuotteita automaattisesti. Kattavan tietomääränsä ja kontekstuaalisen ymmärryksen ansiosta se voi myös toimia tiedon yhteisluojana.
<b>Tiedon jakaminen</b>	Generatiivinen tekoäly kykenee omaksumaan organisaation sisäistä tietoa ja täten mahdollistaa syötteiden kautta helpon käyttöliittymän sen käsittelyyn ja siirtämiseen. Generatiivinen tekoäly kykenee kääntämään tietoa eri kielille, mikä mahdollistaa tiedon jakamisen laajemmalle ihmisryhmälle. Generatiivinen tekoäly voi toimittaa tietoa henkilökohtaisesti räätälöidyillä tavoilla, mikä voi madaltaa tiedon hankintaan liittyviä muureja. Generatiivinen tekoäly voi myös tukea hiljaisen tiedon jakamista siirtämällä esimerkiksi käyttäjien syötteistä tai organisaation dokumenteista opittua tietoa toisille käyttäjille.
<b>Tiedon hyödyntäminen ja palaute</b>	Generatiivinen tekoäly helpottaa tietoon pääsyä luontaisen kielen käyttöliittymän kautta, helpottaen samalla tiedon hyödyntämistä oikeaan aikaan, etenkin teknisesti orientoitumattomille. Generatiivinen tekoäly kykenee ymmärtämään käyttäjänsä kontekstia, jolloin se kykenee tarjoamaan paremmin käyttäjän tietotarpeisiin vastaavaa tietoa. Generatiivinen tekoäly voi myös auttaa tiedon tulkitsemisessa, esittäen suuria datamääriä rakenteellisesti ja selkeästi. Generatiivinen tekoäly voi myös antaa käyttäjilleen palautetta ja kehoitteita, korostaen sen roolia tiedon ja arvon yhteisluojana.

Taulukosta 5 voidaan tunnistaa myös generatiivisen tekoälyn tärkeimmät kyvykkyydet liiketoimintatiedon hallinnan kannalta. Kyvykkyydet on tunnistettu toistuvuuksien perusteella ja tarvittaessa samankaltaiset kyvykkyydet on yhdistetty laajemmiksi kattotermeiksi. Näitä pääkyvykkyyksiä ovat käyttäjän kontekstin ymmärtäminen, monikielinen ja helppokäyttöinen luonnollisen kielen käyttöliittymä, suurten datamäärien prosessointikyky, organisaation tiedon sisäistämiskyky sekä kontekstissa oppimisen kyky.

Generatiivisen tekoälyn käyttöön liittyy myös liiketoimintatiedon hallintaan, tiedon jalostamiseen ja päätöksentekoon liittyviä haasteita. Tutkimuksessa tunnistettiin haasteita ja ne jaoteltiin seuraaviin kategorioihin: tiedon laatu, tietoturva, toimialakohtaisuus ja eettiset haasteet. Haasteet on koottu taulukkoon 6.

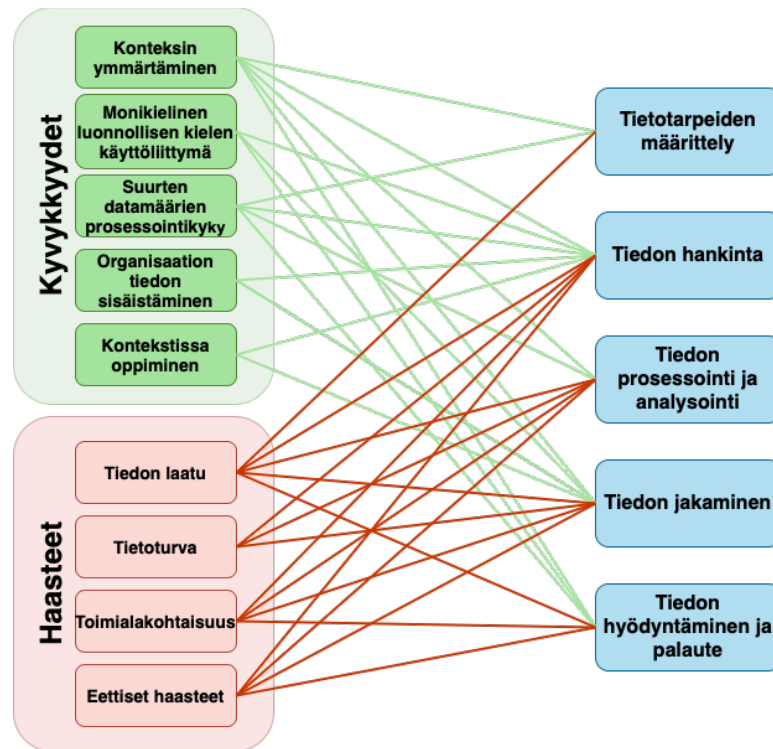
**Taulukko 6.** Generatiiviseen tekoälyyn liittyvät haasteet

Kategoria	Haasteet
<b>Tiedon laatu</b>	<ul style="list-style-type: none"> <li>- Hallusinaatio sekä harhaanjohtavat ja asiasisällöltään virheelliset tulokset</li> <li>- Koulutusdatan laatu</li> <li>- Koulutusdatan alkuperä</li> <li>- Vanhentunut tieto</li> <li>- Rajattu konteksti</li> <li>- Liiallinen luottamus generatiiviseen tekoälyyn tietolähteenä</li> </ul>
<b>Tietoturva</b>	<ul style="list-style-type: none"> <li>- Arkaluontaisen tiedon vuotaminen koulutusdataksi</li> <li>- Arkaluontoiseen tietoon pääsyn rajoittaminen on haastavaa</li> <li>- Jopa organisaatiokohtaisissa malleissa arkaluontoinen tieto voi levitä palautesilmukoiden kautta</li> </ul>
<b>Toimialakohtaisuus</b>	<ul style="list-style-type: none"> <li>- Suurien kielimallien toimialakohtainen tieto voi olla puutteellista</li> <li>- Mallien tietämyksen täydentäminen on teknisesti haastavaa</li> </ul>
<b>Eettiset haasteet</b>	<ul style="list-style-type: none"> <li>- Tasapuolisuus ja ennakoasenteet eri ihmisryhmiä kohtaan</li> <li>- Generatiiviset tekoälymallit eivät ole poliittisesti neutraaleja</li> <li>- Moraalisesti vaihtelevat vastaukset</li> <li>- Epäjohdonmukaisuus päättelyssä</li> <li>- Mallien tiedon ja "ajatusprosessin" läpinäkyvyys</li> <li>- Tekijänoikeuksien noudattaminen</li> <li>- Misinformaation jakaminen</li> <li>- Yksityisyydensuoja</li> </ul>

Taulukon 6 mukaisesti suurin osa generatiiviseen tekoälyyn liittyvistä haasteista liittyy suoraan mallien toimintaperiaatteeseen ja niiden koulutusdatan laatuun. Generatiivista tekoälyä hyödyntäessä tulisikin varmistua mallien antaman tiedon oikeellisuudesta sekä kiinnittää huomiota erityisesti organisaation sisäisen tiedon käsittelyyn generatiivisissa tekoälymalleissa. Päätöksenteossa mallien antamaa tietoa tulisi myös arvioida tasapuolisuuden linssin kautta, sillä generatiivisten tekoälymallien vastaukset voivat sisältää ennakoasenteita esimerkiksi eri ihmisryhmiä kohtaan.

Tutkimuksen tulosten avulla on mahdollistaa muodostaa kokonaiskuva generatiivisen tekoälyn hyödyntämiseen liiketoimintatiedon hallinnassa. Ottaen huomioon erityisesti tiedon oikeellisuuteen, rajattuun kontekstiin, päättelyn epäjohdonmukaisuuteen ja toimialakohtaisen tiedon puutteeseen liittyvät haasteet, generatiivisen tekoälyn hyödyntäminen tilanteissa, jotka vaativat useiden yksityiskohtien käsittelyä, syötteen ulkopuolista faktatietoa tai loogista päättelyä voi johtaa harhaanjohtaviin tuloksiin. Tasapuolisuuteen liittyvät haasteet asettavat myös haasteita eettiseen päätöksentekoon. Toisaalta generatiivinen tekoäly on erinomainen tiedon omaksumisessa, tiedon analysoinnissa ja ideoinnissa. Lisäksi sen kattava tietämyspohja mahdollistaa nopean mukautumisen erilaisiin

tehtäviin, jotka eivät ole liian toimialaspesifejä. Kokonaiskuvan saamiseksi generatiivisen tekoälyn liiketoimintatiedon hallintaan liittyvät kyvykkyydet ja haasteet on kytketty liiketoimintatiedon hallinnan prosessin vaiheisiin kuvassa 7.



**Kuva 7:** Generatiivisen tekoälyn kyvykkyydet ja haasteet liiketoimintatiedon hallinnan kontekstissa

Kuvan 7 mukaisesti generatiivisen tekoälyn kyvykkyydet tukevat eri liiketoimintatiedon hallinnan prosessin vaiheita. Toisaalta jokaiseen prosessin vaiheeseen liittyy myös haasteita. Generatiivinen tekoäly soveltuu hyvin esimerkiksi tietotarpeiden määrittelyyn, sillä tämän vaiheen tarkoituksena on uuden sisällön luominen ja ideoiminen absoluuttisten faktojen hakemisen sijasta. Hallusinaatiot voivat kuitenkin osaltaan vaikuttaa tähän prosessiin, luoden esimerkiksi sekavia tietotarpeita.

Tiedon hyödyntämisen ja palautteen kannalta generatiivisella tekoälyllä on potentiaalia auttaa käyttäjää esimerkiksi tulkitsemaan tietoa. Tiedon tulkitsemiseen voi kuitenkin liittyä niin laadullisia haasteita, kuin eettisiä haasteita, esimerkiksi epäjohtonmukaisuutta ja ennakoasenteita.

Tiedon hankinnan, tiedon prosessoinnin ja analysoinnin sekä tiedon jakamisen vaiheiden osalta generatiivinen tekoäly tarjoaa hyötyjä, mutta toisaalta myös enemmän haasteita kuin muihin vaiheisiin verrattuna. Tiedon hankinnan osalta generatiivinen tekoäly kykenee tarjoamaan tietoa nopeasti moneen eri tietotarpeeseen. Generatiivisen tekoälyn tarjoaman tiedon laatu voi olla kuitenkin kyseenalaista. Hankittu tieto voi myös sisältää

esimerkiksi tekijänoikeudella suojattua sisältöä. Vaihtoehtoisesti toimialakohtaisen tietämyksen puute voi johtaa harhaanjohtavaan vastauksiin.

Generatiivinen tekoäly soveltuu tiedon prosessointiin ja analysointiin varauksella, sillä se voi esimerkiksi jalostaa alustavia näkemyksiä suuresta datamäärästä suhteellisen nopeasti. Kuitenkin otettaessa tähän vaiheeseen linkittyvät haasteet huomioon, ei näitä tuloksia voida pitää absoluuttisina faktoina, vaan juuri alustavina lähtökohtina.

Tiedon jakamisen osalta generatiivinen tekoäly mahdollistaa esimerkiksi organisaationallisen tiedon omaksumisen ja levittämisen useilla eri kielillä. Generatiivisen tekoälyn toimintaperiaatteen vuoksi jaettava tieto voi kuitenkin sekoittua mallin olemassa olevan tiedon kanssa ja täten jopa muuttaa merkitystään. Kuten myös tiedon hankinnassa, jaetun tiedon laatu voi olla heikkoa. Mallit saattavat myös esimerkiksi levittää arkaluontoista tietoa.

Yleisesti katsottuna generatiivisen tekoälyn potentiaali liiketoimintatiedon hallinnassa on suuri. Organisaatioiden ymmärtäessä generatiiviseen tekoölyyn liittyvät sovellukset ja haasteet sekä luomalla raamit sen käyttöön omassa liiketoimintaympäristössään, on niillä mahdollisuus tehostaa omaa liiketoimintatiedon hallintaansa.

## 6.2 Arviointi ja jatkotutkimusehdotukset

Tutkimus toteutettiin kirjallisuuskatsauksena, jolloin lähdeaineiston merkitys oli tutkimuksen kannalta kriittinen. Aineiston löytäminen ei ollut helppoa, sillä aihepiiri on uusi ja tehtyä tutkimusta on olemassa vain vähän. Suurin osa tietokannoista löydetyistä materiaalista oli ei-tieteellistä sisältöä, esimerkiksi uutiskirjeitä tai kuluttajalehtiartikkeleita. Suoraan tutkimuskysymyksiin vastaavia artikkeleita ei löytynyt. Löydetyt artikkelit lähestyivät tutkimuksen aihepiiriä kuitenkin eri tiedonhallinnan osalueiden näkökulmista. Löydetyt tutkimukset olivat myös aineistoltaan kattavia ja ne tukivat toistensa väitteitä.

Aineiston vähäisyydestä huolimatta tutkimus tuotti tietoa generatiivisen tekoälyn hyödyntämisestä liiketoimintatiedon hallinnassa. Tutkimusaineiston kokoamisen ja yhdistelemisen ansiosta kaikkiin alatutkimuskysymyksiin oli mahdollista saada kattavat vastaukset. Tuloksista oli mahdollista myös tehdä päätelmiä, vastaten päättämiskysymykseen. Toisaalta olemassa olevan aineiston vähäisyys voi myös viestiä tutkimuksen uutuusarvosta. Tutkimuksen uutuusarvo syntyi tutkimusaineiston artikkelien eri näkökulmien kokoamisesta ja yhdistelemisestä juuri tämän työn näkökulmien kontekstissa. Tutkimuksen ajankohtaisuudesta kertoo myös se, että uusia, aihepiiriä sivuavia julkaisuja on tehty tutkimuksen aikana. Lisäksi tutkimus on toteutettu

Finkin (2014) mallin mukaisesti ja mahdollisimman läpinäkyvästi, jolloin tutkimus on toistettavissa ja sen tuottama tieto voi aidosti olla arvokasta.

Jatkotutkimuksessa tulisi erityisesti kiinnittää huomiota tutkimuksessa esitettyjen haasteiden ratkaisemiseen. Voisi olla hyödyllistä tutkia käytännön toteutusta tässä tutkimuksessa esitetyistä generatiivisen tekoälyn liiketoimintatiedon hallinnan sovelluksista erilaisissa liiketoimintaympäristöissä. Myös mainittujen haasteiden näkymistä käytännössä voisi olla hyödyllistä tutkia esimerkiksi case-tutkimuksella. Yksi huomionarvoinen tutkimuskohde voisi olla myös organisaation hiljaisen tiedon automaattinen kartoitus ja eksplikoiminen generatiivisen tekoälyn avulla.

# LÄHTEET

- Akhondi-Asl, A., Yang, Y., Luchette, M., Burns, J. P., Mehta, N. M. & Geva, A. (2024). Comparing the Quality of Domain-Specific Versus General Language Models for Artificial Intelligence-Generated Differential Diagnoses in PICU Patients. *Pediatric Critical Care Medicine*, 10-1097.
- Alavi, M., Leidner, D.E. & Mousavi, R. (2024). Knowledge Management Perspective of Generative Artificial Intelligence. *Journal of the Association for Information Systems*. Saatavissa: <https://aisel.aisnet.org/jais/vol25/iss1/15>
- Alghamadi, N.A. & Al-Baity H.H. (2022). Augmented Analytics Driven by AI: A Digital Transformation beyond Business Intelligence. *Sensors* 2022, Vol. 22, s. 8071. Saatavissa: <https://doi.org/10.3390/s22208071>
- Banh, L. & Strobel, G. (2023). Generative artificial intelligence. *Electronic markets*, Vol. 33(1), s. 1–17.
- Baskarada, S. & Koronios, A. (2013). Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and its Quality Dimension. *Australasian Journal of Information Systems*, 18(1). Saatavissa: <https://doi.org/10.3127/ajis.v18i1.748>.
- Bellinger, G., Castro, D. & Mills, A. (2004). Data, information, knowledge, and wisdom.
- Benbya, H., Strich, F. & Tamm, T. (2024). Navigating generative artificial intelligence promises and perils for knowledge and creative work. *Journal of the Association for Information Systems*, Vol. 25(1), s. 23–36. doi:<https://doi.org/10.17705/1jais.00861>
- Borji, A. (2023). A categorical archive of chatgpt failures. arXiv preprint arXiv:2302.03494.
- Brea, E. & Ford, J. A. (2023). No silver bullet: Cognitive technology does not lead to novelty in all firms. *Technovation*, Vol. 122, Article 102643.
- Bishop, C. (2006). *Pattern recognition and machine learning*. Springer, New York
- Deng, L. & Liu, Y. (2018). *Deep learning in natural language processing*. Springer Singapore Pte.
- Feuerriegel, S., Hartmann, J., Janiesch, C. & Zschech, P. (2024). Generative ai. *Business & Information Systems Engineering*, Vol. 66(1), s. 111–126.
- Fill, H.G., Fettke, P. & Köpke, J. (2023). Conceptual modeling and large language models: impressions from first experiments with ChatGPT. *EMISAJ*, Vol. 18(3), s. 1–15. <https://doi-org.libproxy.tuni.fi/10.18417/emisa.18.3>
- Fink, A. (2014). *Conducting research literature reviews: From the Internet to paper*. 4. ed. Sage, Thousand Oaks.
- Foley, É. & Guillemette, M. G. (2010). What is business intelligence?. *International Journal of Business Intelligence Research (IJBIR)*, Vol. 1(4), s. 1–28.
- Foley, S. (2023). Accenture chief executive says most companies are not ready for AI rollout. FT.Com.
- Google. (2024). Gemini for Google Workspace [verkkoaineisto]. Google. Viitattu 12.4.2024. Saatavissa: <https://workspace.google.com/solutions/ai/>
- Google Cloud Tech. (2023). Youtube-video. 22:07 minuuttia. Introduction to Generative AI. Saatavissa: <https://www.youtube.com/watch?v=G2fqAlgmoPo> (Viitattu 21.3.2024)

Hacker, P., Engel, A. & Mauer, M. (2023). Regulating ChatGPT and other large generative AI models. arXiv. <https://arxiv.org/abs/2302.02337>.

Hartmann, J., Schwenzow, J. & Witte, M. (2023). The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation. arXiv preprint arXiv:2301.01768.

Hu, K. (2023). ChatGPT sets record for fastest-growing user base – analyst note. Reuters (verkkoaineisto). Saatavissa (viitattu 5.2.2024): <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>

IBM. (2024a). watsonx. IBM (verkkoaineisto). Saatavissa (viitattu 12.4.2024): <https://www.ibm.com/watsonx>

IBM. (2024b). watsonx.data. IBM (verkkoaineisto). Saatavissa (viitattu 12.4.2024): <https://www.ibm.com/products/watsonx-data>

Iorliam, A. & Ingio, J. A. (2024). A Comparative Analysis of Generative Artificial Intelligence Tools for Natural Language Processing. *Journal of Computing Theories and Applications*, Vol. 2(1), s. 91–105.

Janiesch, C., Zschech, P. & Heinrich, K. (2021). Machine learning and deep learning. *Electronic markets*, Vol. 31(3), s. 685–695.

Kishore, Y. (2024). Optimizing Enterprise Conversational AI: Accelerating Response Accuracy with Custom Dataset Fine-Tuning. *Intelligent Information Management*, Vol. 16(2), s. 65–76.

Korzynski, P., Mazurek, G., Altmann, A., Ejdys, J., Kazlauskaite, R., Paliszkiwicz, J., et al. (2023). Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Central European Management Journal*, Vol. 31(1), s. 3–13.

Kühl, N., Schemmer, M., Goutier, M. & Satzger, G. (2022). Artificial intelligence and machine learning. *Electronic markets*, Vol. 32(4), s. 2235–2244.

Kämäräinen, J. (2023). Koneoppimisen perusteet. Otatieto.

Li, H., Ai, Q., Chen, J., Dong, Q., Wu, Z., Liu, Y., et al. (2024). Blade: Enhancing black-box large language models with small domain-specific models. arXiv preprint arXiv:2403.18365.

Liga, D. & Robaldo, L. (2023). Fine-tuning gpt-3 for legal rule classification. *Computer Law & Security Review*, 51, 105864.

Linkon, A. A., Shaima, M., Sarker, M. S. U., Nabi, N., Rana, M. N. U., Ghosh, S. K., et al. (2024). Advancements and Applications of Generative Artificial Intelligence and Large Language Models on Business Management: A Comprehensive Review. *Journal of Computer Science and Technology Studies*, Vol. 6(1), 225–232.

Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., et al. (2023). Summary of chatgpt-related research and perspective towards the future of large language models. *Meta-Radiology*, 100017.

Magnier-Watanabe, R. & Senoo, D. (2009). Congruent knowledge management behaviors as discriminate sources of competitive advantage. *Journal of Workplace Learning*, Vol. 21(2), s. 109–124.

Mandapuram, M., Gutlapalli, S. S., Bodepudi, A. & Reddy, M. (2018). Investigating the Prospects of Generative Artificial Intelligence. *Asian Journal of Humanity, Art and Literature*, Vol. 5(2), s. 167–174.

- Marr, B. (2019). The 10 Best Examples Of How AI Is Already Used In Our Everyday Life. Forbes (verkkoaineisto). Saatavissa (viitattu 21.4.2024): <https://www.forbes.com/sites/bernardmarr/2019/12/16/the-10-best-examples-of-how-ai-is-already-used-in-our-everyday-life/?sh=4be196561171>
- Martineu, K. (2023). What is generative AI?. IBM Research (verkkoaineisto). Saatavissa (viitattu 21.3.2024): <https://research.ibm.com/blog/what-is-generative-ai>
- Microsoft. (2024). Copilot for Microsoft 365. Microsoft (verkkoaineisto). Saatavissa (viitattu 12.4.2024): [https://www.microsoft.com/en-us/microsoft-365/enterprise/copilot-for-microsoft-365#tabs-pill-bar-oca455\\_tab0](https://www.microsoft.com/en-us/microsoft-365/enterprise/copilot-for-microsoft-365#tabs-pill-bar-oca455_tab0)
- Nazeer, S., Sumbal, M. S., Liu, G., Munir, H. & Tsui, E. (2023). The next big thing: role of ChatGPT in personal knowledge management challenges and opportunities for knowledge workers across diverse disciplines. *Global Knowledge, Memory and Communication*.
- Negash, S. (2004). Business intelligence. *Communications of the association for information systems*, Vol. 13(1), 15.
- Nonaka, I., Toyama R., Hirata T., Bigelow S., Hirose A. & Kohlbacher F. (2008). *Managing Flow*. London: Palgrave Macmillan UK. Saatavissa: <https://doi.org/10.1057/9780230583702>.
- OpenAI. (2023). Fine-tuning. OpenAI (verkkoaineisto). Saatavissa (viitattu 12.4.2024): <https://platform.openai.com/docs/guides/fine-tuning>
- Our World in Data. (2024). Annual scholarly publications on artificial intelligence. Our World in Data (verkkoaineisto). Saatavissa (viitattu 18.1.2024): <https://ourworldindata.org/grapher/annual-scholarly-publications-on-artificial-intelligence?tab=table&time=2015..latest>
- Rane, N. (2023). Role and Challenges of ChatGPT and Similar Generative Artificial Intelligence in Business Management. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4603227>
- Rane, N., Choudhary, S. & Rane, J. (2024). Gemini Versus ChatGPT: Applications, Performance, Architecture, Capabilities, and Implementation. *Performance, Architecture, Capabilities, and Implementation*.
- Renaud, K., Warkentin, M. & Westerman, G. (2023). From ChatGPT to HackGPT: Meeting the cybersecurity Threat of generative AI. *MIT Sloan Management Review*, Vol. 64(3), 1–4.
- Russell, S. & Norvig, P. (2016). *Artificial intelligence: a modern approach*. 3rd edition. Boston: Pearson.
- Stephenson, D. (2018). *Big data demystified: How to use big data, data science and AI to make better business decisions and gain competitive advantage*. 1st edition. Pearson Education Limited.
- Sumbal, M. S. & Amber, Q. (2024). ChatGPT: a game changer for knowledge management in organizations. *Kybernetes*.
- Vadari, S. & Malladi, C. (2024). Generative Knowledge Management for Financial Inclusion Through Financial Literacy: A Systematic Review. *The ICFAI University Journal of Knowledge Management*, Vol. 22(1), 39–75.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Advances in neural information processing systems*, Vol. 30.
- Wei, X., Cui, X., Cheng, N., Wang, X., Zhang, X., Huang, S., et al. (2023). Zero-shot information extraction via chatting with chatgpt. *arXiv preprint arXiv:2302.10205*.

Wu, S., Irsoy, O., Lu, S., Dabrovolski, V., Dredze, M., Gehrmann, S., et al. (2023). Bloomberggpt: A large language model for finance. arXiv preprint arXiv:2303.17564.

Xia, L., Li, C., Zhang, C., Liu, S. & Zheng, P. (2024). Leveraging error-assisted fine-tuning large language models for manufacturing excellence. *Robotics and Computer-Integrated Manufacturing*, Vol. 88, s. 102728-. <https://doi.org/10.1016/j.rcim.2024.102728>

Yenduri, G., Srivastava, G., Maddikunta, P. K. R., Jhaveri, R. H., Wang, W., Vasilakos, A. V., et al. (2023). Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions. arXiv preprint arXiv:2305.10435.

## LIITE 1: TUTKIMUSAINEISTO

Tekijät	Julkaisu- vuosi	Otsikko	Sisällön kuvaus
Feuerriegel et al.	2023	Generative AI	Tarkastelee generatiivista tekoälyä liiketoiminnan ja tietojärjestelmien (Business & Information Systems Engineering) näkökulmasta sekä soioteknisestä näkökulmasta
Korzynski et al.	2023	Generative artificial intelligence as a new context for management theories: analysis of ChatGPT.	Tarkastelee generatiivista tekoälyä liikkeenjohdon ja tietojohdamisen teorioiden ja konseptien kontekstina
Nazeer et al.	2023	The next big thing: role of ChatGPT in personal knowledge management challenges and opportunities for knowledge workers across diverse disciplines.	Tarkastelee ChatGPT:n roolia henkilökohtaisessa tietojohdamisessa
Rane, N.	2023	Role and Challenges of ChatGPT and Similar Generative Artificial Intelligence in Business Management.	Tarkastelee ChatGPT:n roolia ja haasteita liiketoiminnassa, mukaan lukien tiedonhallinnan ja liiketoimintatiedon analyysin näkökulmat
Alavi et al.	2024	Knowledge Management Perspective of Generative Artificial Intelligence	Käsittelee generatiivista tekoälyä tiedon luomisen, varastoinnin, jakamisen ja hyödyntämisen näkökulmista
Benbya et al.	2024	Navigating Generative Artificial Intelligence Promises and Perils for Knowledge and Creative Work	Tarkastelee generatiivisen tekoälyn vaikutuksia tietotyöhön tiedon luomisen, hankkimisen, jakamisen ja hyödyntämisen näkökulmista
Sumbal & Amber	2024	ChatGPT: a game changer for knowledge management in organizations	Konseptuaalinen tarkastelu ChatGPT:stä Nonakan SECI-mallin näkökulmasta.
Vadari & Malladi	2024	Generative Knowledge Management for Financial Inclusion Through Financial Literacy: A Systematic Review	Käsittelee generatiivisen tekoälyn roolia tiedonhallinnassa finanssisektorin näkökulmasta. Esittelee termin "Generative Knowledge Management".

## LIITE 2: TUTKIMUSAINEISTON RISTIINTAULUKOINTI

Tyyppi	Artikkeli	Feuerriegel et al. 2023	Korzynski et al. 2023	Nazeer et al. 2023	Rane, N. 2023	Alavi et al. 2024	Benbya et al. 2024	Sumbal & Amber 2024	Vadari & Mal-ladi 2024	Yht.
Sovellukset	Tietotarpeiden määrittely	X	X		X	X	X		X	6
	Tiedon hankinta					X	X	X	X	4
	Tiedon prosessointi ja analysointi	X		X	X	X	X			5
	Tiedon jakaminen	X	X			X	X	X	X	6
	Tiedon hyödyntäminen ja palaute	X		X		X	X		X	5
Haasteet	Tiedon laatu	X			X	X	X		X	5
	Tietoturva			X		X	X			3
	Toimialakohtaisuus	X			X				X	3
	Eettiset haasteet	X		X	X	X			X	5
Yht.		7	2	4	5	8	7	2	7	