

Alexi Mutka

**TEKOÄLYN JA DATAINTEGRAATIOIDEN HYÖDYNTÄMINEN YRITYSTEN HIILIJALANJÄLJEN LASKENNASSA**



JYVÄSKYLÄN YLIOPISTO  
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA  
2023

# TIIVISTELMÄ

Mutka, Aleksi

Tekoölyn ja dataintegraatioiden hyödyntäminen yritysten hiilijalanjäljen laskennassa

Jyväskylä: Jyväskylän yliopisto, 2023, 22 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja(t): Vuorinen, Jukka

Hiilijalanjäljen laskenta on olennainen osa yritysten päästöjen hallintaa. Päästöjen laskennan avulla yrityksen on mahdollista paikantaa isoimmat päästölähteet ja näin ollen myös pienentää niitä. Hiilijalanjäljen laskenta on myös osalle yrityksistä pakollista erilaisten säännösten ja direktiivien myötä. Lisäksi yritysten vastuullisuudesta ovat kiinnostuneita myös erilaiset sidosryhmät, kuten sijoittajat ja asiakkaat. Tässä kandidaatintutkielmassa tarkastellaan hiilijalanjäljen laskentaa ja selvitetään, miten tekoölyn ja dataintegraation avulla voitaisiin tehostaa nykyisiä menetelmiä. Tutkielma tehtiin kirjallisuuskatsauksena. Aiheen tutkiminen on tärkeää, koska siitä on toistaiseksi vain vähän ajankohtaista tutkimusta. Lisäksi hiilijalanjäljen laskenta tulee lisääntymään merkittävästi tulevien vuosien aikana, joten on tärkeää, että käytettävät menetelmät ovat mahdollisimman tehokkaita.

Avainsanat: hiilijalanjälki, hiilijalanjäljen laskenta, tekoöly, dataintegraatio

## **ABSTRACT**

Mutka, Aleksi

Utilizing artificial intelligence and data integrations in the calculation of companies' carbon footprints

Jyväskylä: University of Jyväskylä, 2023, 22 pp.

Information systems science, bachelor's thesis

Supervisor(s): Vuorinen, Jukka

Carbon footprint calculation is an essential part of corporate emissions management. With emissions calculation, it is possible for a company to locate the largest sources of emissions and thus also reduce them. Carbon footprint calculation is also mandatory for some companies due to various regulations and directives. In addition, different stakeholders, such as investors and customers, are also interested in the responsibility of companies. This bachelor's thesis examines carbon footprint calculation and investigates how artificial intelligence and data integration could be used to enhance current methods. The thesis was done as a literature review. Researching the topic is important because there is currently only a little up-to-date research on it. In addition, carbon footprint calculation will significantly increase in the coming years, so it is important that the methods used are as effective as possible.

Keywords: carbon footprint, carbon footprint calculation, artificial intelligence, data integration

# SISÄLLYS

TIIVISTELMÄ  
ABSTRACT

KÄSITEHAKEMISTO .....	5
1 JOHDANTO.....	7
2 HIILIJALANJÄLKI JA SEN LASKENTA .....	9
2.1 Hiilijalanjäljen määritelmä .....	9
2.2 Hiilijalanjäljen laskennan syitä .....	10
2.3 Hiilijalanjäljen laskenta tällä hetkellä .....	12
3 TEKOÄLY JA DATAINTEGRAATIOT .....	14
3.1 Tekoäly ja dataintegraation määritelmä.....	14
3.2 Tekoilyn ja dataintegraatioiden yleisimmät käyttötarkoitukset.....	14
4 TEKOÄLY JA DATAINTEGRAATIO HIILIJALANJÄLJEN LASKENNASSA .....	16
4.1 Tekoilyn ja dataintegraatioiden hyödyntäminen hiilijalanjäljen laskennassa tällä hetkellä.....	16
4.2 Tekoilyn ja dataintegraatioiden hyödyntäminen hiilijalanjäljen laskennassa tulevaisuudessa.....	17
5 YHTEENVETO .....	19
LÄHTEET .....	21

## KÄSITEHAKEMISTO

**CSRD** (Corporate Sustainability Reporting Directive) = EU:n määrittämä, vuonna 2023 voimaan astunut vastuullisuusraportointidirektiivi yrityksille (Euroopan Unioni, ei pvm.)

**Dataintegraatio** = Datan kerääminen useasta tietolähteestä yhteen kohteeseen

**ESG** (Environmental, social and governance) = ESG-raportointi tarkoittaa vastuullisuusraportointia niin, että se käsittää luonnon, yhteiskuntavastuun sekä hallintotapaan liittyvät asiat

**GHG** (Greenhouse gas) = Kasvihuonekaasu

**GHG-protokolla** = Kasvihuonekaasu-protokolla, jonka avulla voidaan laskea yrityksen päästöjä

**Hiilijalanjälki** = Hiilijalanjälki kertoo esimerkiksi yrityksen suorasti tai epäsuorasti aiheuttamista hiilidioksidipäästöistä (Wiedmann & Minx, 2008)

**Hiilijalanjäljen laskenta** = Hiilijalanjäljen laskennalla voidaan mitata esimerkiksi yrityksen aiheuttamat päästöt

**Koneoppiminen** (Machine learning) = Tekoälyn muoto, jonka tarkoituksena on opettaa tietokonetta datan avulla

**NFRD** (Non-Financial Reporting Directive) = EU:n määrittämä, CSRD:tä suppeampi ja vanhempi vastuullisuusraportointidirektiivi (Euroopan Unioni, ei pvm.)

**Scope** = Hiilijalanjäljen laskennassa käytetään kolmea eri scopea, joilla rajataan laskentaa esimerkiksi yrityksen suoriin ja epäsuoriin päästöihin

**Tekoäly (AI)** = Järjestelmä tai kone, joka kykenee ihmismäiseen ajatteluun tai toimintaan

**Viherpesu (Greenwashing)** = Viherpesua voi tapahtua esimerkiksi yrityksen viestinnässä, jos se väittää olevan vastuullisempi ympäristöön liittyvissä asioissa kuin se todellisuudessa on

# 1 JOHDANTO

Ilmaston lämmetessä päästöjen hallinta on noussut tärkeäksi teemaksi. Erityisen tärkeää se on yrityksille, joilla on suuri vaikutus koko maailmanlaajuisiin päästöihin. Päästöjen laskenta on jo arkipäivää suurimmille yrityksille, mutta siihen liittyvät menetelmät ja säädökset ovat silti epäselviä ja epä johdonmukaisia. Yritysten aiheuttamia päästöjä pyritään seuraamaan jatkossa tarkemmin esimerkiksi EU:n määräämien direktiivien kautta. Päästöjen laskennan yleistyessä on myös tärkeää, että siihen käytettävät menetelmät ovat mahdollisimman tutkittuja ja tehokkaita. Aiheesta on toistaiseksi melko vähän ajantasaisia tutkimuksia, koska aihe ja menetelmät ovat niin uusia ja siten myös muuttuvat usein. Monet hiilijalanjäljen laskentaan liittyvät aiemmat tutkimukset keskittyvät esimerkiksi yksittäisen tuotteen tai valtion hiilijalanjäljen laskentaan. Merkittävimmät tutkimukset ovat myös melko vanhoja (esim. Wiedmann & Minx, 2008), joten niiden jälkeen esimerkiksi laskentaan liittyvät menetelmät ja laskennan motiivit ovat jo kerenneet muuttua. Tämä tutkielma käsittelee aihetta erityisesti koko yrityksen hiilijalanjäljen näkökulmasta, josta on varsin vähän aiempaa tutkimusta. Tässä tutkielmassa tutkitaan aihetta kirjallisuuskatsauksen muodossa, tutkimuskysymyksen ollessa seuraava:

- Miten yritysten hiilijalanjäljen laskentaa voitaisiin tehostaa hyödyntämällä tekoälyä sekä dataintegraatioita osana sitä?

Tutkielma koostuu johdannon ja yhteenvedon lisäksi kolmesta sisältöluvusta. Ensimmäisessä sisältöluvussa käsitellään hiilijalanjälkeä käsitteenä, hiilijalanjäljen laskentaan liittyviä syitä sekä laskentaan käytettyjä menetelmiä. Toisessa luvussa esitellään tekoälyn sekä dataintegraation käsitteet ja perehdytään niiden yleisimpiin käyttötarkoituksiin ja käytön hyötyihin. Viimeinen sisältöluke pitää sisällään katsauksen siitä, kuinka kyseisiä teknologioita hyödynnetään laskennassa tällä hetkellä. Lisäksi siinä pohditaan, miten niitä voitaisiin hyödyntää laskennassa tulevaisuudessa sekä sitä, saadaanko niiden avulla tehostettua laskentamenetelmiä. Tiedonhakuun on käytetty pääsääntöisesti Google Scholar -palvelua sekä Jyväskylän yliopiston kirjaston JYKDOK-palvelua. Ha-

kusanoina käytettiin esimerkiksi sanoja "carbon footprint calculation", "carbon footprint estimation methods" sekä "corporate carbon footprint". Lähteinä on käytetty pääsääntöisesti englanninkielisiä tieteellisiä artikkeleita. Lisäksi lähteinä on käytetty muutamia suomenkielisiä artikkeleita esimerkiksi hiilijalanjäljen laskentaa konsultointipalveluna tarjoavien yritysten nettisivuilta.



## 2 HIILIJALANJÄLKI JA SEN LASKENTA

Tämä luku esittelee hiilijalanjäljen määritelmän ja kertoo syitä hiilijalanjäljen laskennalle yritysten näkökulmasta. Lisäksi luvussa kerrotaan, miten hiilijalanjäljen laskentaa tällä hetkellä tehdään.

### 2.1 Hiilijalanjäljen määritelmä

Wiedmannin ja Minxin (2008) tekemän kirjallisuuskatsauksen perusteella hiilijalanjäljelle on vaikea löytää yksiselitteistä määritelmää. Käsitettä lähestytään erilaisista näkökulmista, kuten yksittäisen tuotteen elinkaaren aiheuttamat kasvihuonekaasupäästöt, tai yrityksen koko toiminnan aiheuttamat päästöt (Wiedmann & Minx, 2008). Wiedmann ja Minx (2008) kuitenkin antoivat itse käsitteelle määritelmän, joka on yhä edelleen yksi käytetyimmistä. Siinä hiilijalanjälki tarkoittaa jonkin toiminnan tai yksittäisen tuotteen elinkaaren suorasti tai epäsuorasti aiheuttamaa hiilidioksidipäästöjen määrää. Tähän sisältyy kaikkien eri toimijoiden, kuten yksilöiden, hallitusten sekä yritysten toiminnan aiheuttamat päästöt. Tulee kuitenkin ottaa huomioon, että tässä tutkielmassa keskitytään nimenomaan yritysten hiilijalanjälkeen ja sen laskentaan. Wiedmannin ja Minxin (2008) määritelmän mukaan hiilijalanjälkeä mitataan nimenomaan massayksiköissä, eikä pinta-alana, vaikka sitäkin joissain yhteyksissä on käytetty. Tämä määritelmä ei ota huomioon hiilidioksidin lisäksi muita aineita, pääasiassa niiden huonon mitattavuuden takia. (Wiedmann & Minx, 2008)

## 2.2 Hiilijalanjäljen laskennan syitä

Maapalloa uhkaava ilmaston lämpeneminen sekä lisääntynyt kasvihuonekaasupäästöjen määrä ovat tärkeimpiä syitä hiilijalanjäljen laskennalle. Erään säännön mukaan vain mitattavat asiat ovat hallittavissa, joka on johtanut eri tahoja laskemaan aiheuttamiensa päästöjä hiilijalanjäljen muodossa. (Pandey, Agrawal & Pandey, 2011) Hiilijalanjäljen laskenta auttaa näin ollen hallinnoimaan päästöjä, kun esimerkiksi toiminnan eri osa-alueiden päästöjä pystytään tarkastelemaan erikseen. Tämä mahdollistaa prosessien muuttamisen ja tehostamisen niin, että päästöjä syntyy jatkossa vähemmän. (Pandey ym., 2011)

Tämän tutkielman keskittyessä erityisesti yritysten hiilijalanjäljen laskentaan, on hyvä käydä läpi nimenomaan yrityksiä laskentaan velvoittavia syitä. Muut tahot, kuten yksityishenkilöt tai valtiot, voivat laskea hiilijalanjälkeä eri syistä tai motiiveista riippuen.

Ensimmäinen merkittävä tekijä hiilijalanjäljen laskennalle yritysten näkökulmasta ovat eri sidosryhmät. Yritykset pyrkivät vastaamaan suorien ja välillisten sidosryhmien tarpeisiin niin, että se kuitenkin ottaa huomioon myös tulevaisuuden sidosryhmät. Nämä sidosryhmät sisältävät esimerkiksi osakkeenomistajat, asiakkaat sekä työntekijät. (Dyllick & Hockerts, 2002, viitattu lähteessä Penz & Polsa, 2018)

Swallowin ja Furnissin (2011) tutkimuksen mukaan suurin osa asiakkaista piti tärkeänä sitä, että ostaa tuotteita ja palveluita ympäristöystävällisesti toimivasta yrityksestä. Tämän tutkimuksen mukaan yhä useampi myös vaihtoi kilpailijan tuotteisiin, mikäli he pyrkivät muuttamaan toimintaansa vastuullisempaan suuntaan. Tutkimuksessa selvisi ekologisten tuotteiden suosion olevan suurinta elintarvikealalla, josta voidaan päätellä, että erityisesti kuluttajamarkkinoilla toimivien yritysten on järkevää panostaa ympäristövastuullisuuteen- ja imagoon. Swallowin ja Furnissin (2011) mukaan vastuullisesti toimivien yritysten arvo nousee etenkin nyt milleniaalien vanhetessa ja vaurastuessa.

Toinen yrityksen kannalta olennainen sidosryhmä on sijoittajat. Sijoittajat vaativat jatkuvasti enemmän tietoa yritysten ympäristövastuullisuudesta ja tietojen tulisi olla sijoittajien mukaan nykyistä helpommin saatavilla. Sijoittajat ovat myös auttaneet yrityksiä yhdenmukaistamaan raportointiansa tekemällä esimerkiksi erilaisia tiedonantokehyksiä (CDP, IIGCC). (Sullivan & Gouldson, 2012)

Toisaalta yritysten kunnianhimoiset ilmastotavoitteet voivat olla myös haasteellisia sidosryhmien kannalta. Monet lupauksia tehneet yritykset keskittyvät edelleen enemmän markkinan reaktioihin ja maineen parantamiseen sen sijaan, että ne pyrkisivät parantamaan ympäristöraportoinnin prosessien yhdenmukaisuutta ja laatua. (In & Schumacher, 2021) Sijoittajien lisääntynyt kiinnostus ympäristövastuullisuudesta näkyy yritysten ympäristöraportoinnissa välillä myös negatiivisella tavalla. Monet yritykset hyödyntävät vielä toistaiseksi puutteellista raportoinnin sääntelyä saadakseen esimerkiksi rahoitusta yritykselleen. Tämän takia eri yritysten raporttien todenmukaisuus voi poiketa

hyvinkin paljon toisistaan, aiheuttaen vaikeuksia raporttien vertailulle. Tämä voi taas aiheuttaa ongelmia esimerkiksi sijoittajille, koska yritysten vertailu on hankalaa. (In & Schumacher, 2021)

Erityisesti suurimpien yritysten kohdalla tulee ottaa huomioon myös erilaisia lainsäädäntöjä. Esimerkiksi Euroopassa toimivien yritysten tulee noudattaa EU:n säännöksiä, joista yhtenä tärkeimpänä CSRD (Corporate Sustainability Reporting Directive), joka velvoittaa erityisesti suuria yrityksiä raportoimaan vastuullisuudestaan. Ensimmäiset direktiivin alla olevat yritykset joutuvat julkaisemaan raportin vuodesta 2025 alkaen, jolloin julkaistaan vuoden 2024 tiedot. Kyseinen direktiivi velvoittaa noin 50 000 suurinta yritystä julkaisemaan vastuullisuusraporttinsa. (Euroopan Unioni, ei pvm.) Toisaalta direktiivi koskettaa välillisesti huomattavasti suurempaa määrää, kun tietoja tullaan kysymään myös arvoketjujen muilta osallisilta. Täten myös pienemmät yritykset joutuvat ottamaan uuden direktiivin huomioon toiminnassaan. (Enne Consulting, 2022) Direktiivi tuo apua aiemmin mainittuun ongelmaan, jossa sijoittajilla ja muilla sidosryhmillä voi olla vaikeuksia löytää luotettavia ja yhdenmukaisia vastuullisuustietoja yrityksistä. EU:lla on jo tällä hetkellä olemassa vastuullisuusraportointiin velvoittava direktiivi (NFRD, Non-Financial Reporting Directive), mutta se velvoittaa vain yli 500 työntekijän yrityksiä. (Euroopan Unioni, ei pvm.) Lisäksi aiempi direktiivi on vaatimuksiltaan huomattavasti uutta kevyempi (Enne Consulting, 2022).

Nämä direktiivit ovat erittäin tarpeellisia hiilijalanjäljen laskennan kannalta, jotta laskennasta saadaan luotettavaa ja yhdenmukaista. Harhaanjohtavat vastuullisuusraportit voivat myös johtaa ”viherpesuun” (greenwashing), jos raporttien sisältö ei vastaa toiminnan todellista vaikutusta ilmastoon (In & Schumacher, 2021).

## 2.3 Hiilijalanjäljen laskenta tällä hetkellä

Yritysten hiilijalanjäljen selvittämistä varten on kehitetty lukuisia eri tapoja. Nämä vaihtoehdot mahdollistavat hiilijalanjäljen laskennan erilaisiin tarpeisiin, kuten yksittäisen tuotteen elinkaarianalyysiin tai koko yrityksen tasoiseen hiilijalanjäljen laskentaan. Tämä luku ei tule kuitenkaan käsittelemään tuotteiden elinkaarten aiheuttamien päästöjen laskentaa, tämän tutkielman keskittyessä nimenomaan koko yrityksen hiilijalanjäljen laskentaan.

Menetelmistä yleisimpiä ovat GHG-protokolla (Greenhouse Gas Protocol), joka soveltuu hiilijalanjäljen laskentaan koko yrityksen tasolla, sekä erilaiset ISO-standardit, jotka soveltuvat tuotteiden hiilijalanjäljen laskentaan. Vuonna 2001 alkunsa saanut GHG-protokolla on maailman suosituin hiilijalanjäljen laskentaan käytetty standardi (OpenCO2, 2022; Walenta, 2021). GHG-protokollassa päästöt jaetaan 3 luokkaan (scope): scope 1, scope 2 sekä scope 3. Scope 1 pitää sisällään yrityksen omasta toiminnasta suoraan aiheutuvat päästöt, kuten käytetyt polttoaineet. Scope 2 taas sisältää ostoenergian, kuten esimerkiksi yrityksen kiinteistöissä käytetty sähkö. Scope 3 taas kattaa kaiken muun, kuten työntekijöiden matkustuksen tai toimitusketjujen päästöt. (Walenta, 2021) Scope 3 -luokan päästöjä ei kuitenkaan aina lasketa, vaikka usein suurin osa yrityksen päästöistä meneekin sen alle. Tähän on johtanut esimerkiksi se, että scope 3 -luokan päästöistä on vaikeampi kerätä dataa, koska ne ovat yrityksen kannalta epäsuoria päästöjä, eikä niihin täten ole yhtä paljon vaikutusvaltaa. (Green Carbon, 2021) Tässä kohdataan myös ensimmäisiä hiilijalanjäljen laskentaan liittyviä haasteita: datan puute tai sen huono saatavuus (OpenCO2, 2022). Laskentaan tarvittavaa dataa on saatavilla myös ulkopuolisilta datatoimittajilta, mutta usein heiltä löytyy dataa vain kaikista suurimmille yrityksille (Busch, 2022). Tämän takia suurimman osan yrityksistä on hankittava tarvittava data itse.

Netistä löytyy paljon yksinkertaisia ja ilmaisia laskureita, jonka avulla yritys voi arvioida päästöjään helposti, mikäli tarvittavat tiedot laskentaa varten ovat jo saatavilla. (De Grosbois & Fennell, 2011). Harangozon ja Szigetin (2017) mukaan nämä laskurit ovat kuitenkin tuloksiltaan osittain puutteellisia, eikä eri laskureiden tuloksia voida vertailla keskenään, vaikka niihin syötettäisiin samat tiedot.

Ilmaisten laskureiden jälkeen jäljelle jäävät keinot ovat yrityksille sekä aikaa vieviä että kalliita (González-González, García-Méndez, De Arriba-Pérez, González-Castaño & Barba-Seara, 2022). Tämä on huomattu myös kotimaassa, näin laskennasta puhuttiin Tivin (2022) artikkelissa:

Hiilijalanjälki selvitetään tyypillisesti niin, että asiantuntijat ja yritysten henkilöstö keräävät lukuja eri lähteistä ja laskevat niistä pitkälti manuaalisesti rivi kerrallaan yrityksen hiilijalanjäljen. Tämä on hankalaa, kallista ja epätarkkaa. Lisäksi laskennassa käytettävät päästökertoimet laahaavat jäljessä vuositolkulla. (Tivi, 2022)

Laskentapalveluita tarjotaan myös ulkopuolisina konsultointipalveluina, mikäli laskentaa ei haluta tehdä yrityksen sisäisesti. Niiden avulla saadaan myös hankittua mahdollisesti tarvittavia sertifikaatteja hiilijalanjälkeen liittyen. (González-González, ym., 2022)

Perinteisten laskentatapojen lisäksi on kuitenkin jo olemassa tehokkaampia, teknologioita paremmin hyödyntäviä menetelmiä. González-González ym. (2022) esittelee menetelmän, jossa hiilijalanjälkeä lasketaan automaattisesti yrityksen pankkitapahtumien perusteella. Tässä menetelmässä jokaisen ostoksen hiilijalanjälki on laskettu hyödyntämällä maksun suuruutta ja sitä, mihin osa-alueeseen maksu kuuluu. Näitä osa-alueita ovat esimerkiksi julkisen liikenteen matkaliput tai sähkölaskut.

Uudenlaisia menetelmiä löytyy käytöstä myös Suomessa. Esimerkiksi Helsingin Yliopiston (ei pvm.) artikkelissa kerrotaan innovaatiosta, jonka avulla yritysten hiilijalanjälkeä pystytään laskemaan taloushallinnon datan perusteella reaaliaikaisesti. Lisäksi Kauppalehti (2022) esittelee Digian uuden raportointijärjestelmän, joka kykenee laskemaan lihantuotannon päästöjä automatisoidusti. Järjestelmän avulla pyritään auttamaan sen käyttäjiä paitsi laskemaan hiilijalanjälkeänsä, mutta myös pienentämään sitä. (Kauppalehti, 2022)

Voidaan siis todeta, että vaikka perinteiset menetelmät ovat edelleen käytetyimpiä (González-González, ym., 2022), on uusia ja tehokkaampia menetelmiä jo markkinoilla. Uusista menetelmistä ei ole vielä tällä hetkellä juurikaan tieteellisiä julkaisuja, mikä voi myös kertoa siitä, että ne ovat pääosin vasta kehittyneillä ja tulossa markkinoille.

## 3 TEKOÄLY JA DATAINTEGRAATIOT

Tämä luku käsittelee tekoälyn ja dataintegraatioiden määritelmiä sekä niiden yleisimpiä käyttötarkoituksia.

### 3.1 Tekoäly ja dataintegraation määritelmä

Tekoälylle on annettu useita erilaisia määritelmiä. Se on määritelty esimerkiksi seuraavilla tavoilla: "Tekoäly on kiinnostunut sellaisten tietokoneiden kehittämisestä, jotka pystyvät ihmismäisiin ajatusprosesseihin kuten oppimiseen, päätelyyn ja itsensä korjaamiseen." (Kok, Boers, Kusters, van der Putten & Poel, 2009) tai "Ihmisen älykkyyden laajentaminen tietokoneiden käytön avulla -." (Kok ym., 2009) Lisäksi tekoälyjärjestelmille on annettu omia määritelmiä, kuten: "Järjestelmä, joka ajattelee kuin ihminen." (Kok ym., 2009) tai "Järjestelmä, joka ajattelee rationaalisesti." (Kok ym., 2009)

Dataintegraatio taas on ratkaisu tilanteisiin, jossa vaaditaan dataa käyttöön erilaisista sovelluksista tai eri organisaatioilta samaan aikaan. Dataintegraatiota verrataan esimerkiksi uuden tietokoneen tilaamiseen, jossa halutaan lyhyessä ajassa saada kasattua eri valmistajien osia sisältävä laite. Tarkoituksena on siis hyödyntää olemassa olevaa dataa mahdollisimman joustavalla ja tehokkaalla tavalla. (Doan, Halevy & Ives, 2012)

### 3.2 Tekoälyn ja dataintegraatioiden yleisimmät käyttötarkoitukset

Davenport ja Ronanki (2018) esittelee kolme yleisintä tekoälyn käyttötarkoitusta yritystoiminnassa. Ensimmäinen niistä on prosessien automatisointi, jonka avulla saadaan esimerkiksi vähennettyä käsityötä datan siirtämisessä paikasta

toiseen. Yleisesti tämän ajatellaan vähentävän työpaikkoja, mutta tutkimusten mukaan ainakaan toistaiseksi automaatio ei juurikaan ole korvannut ihmisiä työpaikoilla. Tulevaisuudessa automaation kehittyessä on kuitenkin mahdollista, että prosessien automatisointi vähentää ihmisten tarvetta työpaikoilla. (Davenport & Ronanki, 2018)

Toinen tekoälyn käyttötarkoitus on ”kognitiiviset näkemykset”, jolla tarkoitetaan tekoälyn hyödyntämistä esimerkiksi ostokäyttäytymisen ennustamisessa. Tekoäly kykenee tulkitsemaan suuria määriä dataa ja tekemään sen perusteella haluttuja ennustuksia tai analyysejä. Samankaltaista analytiikkaa tehdään myös perinteisillä keinoilla, mutta Davenportin ja Ronankin (2018) mukaan tekoälyn hyödyntäminen mahdollistaa suurten datamäärien käsittelyn nopeammin. Lisäksi koneoppimisen kautta tekoälyn on mahdollista kehittyä analyyseissa ja ennusteissa sen käsitellessä enemmän ja enemmän dataa. Artikkelissa tätä kuvataan ”analytiikaksi steroideissa”. (Davenport & Ronanki, 2018)

Kolmanneksi yleisimpänä käyttötarkoituksena mainitaan tekoälyn hyödyntäminen esimerkiksi asiakaspalvelussa tai työntekijöiden teknisten ongelmien ratkaisussa. Tekoälyn avulla voidaan luoda asiakkaan tai työntekijän äidinkielellä toimivia chatbotteja, jotka pyrkivät ratkaisemaan ongelman ilman ihmisen työpanosta. Lisäksi tekoälyä voidaan hyödyntää esimerkiksi hoitosuunnitelman kustomoinnissa tai tuotteen suosittelussa asiakkaalle aiemman historian perusteella. (Davenport & Ronanki, 2018)

Doan ym. (2012) kertovat kaksi erilaista käyttötapausta dataintegraatiosta. Ensimmäisessä esimerkissä kerrotaan kuvitteellisesta Yhdysvalloissa toimivasta yrityksestä, joka haluaa laajentaa toimintaansa ostamalla eurooppalaisen, toisella alalla toimivan yrityksen. Tässä tapauksessa dataintegraatiota tarvitaan esimerkiksi silloin, kun halutaan hakea tietoa yrityksen kaikista asiakkaista samaan aikaan. Lisäksi työntekijät ovat kahdessa eri tietokannassa, joten myös koko yrityksen työntekijöihin liittyvät kyselyt vaativat dataintegraatiota.

Toinen esimerkki kertoo työnhakijasta, joka haluaa etsiä uusia työpaikkoja eri sivuilta ilman, että hänen tarvitsisi täyttää hakukenttiä erikseen joka sivulla. Dataintegraation avulla hänen on mahdollista hakea yhdellä haulla kaikkien niiden sivujen työpaikkailmoitukset, mistä on dataa saatavilla. (Doan ym., 2012)

## 4 TEKOÄLY JA DATAINTEGRAATIO HIILIJALANJÄLJEN LASKENNASSA

Tässä luvussa käsitellään tekoälyn ja dataintegraation hyödyntämistä hiilijalanjäljen laskennassa. Luvussa esitellään tämänhetkinen tilanne ja selvitetään, miten näitä teknologioita voitaisiin mahdollisesti tulevaisuuden laskennassa hyödyntää.

### 4.1 Tekoälyn ja dataintegraatioiden hyödyntäminen hiilijalanjäljen laskennassa tällä hetkellä

Kuten luvusta 2.3 huomattiin, on yritysten hiilijalanjäljen laskenta edelleen pääosin manuaalista. Automatisoituja ratkaisuja kuitenkin löytyy jo markkinoilta, ja niitä tulee selvästi jatkuvasti lisää. Etenkin yksittäisten henkilöiden hiilijalanjäljen laskentaan löytyy jo runsaasti automatisoituja laskureita, jotka käyttävät hyödykseen esimerkiksi pankkitapahtumia (González-González, ym., 2022). Selvästi myös monet yrityksille suunnatut automatisoidut laskentamenetelmät hyödyntävät pankkitapahtumia. Tästä huomataan, että teknologioiden avulla pyritään muuttamaan jo olemassa olevaa dataa hiilijalanjäljen laskentaan sopivaksi. Esimerkiksi González-González ym. (2022) menetelmässä hyödynnetään tekoälyä maksudatan muuttamiseksi sellaiseen muotoon, josta se voidaan automaattisesti muuntaa hiilijalanjäljeksi. Heidän esittelemässä menetelmässä data oli valmiina ja se syötettiin suoraan järjestelmään, joten dataintegraatiota ei ole tehty. Voidaan kuitenkin päätellä, että esimerkiksi pankista tehtävän dataintegraation avulla laskentaprosessia voitaisiin tehostaa huomattavasti, koska data pystytään tekoälyn avulla muuttamaan laskentakelpoiseksi. Ongelmaksi saattaa kuitenkin muodostua se, ettei esimerkiksi kaikkea laskentaan tarvittavia tietoja ole saatavilla pankkitapahtumien perusteella.

Myös Nguyen, Diaz-Rainey ja Kurupparachchi (2021) esittelevät automatisoidun laskentamenetelmän, joka hyödyntää koneoppimista. Tässä menetelmässä data syötetään järjestelmään suoraan avoimista tietokannoista, joista



saadaan laskentaan tarvittavia tietoja lähinnä suurimmille yrityksille. Puuttuvien tietojen kohdalle (lähinnä scope 3) otetaan dataa joko aiempien vuosien tai vastaavien yritysten tiedoista. Laskentaprosessi ja datalle tehtävät tarvittavat toimenpiteet ovat täysin automatisoituja. (Nguyen, ym., 2021)

Yrityksen hiilijalanjälkeä pyritään toisinaan laskemaan myös yrityksen ulkopuolelta. Vastuullisuusraportista ovat kiinnostuneita esimerkiksi sidosryhmät, ja koska kaikki yritykset eivät ainakaan toistaiseksi ole velvoitettuja laskemaan ja raportoimaan hiilijalanjälkeensä, on kehitetty ratkaisuja yrityksen ulkopuoliseen laskentaan. Ratkaisuksi esitellään tekoälyyn pohjautuvaa laskentamenetelmää, joka arvioi yrityksen hiilijalanjäljen saatavilla olevien tietojen perusteella. Menetelmä käyttää laskentaan yrityksestä vapaasti löytyviä tietoja sekä yrityksen toimialueeseen liittyviä tietoja. (Assael, Heurtebize, Carlier & Soupe, 2023)

Kuten näistä esimerkeistä huomataan, automatisoituun hiilijalanjäljen laskentaan on jo tällä hetkellä olemassa menetelmiä. Ne ovat kuitenkin hyvin tuoreita, eivätkä ne ole vielä kovin laajassa käytössä. Esimerkeistä voidaan päätellä, että laskennan tärkeimmissä vaiheissa, eli tiedonkeruussa sekä itse laskennassa voidaan jo nyt hyödyntää dataintegraatioita ja tekoälyä. Erityisen olennaista on, että eri tietolähteistä kerättyä ja erityyppistä dataa voidaan hyödyntää laskennassa, koska se pystytään tekoälyn avulla muuttamaan laskentakelpoiseksi. Myös itse laskentaprosessi voidaan automatisoida, kuten aiemmissa esimerkeissä sekä esimerkiksi netistä löytyvissä ilmaisissa laskureissa.

## **4.2 Tekoälyn ja dataintegraatioiden hyödyntäminen hiilijalanjäljen laskennassa tulevaisuudessa**

Kuten aiemmassa alaluvussa huomattiin, teknologioiden hyödyntäminen hiilijalanjäljen laskennassa on jatkuvasti kehittymässä eteenpäin. Vaikka teknologioita ei vielä suurimmassa osassa tapauksia hyödynnetäkään, alkaa esimerkkejä soveltuvista menetelmistä jo löytymään.

Esimerkiksi tulevien EU-direktiivien takia hiilijalanjäljen laskenta tulee yleistymään tulevina vuosina paljon. Sen takia on tärkeää, että laskentamenetelmät ovat mahdollisimman tehokkaita ja edullisia toteuttaa, jotta mahdollisimman moni yritys saataisiin laskemaan hiilijalanjälkeään. Tärkeimpänä asiana hiilijalanjäljen laskenta mahdollistaisi kuitenkin päästöjen seurannan ja sitä kautta niiden pienentämisen. Jotta yritys hyötyy aidosti laskennasta, tulee sen antaman tuloksen olla tarkka. Tässä on myös yksi isoimmista ongelmista liittyen automatisoituun laskentaan.

Aiemmassa luvussa esitellyt esimerkit ovat kuitenkin jo huomattavasti tarkempia verrattuna jo olemassa oleviin menetelmiin. Esimerkiksi Nguyen ym. (2021) esittelemä menetelmä on parhaimmillaan jopa 25%-30% aiempia menetelmiä tarkempi. Voidaan siis todeta laskentamenetelmien kehittyvän nopeasti. Onkin todennäköistä, että laskennan yleistyessä yhä useampi yritys käyttää

automatisoituja laskentamenetelmiä. Ongelmia voi kuitenkin tulla laskennan tarkkuuden suhteen, sekä esimerkiksi erilaisten sertifikaattien saaminen automatisoidulla laskennalla voi olla vaikeaa. Automatisoidun laskennan tarkkuuteen vaikuttaa myös esimerkiksi se, minkälaista dataa siihen käytetään. Jos käytettävästä datasta puuttuu laskennassa olennaisia tietoja, se joudutaan ottamaan joko avoimista tietokannoista tai generoimaan esimerkiksi historiatietojen perusteella. Positiivista on kuitenkin se, että esimerkkien perusteella hiilijalanjälkeä voidaan laskea tekoälyn avulla hyvin erilaisten datalähteiden pohjalta. Ensimmäisessä esimerkissä käytettiin pankkitapahtumia, kun taas toinen esimerkki käytti laskentaan avoimia tietokantoja.

## 5 YHTEENVETO

Tämän kirjallisuuskatsauksen tarkoituksena oli selvittää, miten nykyisiä hiilijalanjäljen laskennan menetelmiä voitaisiin tehostaa hyödyntämällä tekoälyä ja dataintegraatioita.

Hiilijalanjälki käsitteenä on suhteellisen tuore, minkä takia siihen liittyy paljon keskustelua siitä, mikä olisi oikea määritelmä käsitteelle. Hiilijalanjälki voidaan käsittää hieman eri tavoin kontekstista riippuen, mutta pääsääntöisesti sen ajatellaan tarkoittavan kasvihuonekaasupäästöjä tai hiilidioksidipäästöjä.

Hiilijalanjäljen käsitteen tavoin sen laskentaan liittyy myös paljon epävarmuutta. Sitä lasketaan usein hieman eri tavoin riippuen esimerkiksi maantieteellisestä sijainnista, toimialasta tai yrityksen kokoluokasta. Esimerkiksi Euroopan laajuisia direktiivejä on jo olemassa, mutta ne eivät siltikään koske kaikkia yrityksiä. Tämän takia yritysten raportit voivat poiketa toisistaan paljonkin, mikä vaikeuttaa niiden vertailua. Kaikki raportit eivät myöskään ole ulkopuolisen tahon tarkastamia, joten viherpesun riski on olemassa. On selkeää, että laskennassa on ongelmia, kuten sen vaatima suuri työpanos. Tämä johtuu esimerkiksi siitä, että laskentaa tehdään pääosin käsityönä.

Laskennan ongelmat tulevat kunnolla esiin todennäköisesti vasta tulevaisuudessa laskennan yleistyessä erilaisten säännösten ja direktiivien takia. Vaikka EU:n uusi direktiivi koskeekin vain suurimpia yrityksiä, tulee se vaikuttamaan välillisesti myös pienempiin. Nimenomaan pk-yritykset ovat se ryhmä, jota laskennan ongelmat pahiten koskettavat. Näillä yrityksillä ei ole käytössään samanlaisia resursseja kuin esimerkiksi listautuneilla yrityksillä, jolloin hiilijalanjäljen laskenta voi tuottaa isoja ongelmia. Tätä varten on tärkeää, että laskentaan käytettyjä menetelmiä tutkitaan ja kehitetään niin, että ne olisivat mahdollisimman tehokkaita.

Kirjallisuuskatsauksen perusteella ongelmaan on jo olemassa ratkaisuja. Vaikka suurin osa laskennasta tapahtuu edelleen ulkopuolisena konsultointipalveluna tai yrityksen sisäisesti tehtävänä käsityönä, on esimerkkejä automatisoiduista laskentatavoista jo kehitetty. Myös Suomen markkinoilta löytyy ainakin muutama ratkaisu, jossa laskenta on automatisoitu.

Kirjallisuuskatsauksessa esitellyt menetelmät hyödyntävät etenkin tekoälyä. Tekoälyä on käytetty mm. datan muuttamisessa laskentakelpoiseen muotoon sekä itse laskentaprosessissa. Dataintegraatioita ei suoranaisesti mainittu käytettäneen, mutta voidaan päätellä sen hyödyntämisen olevan mahdollista. Esimerkeissä nimittäin laskettiin hiilijalanjälkeä niin pankkitapahtumien kuin erilaisten tietokantojenkin datan pohjalta. Erilaista dataa voidaan siis muokata niin, että sitä voidaan hyödyntää laskennassa. On siis selvää, että näiden teknologioiden avulla laskentaa saadaan tehtyä huomattavasti tehokkaammaksi. Dataintegraatio poistaa tarpeen datan manuaaliselta keräämiseltä, kun taas tekoälyn avulla voidaan muuntaa data halutunlaiseen muotoon ja lisäksi suorittaa laskentaprosessi automatisoidusti.

Tekoälyn ja dataintegraatioiden hyödyntäminen hiilijalanjäljen laskennassa ei kuitenkaan ole täysin ongelmaton. Ongelmia voi syntyä esimerkiksi puutteellisen datan kohdalla, jos kaikkea laskentaan tarvittavaa dataa ei ole saatavilla. Tällöin puuttuvaa dataa on mahdollista korvata esimerkiksi aiempien vuosien tiedoilla tai generoimalla sitä, mutta samalla myös laskennan tarkkuus laskee. Samoin jos prosessi on täysin automatisoitu, on tarkkuus luonnollisesti matalampi kuin täysin manuaalisesti tehdyssä laskennassa. Kuitenkin, vaikka laskentaprosessia automatisoitaisiin vain osittain, voidaan sitä tehostaa huomattavasti näiden teknologioiden avulla.

Kirjallisuuskatsauksessa tuli ilmi, että aiheesta löytyy toistaiseksi melko vähän aiempaa tutkimusta. Lisäksi monet aiheeseen liittyvät tutkimukset ovat niin vanhoja, että esimerkiksi monet säädökset ja laskennan motivaatiot ovat jo kerenneet muuttua. On mielenkiintoista nähdä, mihin suuntaan hiilijalanjäljen laskenta etenee, etenkin, kun se tulee varmasti yleistymään lähivuosina. Olisi mielenkiintoista tehdä aiheesta uusi kirjallisuuskatsaus esimerkiksi viiden vuoden kuluttua ja katsoa, ovatko laskentaan käytettävät menetelmät muuttaneet.

## LÄHTEET

- Assael, J., Heurtebize, T., Carlier, L., & Soupé, F. (2023). Greenhouse gases emissions: estimating corporate non-reported emissions using interpretable machine learning. *Sustainability*, 15(4), 3391.
- Busch, T., Johnson, M., & Pioch, T. (2022). Corporate carbon performance data: Quo vadis?. *Journal of Industrial Ecology*, 26(1), 350-363.
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard business review*, 96(1), 108-116.
- De Grosbois, D., & Fennell, D. (2011). Carbon footprint of the global hotel companies: Comparison of methodologies and results. *Tourism Recreation Research*, 36(3), 231-245
- Doan, A., Halevy, A., & Ives, Z. (2012). *Principles of data integration*. Elsevier.
- Enne Consulting. (8.8.2022). Vastuullisuusraportoinnin rima nousee ja siitä tulee pakollista – mitä tulevasta EU-direktiivistä tulisi tietää?  
<https://enneconsulting.fi/blogikirjoitus/2022/8/8/vastuullisuusraportoinnin-direktiivi-csr-d>
- Euroopan Unioni. (ei pvm.). Corporate sustainability reporting.  
[https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting\\_en](https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en)
- González-González, J., García-Méndez, S., De Arriba-Pérez, F., González-Castaño, F. J., & Barba-Seara, Ó. (2022). Explainable Automatic Industrial Carbon Footprint Estimation From Bank Transaction Classification Using Natural Language Processing. *IEEE Access*, 10, 126326-126338.
- Green Carbon. (30.3.2021). Mikä ihmeen scope 1, 2, 3? <https://greencarbon.fi/mika-ihmeen-scope-1-2-3/>
- Harangozo, G., & Szigeti, C. (2017). Corporate carbon footprint analysis in practice—With a special focus on validity and reliability issues. *Journal of cleaner production*, 167, 1177-1183.
- Helsingin Yliopisto. (ei pvm). Carbonlink tarjoaa uudenlaista hiilijalanjälkilaskuria.  
<https://www.helsinki.fi/fi/innovaatiot-ja-yhteistyö/innovaatiot/tutkimuksen-kaupallistaminen/carbonlink>

- In, S. Y., & Schumacher, K. (2021). Carbonwashing: a new type of carbon data-related ESG greenwashing. Available at SSRN 3901278.
- Kauppalehti. (28.9.2022). Data varmentaa maatalojen ilmastotekojen vaikuttavuuden. <https://www.kauppalehti.fi/kumppanisisallot/digia/data-varmentaa-maatilojen-ilmastotekojen-vaikuttavuuden/>
- Kok, J. N., Boers, E. J., Kusters, W. A., Van der Putten, P., & Poel, M. (2009). Artificial intelligence: definition, trends, techniques, and cases. *Artificial intelligence*, 1, 270-299.
- Nguyen, Q., Diaz-Rainey, I., & Kuruppuarachchi, D. (2021). Predicting corporate carbon footprints for climate finance risk analyses: a machine learning approach. *Energy Economics*, 95, 105129.
- OpenCO2. (19.9.2022). Mikä on scope 3 ja mitä sen laskennassa tulee huomioida? <https://www.openco2.net/fi/artikkelit/mita-tarkoittaa-scope-3>
- Pandey, D., Agrawal, M., & Pandey, J. S. (2011). Carbon footprint: current methods of estimation. *Environmental monitoring and assessment*, 178, 135-160.
- Penz, E., & Polsa, P. (2018). How do companies reduce their carbon footprint and how do they communicate these measures to stakeholders?. *Journal of Cleaner Production*, 195, 1125-1138.
- Sullivan, R., & Gouldson, A. (2012). Does voluntary carbon reporting meet investors' needs?. *Journal of Cleaner Production*, 36, 60-67.
- Swallow, L., & Furniss, J. (2011). Green business: Reducing carbon footprint cuts costs and provides opportunities. *Montana Business Quarterly*, 49(2), 2.
- Tivi. (25.03.2022). Hiilijalanjäljen pienentäminen ja päästöjen laskenta mullistuvat – Digia pioneerina yritysten it-järjestelmien murroksessa. <https://www.tivi.fi/kumppanisisallot/digia/hiilijalanjaljen-pienentaminen-ja-paastojen-laskenta-mullistuvat-digia-pioneerina-yritysten-it-jarjestelmien-murroksessa/>
- Walenta, J. (2021). The making of the corporate carbon footprint: the politics behind emission scoping. *Journal of Cultural Economy*, 14(5), 533-548.
- Wiedmann, T., & Minx, J. (2008). A definition of 'carbon footprint'. *Ecological economics research trends*, 1(2008), 1-11.