



Tekoäly musiikkituottajan tukena: musiikillisen luovuuden edistäminen tekoälyn avulla

Oulun yliopisto
Tietojenkäsittelytiede
LuK-tutkielma
Juuso Luttinen
2023

Tiivistelmä

Tekoäly on yleistynyt luovien alojen apuvälineenä viime vuosina. Esimerkkeinä tästä ovat ChatGPT ja DALL-E, joista ensimmäinen pystyy tuottamaan ihmismäistä tekstiä ja jälkimmäinen kuvamateriaalia. Musiikkiala on todennäköisesti seuraava luova ala, joka kokee tekoälyn vallankumouksen. Tekoäly on nousemassa yhä tärkeämmäksi osaksi musiikkituotantoa, sillä se tarjoaa monia hyödyllisiä työkaluja ja menetelmiä musiikin luomiseen, jakeluun ja hakemiseen. Tämä kirjallisuuskatsaus tarkastelee tekoälyn käyttöä musiikkituotannon tukena ja luovuuden edistäjänä. Keskeinen havainto on, että tekoäly voi olla hyödyllinen apuväline musiikkituotannossa ja edistää luovuutta, mutta siihen liittyy vielä rajoituksia. Yksi tekoälyn käyttötapaus musiikkituotannossa on generatiivinen musiikki, jossa tekoäly luo musiikkia automaattisesti erilaisten sääntöjen ja mallien perusteella. Tekoälyä voi käyttää myös ikään kuin käänteisesti musiikin analysointiin. Tällainen transkribointi, jolla tunnistetaan elementtejä ja musiikkirakenteita musiikista, voi auttaa musiikin ymmärtämisessä ja oppimisessa. Jotta tekoälyä voisi käyttää miksaamisessa eli ääniraitojen yhteensovittamisessa ennen masterointia, tarvitaan enemmän tutkimusta aiheesta. Tekoälyä pystyy käyttämään musiikin masteroinnissa eli viimeistelyssä, mutta tekoälymasterointi ei yllä ammattilaistasolle. Vaikuttaa siltä, että musiikkiala on valmis ottamaan tekoälyn avosylin vastaan ja hyödyntämään sen tarjoamia mahdollisuuksia, kunhan tekoälysovellukset saadaan integroitua luonnollisesti musiikkituottajien työnkulkuun.

Avainsanat

RNN, LSTM, GAN, DAW, tekoäly, musiikki, musiikkituotanto

Ohjaaja

FT, yliopistonlehtori Mikko Rajanen

Sisällysluettelo

Tiivistelmä	2
Sisällysluettelo	3
1. Johdanto	4
1.1 Motivaatio tutkielman taustalla	4
1.2 Haasteet aiheen ympärillä.....	5
1.3 Tutkielman tavoite ja tutkimuskysymykset	6
1.4 Rajoitukset	6
1.5 Tutkielman rakenne	6
2. Tutkimusmenetelmät.....	8
3. Tekoäly tukena musiikkituotannossa	9
3.1 Musiikkituotantoprosessi ennen ja nyt	9
3.2 Mitä luovuus on ja voiko tekoäly olla luova?.....	10
3.3 Tekoäly säveltämisen tukena	10
3.3.1 Sointurakenteen analysointi	10
3.3.2 Melodian generointi	11
3.3.3 Bassolinjan generointi	12
3.3.4 Rumpubiittien analysointi, generointi ja keinotekoiset rummut	12
3.4 Tekoäly miksaamisen ja masteroinnin apuna	14
4. Pohdinta.....	16
4.1 Sävellys.....	16
4.2 Miksaus ja masterointi	17
4.3 Musiikkituottajien kokemukset tekoälystä	17
4.4 Tekoälyn ja musiikin tulevaisuus	18
5. Yhteenveto	20
Lähdeluettelo.....	21

1. Johdanto

Tekoäly on sen uusien aluevaltausten myötä saanut paljon huomiota mediassa sekä positiivisessa että negatiivisessa mielessä. Aikaisemmin on ajateltu, että luovat alat olisivat kokonaan tekoälyn ulottumattomissa, koska tällaisten alojen taiteellinen työ vaatii paljon intuitiota, affektiivista tietoa, aavistusta, mututuntumaa ja muita ihmisperäisiä piirteitä, joita on vaikea jäljentää algoritmein ja älykkäin järjestelmin (Birtchnell & Elliott, 2018). Näin ei kuitenkaan ole. DALL-E-tekoälyohjelma, joka pystyy luomaan tekstimuotoisesta syötteestä monipuolisia ja hienoja kuvia, on jo jollain tapaa vaikuttanut siihen, miten silmin havaittavaa taidetta luodaan tulevaisuudessa. DALL-E:n ohessa ChatGPT-chattibotti, joka kykenee vastaamaan tekstisyötteisiin hämmästyttävän tarkasti ja uskottavasti, on taas implikoinut, että tekstin kirjoittamisprosessi muuttuu hyvin erilaiseksi.

Ihmiset ympäri maailmaa ovat siis jo päässeet käsiksi julkisiin tekoälysovelluksiin, jotka auttavat heitä ilmaisemaan itseään sanallisesti ja kuvallisesti, mutta musiikillinen ilmaisu ja sen tukeminen on jäänyt vähemmälle huomiolle. Derutyn ym. (2022) mukaan AI-teknologian hyödyntäminen musiikkituotannossa on vielä alkutekijöissään, mutta sillä on potentiaalia jättää pysyvä jälki siihen, miten musiikkia tuotetaan. Myös Alaeddinen ja Tannouryn (2021) mukaan tekoälyllä tulee olemaan suuri vaikutus musiikkiteollisuuteen. Heidän mukaansa sitä ei kuitenkaan pidä nähdä teknologiana, joka korvaa musiikkituotannon luovan puolen, vaan asiana, joka kehittää luovuutta ja antaa muusikolle enemmän mahdollisuuksia olla luova. Tekoälyn hyödyntäminen nykyaikaisessa musiikkituotannossa on siis mahdollista, mutta miten tekoäly tarkalleen ottaen on hyödyllinen musiikkituotannossa tällä hetkellä? Mitä tekoälyllisiä keinoja on olemassa musiikkituotannossa ja voivatko ne aidosti auttaa olemaan musiikillisesti luovempi? Tämän tutkimuksen tarkoitus on vastata näihin sekä myöhemmin esiteltäviin tutkimuskysymyksiin.

Musiikki on siis yksi monista tavoista, joilla ihminen pystyy ilmaisemaan itseään. Musiikin tuottaminen on kuitenkin vaikeaa, koska siinä ihmisen täytyy olla luova, mutta samalla noudattaa tiukkoja sääntöjä (Chu ym., 2022). Musiikin tuottamista vaikeuttaa nykyään myös se, että kappaleen teon kaikki vaiheet ja osa-alueet saattavat olla ainoastaan yhden henkilön harteilla, kun ennen kappaleiden takana saattoi olla isoja ryhmiä (Deruty ym. 2022). Musiikintuotantoon kuuluu paljon resursseja: aikaa, energiaa ja materiaaleja (Liu, 2022). Yksilön rajallinen, musiikillinen tietämys ja nykymusiikin pitkä tuotantoprosessi johtavat siihen, että ihminen ei välttämättä pysty ilmaisemaan itseään niin kuin haluaisi. Musiikin tekeminen on tietenkin taito, jossa ihminen voi kehittyä hyväksi, ellei erinomaiseksi, jos hän vain panostaa siihen määrätietoisesti vuosien ja vuosikymmenien ajan, mutta voisiko tekoäly auttaa ihmistä ilmaisemaan itseään luovemmin ilman, että hänen täytyisi olla niin hyvä musiikkiteorian ja tuotantoon liittyvän teknisen osaamisen suhteen? Vastaus on kyllä, sillä modernin musiikkituotannon kaavamaisuus ja normatiivisuus tekevät musiikkituotannosta hyvän käyttökohteen koneoppimiselle (Liu, 2022).

1.1 Motivaatio tutkielman taustalla

Tämän tutkimuksen tekemistä motivoi paljon allekirjoittaneen oma musiikkituotantoharrastus, joka on alun perin lähtenyt liikkeelle peruskoulupohjaisella tietämyksellä musiikkiteoriasta. Vuosien varrella on tullut opittua runsaasti, mutta puutteita musiikillisessa tietämyksessä on vielä paljon. Internetin foorumeilla on tullut vastaan paljon muita harrastajia, jotka ovat kamppailleet samojen haasteiden kanssa kuin itse; rajallisen musiikillisen ja ääniteknisen tietämyksen vuoksi ei ole voinut tehdä musiikkia niin kuin haluaisi. Haastavaa on ollut esim. mielenkiintoisten sointukulkujen rakentaminen, rytmitys ja äänien miksaaminen. Kun on harjoitellut musiikin tekemistä tietokoneella, on kaivannut tukea prosessin monotonisiin ja teknisiin vaiheisiin, jotka syövät aikaa luovalta työtä. On myös toivonut, että saisi palautetta muutenkin kuin esittelemällä keskeneräistä työtä muille. Musiikkituottamisesta tietämätön kuulija ei usein osaa ottaa kantaa keskeneräisen työn tekniseen

puoleen eikä antaa korjausehdotuksia. Tutkimusta motivoi myös aihealueen eli tekoälyn ajankohtaisuus sekä halu oppia lisää tekoälystä ja sen tuomista mahdollisuuksista. Lisäksi motivaattorina toimii kiinnostus ääntä kohtaan, halu oppia miten ääntä tuottavat ohjelmistot toimivat ns. ”konepellin alla” sekä halu ymmärtää musiikkiin liittyviä käsitteitä paremmin.

1.2 Haasteet aiheen ympärillä

Alaeddine ja Tannoury (2021) tuovat esiin useita haasteita tekoälyyn ja musiikkiin liittyen. Näitä ovat ihmisen epätäydellisyys, mallien yhteensopivuus eri genrejen kanssa ja tulevaisuudenkestävyys, mallien ns. on-the-spot opettaminen, musiikin pitkäaikainen rakenne ja polyfoniset äänet. Kaksikko viittaa ihmisen epätäydellisyydellä siihen, että ihmisen virheet antavat musiikille sielun. Liian täydellinen musiikki kuulostaa geneeriseltä ja tunteettomalta, joten epätäydellisyyden kopioiminen on yksi haaste, kun luodaan tekoälysovelluksia (Alaeddine & Tannoury, 2021). Alaeddine ja Tannoury (2021) kertovat, että mallien opettaminen generoimaan generiippumatonta musiikkia on haastavaa; genrejä on valtavasti ja uusia genrejä syntyy päivittäin. Mallien opettaminen on resurssiriippuvaista ja aikaa vievää, ja prosessin optimointiin liittyy haasteita (Alaeddine & Tannoury, 2021). Musiikkikappaleen rakenne koostuu lyhyistä osista, jotka taas koostuvat toistuvista kaavoista (Alaeddine & Tannoury, 2021). Heidän mukaansa ihminen pystyy muistinsa ansiosta sisäistämään näitä kaavoja ja pääsemään mukaan kappaleeseen, mutta tekoälylle muistaminen on vaikeaa, joten pitkien, loogisten musiikkirakenteiden generoiminen tuo oman haasteensa. Viimeisenä he tuovat esiin musiikkikappaleiden melodioiden tunnistamiseen liittyvän haasteen. Musiikkikappaleissa soi tyypillisesti useampi instrumentti ja melodia samanaikaisesti, mitä kutsutaan polyfoniaksi. Ihmiselle polyfonian eli eri äänien ja melodioiden tunnistaminen on triviaalia, mutta tietokoneelle niiden erottaminen on vaikeaa. Kaksikon mukaan on olennaista erottaa kappaleen eri melodiat toisistaan ennen kuin kappaletta voi käyttää tekoälymallin opettamiseen. Musiikkirakenteet sisältävät valtavan määrän tietoa ja tekoälyn tiedetään loistavan suurten datamäärien kanssa, mutta tekoälyn käyttäminen hyvän musiikin luomiseen on silti haastavaa (Alaeddine & Tannoury, 2021).

Grachten, Lattner ja Deruty (2020) tuovat esiin haasteen käyttäjän kontrollin ja musiikkigeneraattorien päätösvallan tasapainottamiseen liittyen. Tällä haasteella tarkoitetaan käytännössä sitä, miten paljon käyttäjälle täytyy suoda valtaa tekoälyn musiikillisiin päätöksiin ja lopputuloksen muokkaukseen. Carnovalini ja Rodà (2020) tuovat esiin omassa artikkelissaan myös kontrolliin liittyviä haasteita: kyvyn generoida musiikkia narratiivisesti ja tunteellisesti. He viittaavat narratiivisuudella MGS:n (= Music Generation System) kykyyn luoda kehityksen tunnetta musiikkitekseen, jotta teos saisi monimutkaisemman merkityksen. Tunteikkuudella he viittaavat tietysti siihen, että MGS kykenisi herättämään tietynlaisia tunteita generoidulla musiikillaan. Kontrolliin nämä haasteet liittyvät siten, että kontrollin antaminen loppukäyttäjälle näiden ominaisuuksien suhteen on haastavaa. Aiemmin esitettyjen haasteiden lisäksi Carnovalini ja Rodà (2020) kirjoittavat haasteista, jotka liittyvät tekniikoiden hybridisaatioon sekä musiikin renderöintiin ja evaluointiin. Hybridisaatiolla tarkoitetaan useamman kuin yhden tekniikan hyödyntämistä yhdessä MGS:ssä, renderöinnillä MGS:n generoiman äänen laatua ja evaluoinnilla MGS:n tuotosten hyödyllisyyden arviointia (Carnovalini & Rodà, 2020).

Aikaisempi tutkimus käsittelee siis hyvin pitkälti sointujen, sointurakenteiden ja melodioiden generoimista tekoälyn avulla. Näin ollen keskeinen haaste tekoälyn ja musiikin yhdistämisessä on se, miten luodaan algoritmeja ja koneoppimismalleja, jotka onnistuvat tekemään ”oikeita” ratkaisuja sävellyksissä. Oikeita-sana on lainausmerkeissä siksi, koska musiikissa oudohko tai väärä sävellysratkaisu voikin olla joskus hyvä. Esimerkiksi jazz-musiikissa rikotaan jatkuvasti sääntöjä, mutta se kuulostaa silti hyvältä. Ehkä paremmin olisi ilmaistu, että miten tekoäly saataisiin tekemään ratkaisuja, jotka ovat melodisesti mielenkiintoisia ja miellyttäviä ihmiskorvalle, mutta eivät liian yllättäviä ja satunnaisia. Musiikin generoimista tekoälyn avulla on käsitelty aiemmin paljon, mutta se, miten tekoäly voitaisiin integroida tehokkaasti nykyaikaiseen musiikkituotantoprosessiin, on

jäänyt selvästi vähemmälle huomiolle.

1.3 Tutkielman tavoite ja tutkimuskysymykset

Tässä tutkielmassa selvitetään miten nykyaikainen muusikko, jonka musiikkitekovälineenä on tietokone, voi hyötyä tekoälystä musiikkituotannon eri vaiheissa. Näin ollen päätutkimuskysymys kuuluu seuraavanlaisesti:

- Miten tekoäly voi tukea nykyaikaista musiikkituottajaa työssään?

Tutkielmassa määritellään, että minkälainen on tyypillinen, nykyaikainen musiikkituotantoprosessi. Prosessi jaetaan vaiheisiin, jotka Deruty ym. (2022) ovat tuoneen esiin tutkimuksessaan, ja pohditaan miten tekoälyä voi hyödyntää luovuuden edistämiseen niiden aikana. Tavoitteena on saada selville, miten juuri tekoäly voi auttaa muusikkoa ilmaisemaan itseään paremmin ja monipuolisemmin sekä miten se voi nopeuttaa sellaisia musiikkituotannon työvaiheita, joita pidetään erityisen työläinä ja pitkästyttävinä. Tutkielmassa kootaan olemassa olevaa tietoa ja tämän pohjalta pohditaan tekoälyn ja musiikin nykyistä tilaa ja tulevaisuutta.

Tavoitteena on tuoda esiin aihealueen keskeisiä konsepteja ja sellaista tietoa, jotka musiikkituottajien ja miksei myös musiikillisten tekoälysovellusten kehittäjien olisi hyvä tiedostaa. Tekoäly on vaikuttanut jo merkittävästi muihin luoviin aloihin, joten musiikkituottajien on hyvä tietää, mitä mahdollisuuksia se voi antaa musiikkituotannolle sekä miten se tällä hetkellä vaikuttaa heidän alaansa ja tulee vaikuttamaan siihen. Tutkielmassa kerrotaan mitä tekoälyratkaisuja on olemassa ja arvioidaan sitä, pystyykö tekoäly tuottamaan tällä hetkellä sellaista materiaalia, joka tuntuu ihmismäiseltä ja herättää tunteita ihmisissä. Selvitetään hieman myös sitä, miten muusikot suhtautuvat tekoälyyn; pidetäänkö sitä uhkana vai mahdollisuutena?

1.4 Rajoitukset

Tutkielmassa keskitytään tarkastelemaan tekoälyn hyödyntämistä nykyaikaisen musiikkituotannon näkökulmasta. Tällainen musiikkituotanto tapahtuu studioympäristössä, jolloin tämän tutkimuksen ulkopuolelle jää perinteisempi musiikkituotanto ja -esitys, jossa ei tallenneta musiikkia, ja jossa musiikkia prosessoidaan vain vähän, esim. live-esityksissä äänen kantavuuden parantaminen. Koska keskitytään itse musiikkituotantoon, jossa tuotetaan uutta ääntä, ulkopuolelle jää myös musiikin jakelu, markkinointi jne. Nykyaikaisella musiikkituotannolla tarkoitetaan tämän tutkielman yhteydessä sellaista musiikkituotantoa, jossa pääpaino on tietokoneella tuotetussa musiikissa. Tutkielmassa ei kehitetä omia tekoälyratkaisuja vaan paneudutaan olemassa oleviin tekoälyjärjestelmiin ja -algoritmeihin ja selvitetään niiden nykyistä tilaa musiikkituotannon kannalta. Huomautuksena, että suuri osa materiaaleista, jotka tulivat vastaan aiheeseen tutustuesssa, ja jotka päätyivät lähdemateriaaleiksi, on julkaistu viime vuosien sisällä. Tämä saattaa kertoa siitä, että aihe on noussut ajankohtaisesti vasta hiljattain ja tutkimus ei ole vielä niin valmista eivätkä esitetyt ratkaisut välttämättä täysin valmiita. Vaikka musiikin ja tekoälyn yhdistäminen on vasta nyt kunnolla nousussa, on aihealueella kuitenkin pidempi historia. Esim. Roads (1985) on käsitellyt musiikkiin ja tekoälyyn liittyviä haasteita jo 80-luvulla.

1.5 Tutkielman rakenne

Tämä tutkielma etenee seuraavalla tavalla. Aluksi kerrotaan tiedonhankinnasta ja tutkimusmenetelmistä, joita on käytetty tämän tutkimuksen luomiseen. Tutkielman päätutkimusmenetelmä on kirjallisuuskatsaus. Kolmannessa luvussa avataan hieman historiaa

musiikkituotannon kehityksestä ja käsitellään tekoälyyn ja musiikkituotantoon liittyvää aikaisempaa tutkimusta eli kerrotaan tekoälyratkaisuista, joita on kehitetty musiikkituotannon avuksi. Neljännessä luvussa analysoidaan aiempaa tutkimusta ja kerrotaan mitä uutta tietoa aikaisempaa tutkimusta analysoimalla on saatu aikaan ja pohditaan miltä tekoälyn ja musiikin nykytila ja tulevaisuus näyttävät. Viidennessä eli viimeisessä luvussa tehdään yhteenveto tutkimuksen tuloksista ja kerrotaan, että millaista jatkotutkimusta aihe vaatii.

2. Tutkimusmenetelmät

Tämän tutkimuksen päätutkimusmenetelmä on kirjallisuuskatsaus, mikä tarkoittaa sitä, että tutkimus on tehty olemassa olevien tutkimusten ja muun kirjallisuuden perusteella. Tässä tutkimuksessa ei siis ole tehty varsinaisesti omaa tutkimusta vaan suuri osa esitetyistä asioista ja pohdinnoista pohjautuu aikaisempaan tutkimukseen.

Aiheeseen tutustuminen alkoi Google Scholarilla, jonka olen todennut hyväksi työkaluksi nopeaan ja iteroivaan tiedonhakuun sekä itse kokotekstitiedostojen löytämiseen ja lataamiseen. Päätelin lähdemateriaalien runsauden ja tuoreuden perusteella, että aihe on toteutettavissa kirjallisuuskatsauksena. Tiedonhaku tapahtui pääasiassa Scholarin kautta. Tiedonhaku alkoi hakulausekkeilla ”ai music production”, ”ai musician help” ja ”ai creativity music”. Edellisiä hakulausekkeitä käytettiin sitten Scopus- ja Web of Science -viitetietokannoissa, joista valittiin uudehkoja lähteitä. Samoja hakulausekkeitä käytettiin myös suoraan ACM Digital Library-, SpringerLink ja ScienceDirect -kokotekstitietokannoissa. Vuosirajausta ei oikeastaan tarvinnut tehdä missään tietokannassa, koska suuri osa relevanttien hakutuloksien teksteistä oli julkaistu viime vuosien sisällä.

Seuraavaksi tehtiin yksityiskohtaisempia hakulausekkeitä. Hakulausekkeilla ”ai music drums”, ”ai drum generation” ja ”deep learning drum pattern” etsittiin rumpuihin liittyviä artikkeleita Scholarista. Hakulausekkeilla ”ai melody generation” ja ”ai music structure” taas etsittiin melodioihin ja musiikkirakenteisiin liittyviä artikkeleita Scholarista. Hakulausekella ”ai music mastering” haettiin äänen masterointiin liittyvää tutkimusta Scholarista. Joitain lähteitä, kuten Grachten ym. (2020) BassNetiä käsittelevä tutkimus, löydettiin valittujen lähteiden lähdeviittauksista. Yhteenvedona tiedonhausta ja hakutuloksista voisi sanoa sen, että nykyinen tekoälyn ja musiikin tutkimus on keskittynyt musiikin analysointiin ja generointiin.

Lähteiden valitseminen eteni seuraavasti. Aluksi tutkimuksen soveltuvuutta lähteeksi arvioitiin lukemalla sen otsikko. Jos otsikko vaikutti vastaavan aihealuetta, tehtiin julkaisukanavahaku Julkaisuforumin JUFO-portaalin kautta. Jos tutkimuksen julkaisukanavan taso oli 1, 2 tai 3, luettiin tutkimuksen abstrakti. Jos abstrakti vahvisti käsitystä, että tutkimus on soveltuva tämän kirjallisuuskatsauksen lähteeksi, tutkimus lisättiin mahdollisiin lähteisiin. Jos julkaisun taso oli 0 tai sille ei löytynyt tasoa tai abstraktin perusteella tutkimus meni oman aihealueen ohi, teksti sivuutettiin. Kokotekstit haettiin ensisijaisesti Google Scholarin kautta. Jos kokotekstiä ei löytynyt Scholarista tai siihen ei ollut käyttöoikeutta julkaisijan omilla sivuilla Oulun yliopiston kirjastoportaalin kautta, se sivuutettiin. Teksti sivuutettiin myös, jos sitä ei ollut saatavilla englanninkielisenä.

Tekstien luotettavuus arviointiin enimmäkseen JUFO-tason, asiasisällöstä huokuvan asiantuntijuuden ja yleisen kieliasun perusteella. Viittausten määrään ei kiinnitetty niin paljon huomiota, koska aiheeseen liittyvät tekstit olivat vielä suhteellisen tuoreita. Google Scholarissa huomiottiin hakutuloksista vain ensimmäiset kolme sivua. Varsinaiseksi lähdemateriaaliksi valittiin lähteet, jotka oli tämän tutkimuksen kirjoitushetkellä julkaistu 10 vuoden sisällä. Varsinaisella lähteellä tarkoitan lähteitä, jotka täyttävät tieteellisen lähteen tunnusmerkit, ja joihin tämä kirjallisuuskatsaus pääosin perustuu.

3. Tekoäly tukena musiikkituotannossa

Tässä luvussa käsitellään aikaisempaa tutkimusta musiikin ja tekoälyn yhdistämisestä. Aluksi kerrotaan siitä, miten musiikkituotantoprosessi on muuttunut ajan saatossa ja tunnistetaan nykyaikaisen musiikkituotannon vaiheet. Tämän jälkeen käsitellään lyhyesti luovuutta ja sitten siirrytään pääasiaan eli miten tekoälyä voi hyödyntää musiikkituotannon eri vaiheissa nykytutkimuksen valossa.

3.1 Musiikkituotantoprosessi ennen ja nyt

Deruty ym. (2022) käsittelevät artikkelissaan musiikkituotannon muutosta ja tekoälyteknologioiden hyödyntämistä nykymusiikissa. He kertovat musiikkituotannon muuttuneet merkittävästi vuosisatojen aikana. Heidän mukaansa musiikkituotantoprosessi oli hyvin lineaarinen ja vähävaiheinen 1700- ja 1800-luvuilla:

Musiikki sävelletään → musiikki esitetään → lopputuotetta kuunnellaan

Derutyn ym. (2022) mukaan ennen vanhaan säveltäjän ja esiintyjän/muusikon roolit olivat tyypillisesti erilliset. Säveltäjä kirjoitti sävellyksen nuottikirjoituksella, jonka jälkeen yksi tai useampi esiintyjää soitti sävellyksen yleisölle edeten nuottikirjoituksen mukaan (Deruty ym., 2022). Tänä päivänä prosessi on hyvin erilainen. Äänentallentamiseen ja prosessointiin liittyvien teknologioiden yleistyminen ja halpeneminen 1980-luvulla mahdollisti sen, että sekä ammattilais- että amatööriartistit pystyivät hyödyntämään tuotannossaan enemmän uutta studioteknologiaa, kuten syntetisaattoreita, rumpukoneita, äänikirjastoja ja virtuaalisia instrumentteja (Deruty ym. 2022). Edellä mainittuja tekniikoita käytetään nykymusiikissa runsaasti; äänitettyjä instrumentteja tuetaan studioteknologian avulla tai kappaleet tehdään jopa alusta loppuun pelkästään tietokonein (Deruty ym. 2022). Teknologian kehitys on muuttanut musiikkituotantoprosessia. Deruty ym. (2022) jakavat nykymusiikin tuotantoprosessin vaiheisiin seuraavanlaisesti:

Musiikki sävelletään → esitetään → äänitetään → muokataan → miksataan → masteroidaan → lopputuotetta kuunnellaan

Nykymusiikin tuotannossa ei välttämättä edetä lineaarisesti, vaan vaiheissa voidaan palata taaksepäin useasti ja iteroida tehtyä työtä (Deruty ym. 2022). Derutyn ym. (2022) mukaan kehitys musiikkituotannossa on johtanut siihen, että rajat eri roolien välillä ovat sumentuneet. Ts. on vaikea määrittää tehtäviä, jotka olisivat ainoastaan säveltäjän, muusikon, ääni-insinöörin ja tuottajan tehtävissä (Deruty ym. 2022). Derutyn ym. (2022) mukaan nykyaikaisessa populäärimusiikissa kappale tai albumi on tyypillisesti luovan ryhmän tekemä, mutta on myös mahdollista, että vastuu jokaisesta musiikkituotannon vaiheesta on ainoastaan yhden henkilön harteilla. Tämän voi nähdä hyvänä perusteena tekoälyn hyödyntämiselle musiikkituotannossa; tekoälyn avulla voidaan tukea sooloartistia, jotta hän voisi olla musiikillisesti luovempi ja tuotteliaampi.

Musiikkituotannon ala on kehittynyt pikkuhiljaa ja uudet teknologiat ovat tuoneet alalle kehittyneitä tekniikoita musiikkituotannon avuksi (Mukherjee ym., 2020). Käsitellään seuraavissa luvuissa luovuutta sekä tekoälytutkimuksen nykyistä antia musiikkituotannolle ja sen eri vaiheille. Musiikin esitys ja äänitys jätetään käsittelemättä, koska tutkielmassa keskitytään vahvasti tietokonelähtöiseen tapaan tehdä musiikkia. Musiikin muokkaus jätetään myös pois. Deruty ym. (2022) viittaavat muokkaamisella oletettavasti yksittäisen äänen prosessointiin, jonka tavoitteena on muokata äänestä halutunlainen. Puhtaasti äänen muokkaamista käsittelevää tutkimusta ei löytynyt, mutta muokkaaminen on kuitenkin toimenpiteenä melko lähellä miksausta ja masterointia, joissa myös prosessoidaan ääntä erilaisilla efekteillä, joten tutkimustieto näistä kahdesta vaiheesta pätee myös

muokkaamiseen. Käsittelemättä jätetään myös lopputuotteen kuuntelu, koska se ei tapahdu musiikkituottajan toimesta studioympäristössä, johon tämä tutkimus on rajattu, vaan se on yleisön toimintaa. Käsiteltäväksi jää siis säveltäminen, miksaaminen ja masterointi.

3.2 Mitä luovuus on ja voiko tekoäly olla luova?

Carnovalini ja Rodà (2020) kertovat, että luovuus mielletään yleisesti yksilön kyvyksi luoda jotain uutta, jota ei ole ollut vielä aikaisemmin olemassa. He tarkentavat, että yleensä tällaiset uudet luomukset saavat alkunsa sellaisista konsepteista, jotka ovat olleet jo olemassa, mutta joita kukaan ei ole tajunnut yhdistää. Luovuus on subjektiivista (Fernández & Vico, 2013). Tietokoneet on perinteisesti mielletty deterministiseksi eli ne antavat aina samalla syötteellä saman, odotetun lopputuloksen, mutta voivatko tietokoneet olla luovia ja toimia ikään kuin vastakohtaisesti deterministisyydelle (Carnovalini & Rodà, 2020)? Fernández ja Vico (2013) tuovat artikkelissaan esiin vanhaa perua olevan väitteen, jonka mukaan mikään Turing-yhteensopiva formalismi ei pysty jäljentämään ihmisluovuutta aidosti. Toisin sanoen, musiikillista luovuutta ei voi laskea tietokoneella, mikä estää tietokoneita imitoimasta ihmissäveltäjiä täydellisesti (Fernández ja Vico, 2013). Laskennallinen luovuus on ala, jonka toimijat ovat kiinnostuneita selvittämään miten ihmisen luovuus toimii ja missä määrin sitä voidaan jäljitellä tietokonejärjestelmien, mukaan lukien tekoälyn, avulla (Carnovalini & Rodà, 2020). Ihmisen luovuuden tutkiminen on aivan oma aiheensa, mutta on hyvä sisäistää mitä luovuus on, jotta ymmärtää tekoälyn ja musiikin yhdistämiseen liittyvän haastavuuden. Musiikkihan on luovuuden aikaansaama tuote.

3.3 Tekoäly säveltämisen tukena

Säveltäminen on sointukulkujen, melodioiden, bassolinjojen ja rumpubiittien luomista sekä näiden kaikkien yhteensovittamista. Seuraavissa alaluvuissa käsitellään miten tekoälyä voi hyödyntää näiden musiikillisten elementtien luomisprosesseissa.

3.3.1 Sointurakenteen analysointi

Musiikillisen sävellyksen kontekstissa päämelodia on keskeisessä osassa, mutta melodiaa tukevan taustamusiikin vaikutusta ei pidä vähätellä (Mukherjee ym., 2020). Taustamusiikin rooli on äärimmäisen tärkeä, koska se vastaa kappaleen tunnelmasta ja harmonisesta sisällöstä (Mukherjee ym., 2020). Yksi taustamusiikin rakennuspalikoista on sointu, joka koostuu kahdesta tai useammasta samaan aikaan soitetusta sävelestä (Mukherjee ym., 2020). Musiikkiin luodaan variaatiota soittamalla toisiinsa luonnollisesti johtavia sointuja peräkkäin, mutta variaatiota voi luoda edelleen hyödyntämällä sointujen inversioita (Mukherjee ym., 2020). Inversiolla tarkoitetaan saman soinnun soittamista niin, että sointuun kuuluvien nuottien järjestystä muutetaan eli yhtä tai useampaa soinnun säveltä soitetään eri oktaavikorkeudella verrattuna perussointuun (Mukherjee ym., 2020). Mukherjeen ym. (2020) mukaan inversioiden tunnistaminen on äärimmäisen tärkeää sävellyksen automaattiselle analysoimiselle ja transkriboimiselle. Musiikissa transkriboinnilla tarkoitetaan sointujen ja melodioiden merkitsemistä sellaiseen muotoon, jonka ihminen ymmärtää, ja jonka perusteella ihminen osaa soittaa saman musiikin uudelleen. Perinteiselle muusikolle nuottikirjoituksen lukeminen on tuttua, mutta nykyaikaiselle musiikkituottajalle MIDI (= Musical Instrument Digital Interface) on kenties tutumpi. MIDI on viestintäprotokolla, jonka kautta syntetisaattorit, tietokoneet ja muu laitteisto voivat viestiä, ohjautua ja synkronoitua toistensa kanssa (ASA, s.a.). MIDI-data visualisoidaan käyttäjälle DAW:ssa (= Digital Audio Workstation) usein näkymänä, jossa X-akseli kuvaa aikaa ja Y-akseli sävelkorkeutta. MIDI ei pysty varsinaisesti välittämään ääntä, vaan ainoastaan sävelkorkeuteen ja musiikin intensiteettiin liittyvää dataa (Liu, 2022).

RNN (= Recurrent neural network) on neuroverkkoarkkitehtuuri, joka on tehokas työkalu reaaliaikaisen, kontekstuaalisen tiedon käsittelyyn (Smagulova & James, 2019). RNN-arkkitehtuurin yleinen tarkoitus on mallintaa sekventiaalista dataa, ja ääni on luonteeltaan sekventiaalista (Jiang ym., 2017). Näin ollen RNN sopii hyvin musiikin käsittelyyn. LSTM on RNN:n alatyyppejä, jolla on sisäinen muisti ja multiplikatiivisia portteja (Smagulova & James, 2019). LSTM:t pystyvät usein "muistamaan" pidempiä kaavoja ja mallintamaan niitä paremmin kuin puhtaasti RNN-pohjaiset neuroverkot (Jiang ym., 2017). Mukherjeen ym. (2020) tutkimuksessa esitellään LSTM-RNN-pohjainen järjestelmä, joka osaa tunnistaa sointumuotoja äärimmäisen lyhyistä nauhoituksista. Mukherjeen ym. (2020) tutkimuksen tulokset vaikuttavat lupaavilta. Heidän järjestelmänsä onnistui tunnistamaan sointumuodot oikein 99,47 % tarkkuudella 40 572 kitaranauhoituksesta, jotka oli nauhoitettu tavanomaisilla varusteilla tavallisessa huoneympäristössä. Heidän mukaansa nauhoitukset tehtiin näin realistisen käyttötilanteen simuloimiseksi; käyttäjät olettavat, että äänipohjaiset sovellukset toimivat myös kannettavilla laitteilla, joiden mikrofonit eivät välttämättä kykene taltioimaan ääntä niin laadukkaasti. Mukherjeen ym. (2020) mukaan inversioiden automaattinen tunnistaminen auttaa taustamusiikin ja päämelodioiden generoimisessa sekä musiikkikappaleiden sointupohjaisessa hakemisessa. He implikoivat musiikkiopiskelijoiden hyötyvän tällaisista työkaluista, koska ne mahdollistavat oppimisen sävellysten analysoinnin kautta.

3.3.2 Melodian generointi

Melodia on peräkkäisistä sävelistä koostuva kappaleen osa, joka jää yleisön päähän ja tekee kappaleesta muistettavan. Alaeddine ja Tannoury (2021) kertovat artikkelissaan koneoppimisen hyödyntämisestä ihmismäisen musiikin tuottamisessa. Heidän menetelmänsä tuottaa musiikin MIDI-muodossa ja jäljittelee ammattilaispianistin ajoituksia ja painallusten voimakkuutta. Menetelmä pohjautuu Google Magentaan, joka mahdollistaa mm. melodioiden generoimisen kokonaan itsenäisesti ja olemassa olevan melodian täydentämisen (Alaeddine & Tannoury, 2021). Alaeddine ja Tannoury (2021) kertovat, että heidän työkalunsa kykenee generoimaan melodiarakenteita nopeasti ja todella tarkasti tiettyyn rajaan saakka. Jos syötetyn MIDI-tiedosto on liian monimutkainen, malli ei osaa ennustaa sentyyppistä musiikkia tarkasti (Alaeddine & Tannoury, 2021). Kaksikon generoima ja analysoima MIDI-näytteiden määrä, 25, on kohtalaisen alhainen, minkä he ovat itsekin tiedostaneet artikkelissaan. Heidän mukaansa se on kuitenkin tarpeeksi paljon kertomaan mallin vahvuuksista ja heikkouksista. Malli osaa ennustaa parhaiten elektronista musiikkia ja pelimusiikkia, ja huonoiten jazz- ja lähi-itämaista musiikkia (taulukko 1). Länsimaalainen musiikki on ennustettavaa, koska se noudattaa musiikkiteoriaa melko tiukasti. Lähi-itämainen musiikki ja Jazz-musiikki tuottavat haastetta mallille, koska Lähi-itämaisessa musiikissa on neljäsosasävelaskelia, joita länsimaalaisessa musiikissa ei ole, ja Jazz-musiikissa on monimutkaisia sointukulkuja sekä nuotteja, jotka saatetaan soittaa yllättävästi ei-tahdissa (Alaeddine & Tannoury, 2021).

Liun (2022) esittelee Multi-Track-Music-GAN-mallin. GAN-arkkitehtuurit (= Generative Adversarial Network) käyttävät syväoppimista generatiiviseen mallintamiseen ja ne koostuvat pääasiassa kahdesta syväoppimismallista: generaattorista ja erottelijasta (Shahriar, 2022). Generaattori opetetaan tuottamaan uusia datapisteitä annetusta kohinasta ja erottelija tunnistaa datapisteet joko aidoiksi tai väärennetyiksi (Shahriar, 2022). GAN:t voivat oppia generoimaan uusia esimerkkejä automaattisesti tietyn datajoukon perusteella, mikä on osoittautunut hyödylliseksi piirteeksi tietokonepohjaisen taiteen luomisessa (Shahriar, 2022). Liun (2022) mallin yksi käyttökohde on melodioiden generointi, ja mallin avulla onnistuttiinkin luomaan sulavia ja esteettisiä musiikkiraitoja. Hänen mukaansa 62,8 % kuuntelijoista ei huomaa eroa oikean ja mallin generoiman melodian välillä. Joissain mallin generoimissa musiikkiraidoissa kuitenkin on liikaa satunnaisia nuotteja, jotka eivät kuulosta miellyttävältä ihmiskorville (Liu, 2022). Mallia täytyy parantaa, jotta vältettäisiin generoitujen tuotosten manuaaliset muokkaukset, kun haetaan tietyn tyylistä estetiikkaa (Liu, 2022).

Taulukko 1. 25 generoidun näytteen analysointi viidessä eri genressä (Mukailtu lähteestä Alaeddine & Tannoury, 2021).

GENRE	Rytmillinen tarkkuus	Sävellaji-tarkkuus	Tempo-tarkkuus	Näytteet
Elektroninen musiikki	100 %	80 %	100 %	5
Jazz	40 %	80 %	40 %	5
Klassinen musiikki	80 %	80 %	100 %	5
Lähi-itämainen musiikki	80 %	20 %	60 %	5
Pelimusiikki	100 %	80 %	100 %	5

3.3.3 Bassolinjan generointi

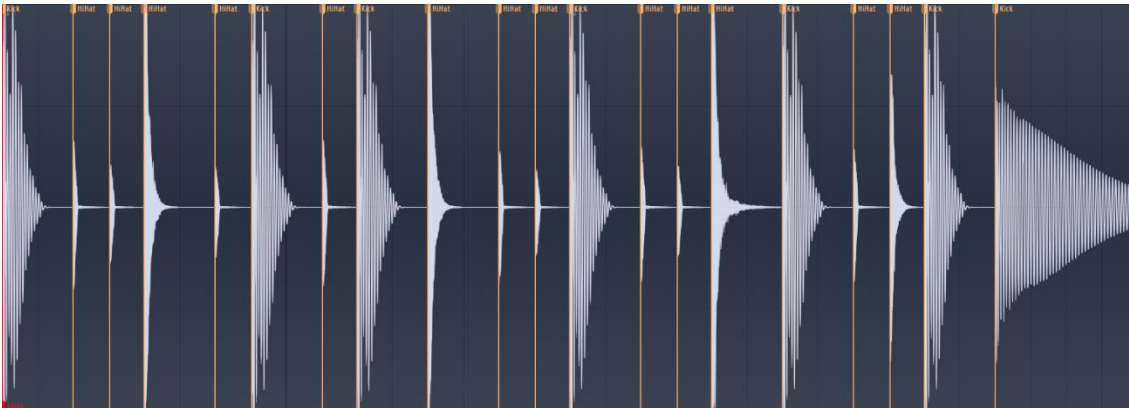
Bassolinjat soivat matalilla taajuuksilla ja ne eivät tavallisesti ole yhtä esillä kuin melodiat. Grachten ym. (2020) esittelevät tutkimuksessaan VAE-pohjaisen (= Variational Autoencoder) BassNet-syväoppimismenetelmän, joka pystyy generoimaan bassokitarasäestyksiä kontekstin eli yhden tai useamman syöteääniraidan perusteella. He kertovat BassNetin yhdistävän tiheää ja harvaa ehdollistamissignaalia bassolinjojen generointiin. Syötteen rytmiset ja harmoniset osat, kuten kitarat, kosketinsoittimet ja rummut, muodostavat tiheän signaalin (Grachten ym. 2020). Mallin harva signaali taas muodostuu käyttäjän hallittavissa olevan kaksiulotteisen koordinaatiston perusteella (Grachten ym. 2020). Grachtenin ym. (2020) mukaan BassNet on ensimmäinen musiikkigeneraattori, joka käyttää moduloimiseen latentiavaruutta. Moduloimisella tarkoitetaan jonkin jaksollista värähtelyä kuvaavan parametrin vaihtelemista (ASA, s.a.)

Grachten ym. (2020) mukaan BassNet tarjoaa mielenkiintoisia käyttötapauksia musiikkituotantoon: BassNet ei luo bassolinjoja tyhjästä vaan käyttäjän syöttämän materiaalin perusteella ja syötetty musiikkimateriaali ei päästä jättäjäisesti lopputulosta vaan käyttäjä voi luoda variaatioita ja muokata bassolinjan luonnetta kaksiulotteista ehdollistamissignaalia muokkaamalla. He tuovat hyvänä puolena esiin myös sen, että mallin opettamiseen ei tarvita kommentoitua dataa, koska harva ehdollistamissignaali oppii; muusikko voi opettaa mallia haluamallaan musiikilla oli se sitten hänen omaa musiikkiaan tai jonkin erityisen alagenren musiikkia. BassNet ei kuitenkaan ole täydellinen. Musiikkituottajien palaute on ollut, että he suosivat seikkailukkaampia generointeja, vaikka niissä olisikin virheitä (Grachten ym. 2020). Sen sijaan että musiikkituottajat käyttäisivät BassNetiä valmiiden, harha-askellettömien ääniraitojen generoimiseen, he haluavat käyttää sitä inspiroitumiseen (Grachten ym. 2020). Tämä siis tukee sitä ajatusta, että tekoälyn halutaan tukevan musiikillisesta luovuutta eikä tekevän kaikkea, kuten myös Alaeddine ja Tannoury (2021) kertovat omassa tutkimuksessaan.

3.3.4 Rumpubiittien analysointi, generointi ja keinotekoiset rummut

Länsimaisessa musiikissa, kuten rockissa, jazzissa ja tanssimusiikissa, rumpujen tehtävä on korostaa musiikkikappaleen rytmistä rakennetta ja tukea sen jakamista eri osiin (Wu ym., 2017). Wu ym. (2017) kuvaa rumpujen eroavan sävelkorkeudella soitetuista instrumenteista siinä, että ne eivät tyypillisesti tuo musiikkiin melodista tai harmonista arvoa. Lyömäsoittimina niillä ei siis yleensä ole selkeää säveltä, mutta sen sijaan niiden piirre on selkeä transientti, jonka voi havaita piikkinä äänenvoimakkuudessa (kuva 1) (Wu ym., 2017). Melodisilla instrumenteilla transientti ei ole niin

voimakas (kuva 2). On kuitenkin olemassa poikkeuksia, kuten piano ja kitara, joilla on transienttimaisia piirteitä (Wu ym., 2017).



Kuva 1. Rumpubiitti havainnollistettuna aaltomuotoisena näkymänä. Transientit havaittavissa helposti.



Kuva 2. Syntetisaattorilla soitettu bassolinja havainnollistettuna aaltomuotoisena näkymänä. Transientit eivät ole niin voimakkaita kuin rummuissa.

Wu ym. (2017) käsittelevät tutkimuksessaan ADT:tä (= Automatic Drum Transcription) ja siihen liittyvää aikaisempaa tutkimusta. ADT-järjestelmät keskittyvät tunnistamaan transienttipiirteiden omaavia, impulssimaisia ääniä eli rumpuja nauhoitteista ja yrittävät luokitella niitä eri rumputyyppeihin, kuten bassorumpuihin, tomtomeihin ja hihateihin (Wu ym., 2017). Wu ym. (2017) kertovat, että ADT:n yksi käyttökohde on musiikkituotanto, jossa sitä voi käyttää sekä rumpujen jälkikäteen korvaamiseen ääniraidassa että mikrofoniin välisten äänivutojen vähentämiseen rumpuäänitteistä, mistä esimerkkinä Drumatom-ohjelmisto. Wu ym. (2017) kertovat, että näiden lisäksi ADT saattaisivat mahdollistaa sisältöperustaisten lähestymistapojen käytön tunnistettujen äänten korvaamiseen, mikä parantaisi tietokoneavusteisen säveltämisen tai jopa automaattisen musiikin säveltämisen tehokkuutta. Jos monoääniset rumputallenteet voitaisiin purkaa luotettavalla transkriboinnilla yksittäisiksi rumpuiskuiksi, ADT voisi myös avustaa remixien teossa eli kappaleiden uudelleenmiksauksessa, mihin esim. Regroover-ohjelmisto pyrkii (Wu ym., 2017).

Wun ym. (2017) yksityiskohtaiset analyysit ja vertailut hallituissa olosuhteissa näyttävät, että RNN-pohjaiset menetelmät vaikuttaisivat lupaavimmilta lähestymistavoilta ADT:hen, ja niitä suositellaan etenkin silloin, kun on mahdollista käyttää suurta ja monipuolista datasettiä, joka on kommentoitu laadukkaasti. ADT-tekniikat toimivat yleisesti hyvin, kun ne analysoivat pelkkiä rumpuäänitteitä, mutta äänitteet, joissa on rumpujen lisäksi melodisia elementtejä, tuovat haasteita (Wu ym. 2017).

Wun ym. (2017) mukaan ADT-tutkimus vaatii vielä paljon työtä. Heidän mukaansa ADT-tutkimus on riippuvainen vapaasti käytettävissä olevasta datasta, ja tällaista korkealaatuista dataa tarvitaan lisää ADT-tutkimuksen edistämiseksi. Wun ym. (2017) mukaan tutkimusta saataisiin vietyä paremmin eteenpäin, jos käytössä olisi avoin arviointikanava, jossa uusimpia edistysaskeleita jaettaisiin. He kritisoivat, että suurin osa ADT-menetelmistä keskittyy vain muutaman keskeisen rumputyypin havaitsemiseen, minkä vuoksi niitä täytyisi kehittää havaitsemaan myös muun tyyppisiä rumpuja. He tuovat esiin myös seuraavat kehitysideat: menetelmien pitäisi huomioida rumpujen dynamiikka eli äänenvoimakkuuden vaihtelu ja menetelmiin pitäisi implementoida esikäsitteilytekniikoita, jotka tukahduttavat epäolennaisia elementtejä ja selkeyttävät rumpujen ääniä ennen varsinaista ADT-prosessointia.

ADT-menetelmät ovat oiva työkalu rumputyöskentelyn ja siten musiikkituotannon tehostamiseen. Riippuen musiikkituottajan työnkulusta, generoivat menetelmät saattavat olla luovuuden tukemiseen parempi vaihtoehto. Deruty ym. (2022) tuovat esiin tutkimuksessaan muutamia eri rumpuäänisyntetisaattoreita. He kertovat Planet Drums -ohjelmistosta, joka VAE-arkkitehtuuriin perustuvana antaa käyttäjälleen mahdollisuuden löytää erilaisia rumpuääniä latentiavaruuden avulla. He kertovat myös DrumGAN:sta ja Impact Drums:sta, jotka taas perustuvat GAN:in. Edellä mainituista rumpugeneraattoreista DrumGAN on ehdollistettu havaintopiirteisiin, joita voidaan käyttää kontrollina (Deruty ym., 2022). DrumNet on BassNetin tapaan generaattori, joka ottaa huomioon kontekstin, johon se generoi rumpuraitaa; se panee merkille olemassa olevat ääniraidat, kuten kitarat ja bassolinjan, ja adaptoituu niiden tempoon ja rytmiin (Deruty ym., 2022). DrumNetissä käyttäjät voivat kokeilla erilaisia rytmillisiä variaatioita latentiavaruuden avulla (Deruty ym., 2022).

3.4 Tekoäly miksaamisen ja masteroinnin apuna

Miksaamisella tarkoitetaan ääniraitojen muokkaamista, jossa säädellään instrumenttien äänenvoimakkuuksia suhteessa toisiinsa ja sovelletaan erilaisia signaalinkäsittelytekniikoita, jotta saataisiin aikaiseksi miksattu kokonaisuus, jossa jokainen elementti kuulostaa miellyttävältä yhdessä (Sterne & Razlogova, 2021). Perusteellinen miksaaminen voi viedä paljon aikaa, kuten voi masterointikin. Masterointi on musiikkituotannon viimeinen vaihe. Siinä tarkoituksena on viimeistellä musiikkiteos eli korjata mahdolliset virheet, tuoda musiikin taajuudet ja äänenvoimakkuus standarditasolle ja varmistaa, että musiikki kuulostaa hyvältä eri formaateissa (MP3, WAV, CD jne.) ja kuunteluympäristöissä (klubit, autot, televisiot, kuulokkeet, älypuhelimet jne.) (Birtchnell, 2018). Miksaaminen ja masterointi ovat samankaltaisia toimenpiteitä ja niissä käytetään samoja työkaluja DAW:n sisällä. Nykyaikaisessa musiikkituotannossa miksaaminen tapahtuu pääasiassa musiikkituottajan toimesta, mutta on mahdollista, että miksausinsinööri ottaa kappaleen vielä käsittelyyn ennen masterointia.

Puhtaasti miksaamista ja tekoälyä käsittelevää tutkimusta on vähän ja syynä voi yksinkertaisesti olla aiheen monimutkaisuus. Sterne ja Razlogova (2021) kertovat artikkelissaan LANDR-nimisestä mediayrityksestä ja sen alkutaipaleesta. Kun yritys vielä alkuaikanaan tunnettiin nimellä MixGenius, yrityksen sisällä huomattiin, että miksaaminen ei ole koneoppimisella ratkaistavissa oleva ongelma (Sterne & Razlogova, 2021). Koneoppiminen ei tässä kontekstissa tuota johdonmukaisia ja oikeasuhteisia tietokokonaisuuksia, joita voitaisiin verrata helposti keskenään, joten se ei ole altis LANDR:n kaltaiselle tietokoneistetulle luokittelulle (Sterne & Razlogova, 2021). Pelkästään ääniraitojen lukumäärien vaihtelu musiikkiprojekteissa on mahdollisesti ääretön ja ei ole olemassa mitään yhteistä standardia tai formaattia genrejen sisällä tai välillä (Sterne & Razlogova, 2021). Miksaamiseen tarvittavan datasetin pitäisi olla niin massiivinen, että koneoppimisen käyttäminen miksausien tukena tuntui mahdottomalta (Sterne & Razlogova, 2021).

Moffat ja Sandler (2019) pohtivat tutkimuksessaan IMP:ä (= Intelligent Music Production), jonka yksi osa-alue on miksaaminen. He kertovat, että tapa, miten ääniraitoja sovitetaan yhteen

miksauksessa, riippuu todella paljon kaikista projektin ääniraidoista. Kuhunkin ääniraitaan sovellettava käsittely on hyvin subjektiivista ja luovaa, mikä tekee miksausprosessin laskennallisesta ymmärtämisestä hyvin vaikeaa (Moffat & Sandler, 2019). Moffatin ja Sandlerin (2019) mukaan ammattilaisen tekemän miksausprosessin simuloiminen vaatii todella tarkkaa ymmärrystä tämän tekemistä päätöksistä. Jos miksaaminen haluttaisiin automatisoida täysin datapohjaisesti, tarvittaisiin suuri määrä laadukasta dataa. On kuitenkin odotettavissa, että tällainen lähestymistapa toimisi lähinnä tyydyttävästi (Moffat & Sandler, 2019). Moffatin ja Sandlerin (2019) mukaan tyydyttäviä tuloksia antavat IMP-järjestelmät voivat kuitenkin olla opettavaisia ja auttaa amatööriartistiteja ensimmäisten tuotostensa kanssa kustannustehokkaasti sekä myös parantaa verkkoon ja suoratoistopalveluihin lähetettävän musiikin miksauslaatua.

Masterointi on perinteisesti ollut tehtävään erikoistuneen ammattilaisen suorittama vaihe, mutta kuten Deruty ym. (2022) kertovat, rajat tehtävien välillä ovat sumentuneet. Nykyään musiikkituottaja saattaa hoitaa masteroinninkin itse. Tällainen tee-se-itse-masterointi voi kuitenkin pilata albumin, koska musiikkituottaja saattaa viimeistelyllään luoda itselleen illuusion, että kappale olisi parempi, kun tosiasiallisesti hän on luonut ongelmia esim. vokaaleihin, rumpuihin tai kappaleen muihin tärkeisiin osiin (Sterne & Razlogova, 2021). Jos ihminen masteroi musiikkia, hänen täytyy kuunnella sitä kriittisessä kuunteluympäristössä (Birtchnell & Elliott, 2018). Tällaisen tilan rakentaminen on kallista, koska täytyy ottaa huomioon huoneen rakenne ja miten ääniaallot kimpoilevat siellä sekä panostaa kaiuttimiin ja muihin varusteisiin, jotka tuottavat äänen mahdollisimman totuudenmukaisesti. Tekoäly ei tarvitse tällaista ympäristöä (Birtchnell & Elliott, 2018). Tekoälymasterointi on siis kustannustehokas ratkaisu esim. muusikoille, joilla ei vielä ole taloudellista tukea takanaan (Sterne & Razlogova, 2021).

Masteroinnin saralla monet tekoälyä ja musiikkia käsittelevät tutkimukset tuovat esiin jo aiemmin mainitun LANDR:n, jonka yksi tarjoama palvelu on automatisoitu, verkko- ja tekoälypohjainen musiikkimasterointipalvelu (Birtchnell, 2018. Sterne & Razlogova, 2021). Valmiiksi miksatun musiikkikappaleen masterointi on paljon rajatumpi ongelma kuin useiden ääniraitojen yhteen miksaaminen, joten koneoppimisen soveltaminen masterointiin ei ollut mahdollista (Sterne & Razlogova, 2021). LANDR yhdistää erilaisia algoritmeja, koneoppimista ja big data -analytiikkaa simuloidakseen ihmisen asiantuntemusta musiikin viimeistelystä yleisöä varten (Birtchnell, 2018). LANDR on opetettu erään levy-yhtiön musiikkikappaleiden masteroimattomilla ja masteroiduilla versioilla, jotka noudattavat vakiintuneita standardeja (Sterne & Razlogova, 2021). Yksinkertaistettuna LANDR toimii niin, että kappaleen digitaalisille ääniaalloille tehdään spektri- ja taajuusanalyysi, jossa ääniaaltoja vertaillaan suuressa datasetissä olevien kappaleiden keskiarvoihin, jotta voitaisiin määrittää signaalinprosessointiparametrit ilman ihmisen puuttumista prosessiin (Birtchnell, 2018). Birtchnellin (2018) mukaan palvelun helppokäyttöinen tiedostonlähetyksen ja -latauksen tekeminen artistista laadunvalvojan, jonka käytös opettaa LANDR:a, ja joka tavallaan saa LANDR:n imitoimaan ihmisen oppimista ja päätöksentekoa.

Birtchnellin (2018) mukaan LANDR-masterointi ei vielä yllä ihmisammattilaisen tasolle, minkä voi päätellä mm. siitä, että palvelun prosessoimat kappaleet eivät ole yltäneet samalla tavalla hittilistoille kuin perinteisesti prosessoitujen kappaleiden. LANDR ei siis vielä kykene masteroimaan musiikkia yhtä laadukkaasti ja miellyttävästi kuin ammattilaiset. Birtchnell (2018) kertoo alan ammattilaisten kokevan, että tekoäly ei ole uhka heidän työlleen. Uhan sijaan he ajattelevat tekoälyn tuovan helpotusta joihinkin äänimasteroinnin rutiinitehtäviin, kuten virheenkorjaukseen ja mediantiedostojen muodostukseen (Birtchnell, 2018). LANDR:a käyttäneet musiikkituottajat ovat kokeneet sen auttavan heitä miksausensa korjaamisessa; he prosessoivat kappaleensa LANDR:n läpi, kuuntelevat mikä siinä on pielessä, miksaavat uudelleen ja syöttävät sitten kappaleensa uudelleen tekoälyn läpi (Birtchnell, 2018).

4. Pohdinta

Tämän tutkielman päätavoitteena oli selvittää, miten tekoäly voi tukea musiikkituottajaa työssään. Ajatukseen tukemisesta sisältyy se, miten tekoäly voisi nopeuttaa työvaiheita ja ohjata kohti musiikkiteollisuuden standarditasoja sekä tuoda esiin artistin luovuutta, jotta hän voisi olla itsevarmempi musiikkinsa suhteen ja julkaista enemmän musiikkia. Aikaisemmasta tutkimuksesta kerrottiin musiikkituotannon vaiheiden kautta, joten on luonnollista jatkaa asian pohdintaa näiden samojen vaiheiden kautta.

4.1 Sävellyks

Kuten jo mainittiin aikaisemmin luvussa kaksi, tiedonhaun hakutuloksista pystyi heti päättelemään, että tekoälyn ja musiikin tutkimus on keskittynyt voimakkaasti säveltämiseen ja tarkemmin ottaen musiikin transkribointiin ja generointiin. Nämä menetelmät voivat auttaa merkittävästi musiikkituottajaa, jos ne vain toimivia hyvin eri käyttötarkoituksissaan. Mukherjee ym. (2020) esiin tuoma sointuinversioiden tunnistamistyökalu vaikuttaa lupaavalta, mutta työkalua oli testattu ainoastaan kitaralla. Ollakseen hyödyllinen, työkalu pitäisi opettaa havaitsemaan myös muita instrumentteja, kuten syntetisaattoreita, jotka ovat tyypillisiä nykyaikaisissa musiikkikappaleissa. Monipuolisemman instrumenttirepertuaariin havaitsemiin liittyvän ongelman he ovat toki itsekkin tiedostaneet. Toisaalta ollakseen aidosti hyödyllinen, työkalun pitäisi kyetä tunnistamaan sointuja myös kokonaisista kappaleista, joissa moni eri instrumentti soi samaan aikaan. Tämä menee kylläkin jo monimutkaiseksi, sillä työkalun pitäisi pystyä erottamaan kaikki instrumentit ja taustahälinät ensin toisistaan ennen sointujen tunnistamista. Eri transkribointimenetelmien hybridisaatio eli menetelmien tuominen samaan ohjelmistoon olisi kuitenkin mittaamattoman arvokasta musiikkituotannolle. Jos musiikkituottaja pystyisi purkamaan suosikkikappaleensa osiin tekoälypohjaisella musiikkianalysointityökalulla, hän näkisi kappaleen eri tavalla ja saattaisi oppia jotain. Nykymusiikissa trendinä on tehdä remix-versioita vanhoista hittikappaleista. Hybridityökalut voisivat tehostaa kappaleiden uudelleenmiksausta merkittävästi.

Koneoppimista hyödyntävien generointityökalujen avulla muusikko voi luoda sovituksia tai variaatioita lyhyessä ajassa, vaikka hänellä ei olisi tarvittavia taitoja (Sturm ym., 2019). Se, että artisti pystyy generoimaan useita melodioita ja sointukulkuja nopeasti, avaa mahdollisuuden erittäin luovaan ja tehokkaaseen työhön. Alaeddine ja Tannoury (2021) havaitsivat että eksoottiset ja monimutkaiset musiikkigenret ovat haastavia tekoälylle. Nykyaikaista populaarimusiikkia tekevän musiikkituottajan kannalta tämä puute ei ole merkittävä, koska populaarimusiikki on verrattain yksinkertaista ja tekoälygeneraattorit kykenevät tuottamaan tällaista musiikkia hyvällä tarkkuudella. Jos musiikkituottaja työskentelee paljon yksin ja ei saa toista perspektiiviä tuotantoonsa, on riskinä, että luovuus tyrehtyy. Tällaista luovuuden puutetta kutsutaan writer's blockiksi eli suomalaisittain kirjoituskammoksi. Jos musiikkituottaja käyttää tekoälyä tukenaan, hän pystyy vierottamaan itseään omista tottumuksistaan, haastamaan itseään ja inspiroitumaan (Sturm ym., 2019). Tekoälystä tulee muusikolle ikään kuin toinen muusikko, joka avustaa mielenkiintoisten tulosten luomisessa ja auttaa luomaan jotain sellaista, mikä ei muuten tulisi esiin (Sturm ym., 2019).

Erlaisia tekoälyllisiä menetelmiä sointujen ja melodioiden generointiin on valtavasti. Tässä tutkielmassa nostettiin esiin näistä vain muutamia mielenkiintoisia, joilla saattaisi olla mahdollisuus päätyä nykyaikaisen musiikkituottajan arsenaaliin. Menetelmät ovat lupaavia, mutta useat niistä ovat vielä tutkimusvaiheessa. Osaa menetelmistä ei ole julkaistu vapaaseen käyttöön, ja niistä on saatavilla vain esimerkkiääniraitoja. Joitain menetelmiä on julkaistu avoimeen käyttöön, mutta niiden käyttöönotto vaatii vielä taitoja, joita ei voi olettaa tavalliselta musiikkituottajalta.

4.2 Miksaus ja masterointi

Musiikin miksauseseen liittyvää tutkimusta on harmillisen vähän. Moffat ja Sandler (2019) ovat toiveikkaita älykkään miksaamisen suhteen, mutta he kuitenkin kertovat aiheesta hyvin konseptimaisesti mainitsematta mitään tiettyä markkinoilla olevaa ratkaisua. Kun ottaa tämän lisäksi huomioon sen, että LANDR ei aikoinaan taipunut miksaamiseen, voi tehdä johtopäätöksen, että nykyaikaiselle musiikkituottajalle tekoälyn hyödyntäminen miksausessa ei ole vielä ajankohtainen asia; tekoälyn ja miksausksen yhdistäminen on todella hankalaa ja aihe vaatii vielä merkittävästi tutkimustyötä ja erityisen laadukasta dataa. LANDR:n epäonnistuminen miksausessa osoittaa, että tekoälyjärjestelmät tarvitsevat sisäisesti johdonmukaisia tietokokonaisuuksia: datan osasten on oltava vertailukelpoisia keskenään, jotta järjestelmä voi toimia hyvin (Sterne & Razlogova, 2021).

Musiikkituottajalla on kuitenkin paremmat edellytykset käyttää tekoälyä musiikin masterointiin, vaikkakin markkinoilla ei ole paljon varaa valita. Äänimasterointia käsittelevissä tutkimuksissa on yhteistä, että niissä on nostettu esiin LANDR. Muut kilpailevat järjestelmät on mainittu vain ohimennen. Esim. Sterne ja Razlogova (2021) mainitsevat Izotopen ja Cloudbouncen. Heidän mielestään Izotopen masterointimenetelmä ei ole tarpeeksi dynaaminen ja reaaliaikainen, koska se asennetaan käyttäjän tietokoneelle ja se pyörii käyttäjän koneella. Cloudbouncesta he taas kertovat, että koneoppiminen ei ole siinä tarpeeksi merkittävässä roolissa. Tutkimuksen kannalta LANDR:in tosin liittyy epäkohta, että sen toiminnasta ei tiedetä tarkasti. LANDR on yksityinen yritys, jolla on tietenkin yrityssalaisuutensa, mikä estää tarkastelemasta miten palvelun masterointi toimii konepellin alla (Sterne & Razlogova, 2021). Musiikkimasterointia ja tekoälyä käsittelevää avointa tutkimusta on niukasti, joten olisi toivottavaa, että avoimessa tekoälytutkimuksessa huomiota annettaisiin myös masteroinnille.

4.3 Musiikkituottajien kokemukset tekoälystä

Derutyn ym. (2022) tutkimukseen osallistui 6 eri artistia. He kertovat, että erilaisten tekoälymenetelmien avulla artistit onnistuivat mm. generoimaan abstrakteja rumpuja, mielenkiintoisia rumpubiittejä ja erilaisia positiivisia sattumuksia. Heidän mukaansa menetelmät yllättivät monesti myönteisesti, kun ne taipuivat sellaisiin käyttötarkoituksiin, joihin niitä ei ollut suunniteltu. Eräs artisti kertoi menetelmien käyttämisen muuttaneen hänen työkulkuaan tuloskeskeisestä prosessikeskeiseksi. Musiikkituottajien kritiikki kohdistuu työkalujen käytettävyyteen; parametrien visualisoimattomuus saa aikaan sen, että tuloksia haetaan yrityksen ja erehdyksen kautta (Deruty ym., 2022). Nykyiset tekoälytyökalut toimivat musiikkituotannossa tyypillisesti niin, että ne antavat aloittamismateriaalia tai ehdottavat variaatioita olemassa olevien materiaalien perusteella (Deruty ym., 2022).

Tähän kirjallisuuskatsaukseen valikoitujen lähteiden perusteella musiikkituottajien mielipide tekoälyyn vaikuttaisi olevan myönteinen. Täytyy kuitenkin nostaa esiin se, että musiikkituottajien ajatuksia ei ole tuotu merkittävästi esiin näissä tutkimuksissa. Spesifiä tutkimusta, joka käsittelee musiikkituottajien suhtautumista tekoälyyn, ei valitettavasti ole olemassa. Lisäksi Derutyn ym. (2022) tutkimukseen ovat saattaneet päätyä sellaiset artistit, jotka ovat olleet jo valmiiksi tekoälymyönteisiä. Nykymusiikin piirteiden perusteella voisi kuitenkin olettaa, että tekoäly nähdään yleisesti mahdollisuutena eikä uhkana. Nykymusiikissa käytetään paljon samoja saundeja sekä syntetisaattoreiden ja efektien esiasetuksia. Toisten lainaaminen tyylin ja soundien muodossa on tavanomaista. Puhutaan ns. samplaamisesta, jossa käytetään toisten kappaleiden osia omissa kappaleissa. Alalla vallitsee myös remix-kulttuuri, jossa tehdään omia versioita toisten kappaleista. Musiikkialalla toisen luovan työn hyödyntäminen on hyväksyttävämpää kuin monella muulla luovalla alalla, joten miksi tekoälyn tuotosten hyödyntäminen nähtäisiin pahana asiana? Laajemman mielipiteen selvittäminen vaatisi kyselytutkimuksen tekemistä musiikkituottajille.

Deruty ym. (2022) kertovat, että tekoälyä hyödyntävät työkalut tarjotaan usein palveluna jonkin tietyn alustan kautta. Tekoälytyökalut eivät siis ole vielä täysin integroitua musiikkituottajan työympäristöön ja -tapoihin, minkä vuoksi tekoäly ei ole vielä niin laajasti käytössä musiikkituottajien keskuudessa (Deruty ym., 2022). Nykyisten tekoälymusiikkityökalujen ymmärtäminen, asentaminen ja itsenäinen käyttäminen ei ole helppoa, sillä ne edellyttävät yleensä teknistä asiantuntemusta ja koneoppimisen peruskäsitteiden tuntemusta (Deruty ym., 2022). Tällaiset työkalut täytyisi siis integroida jollain tapaa paremmin musiikkituottajan työnkulkuun. Musiikkituottajan työ tapahtuu pääasiassa DAW:n sisällä. Jos musiikkituottaja joutuu jatkuvasti poistumaan DAW:sta, se vaikuttaa työn tehokkuuteen. Kun kehitetään tekoälysovelluksia musiikkituotantoon, ne täytyy kehittää musiikkituottajan ehdoilla. VST (= Virtual Studio Technology) on standardiformaatti efekteille ja instrumenteille. VST-pluginit toimivat riippumatta siitä, mitä DAW:a musiikkituottaja käyttää, joten tekoälysovellusten kehittäminen tässä muodossa olisi tekoälyn käyttöönottoasteen kannalta suotavaa. Toinen vaihtoehto olisi integroida tekoälyä suoraan DAW:hin, kuten Abletoniin, FL Studioon ja Logic Prohon. Tämä vaatisi tietenkin paljon työtä DAW:a kehittävilta ohjelmistofirmoilta. Civit ym. (2022) kertovat, että monet alan kaupalliset toimijat ovat alkaneet kiinnittämään huomiota käyttöliittymiin ja DAW-integraatioon, mikä kertoo sen, että itse tekoälyteknologia alkaa olla tarpeeksi kypsää laajempaa käyttöönottoa varten.

Fridin, Gomesin ja Jinin (2020) mukaan käyttäjätutkimukset näyttävät, että on tärkeää tasapainottaa tekoälyn kontrolli ja ihmisen syöte. Käyttäjistä tuntuu, että musiikki ei ole ollenkaan heidän luomaansa, jos tekoälyllä on liikaa valtaa musiikkitekoprosessissa (Frid ym., 2020). Tämä on ristiriidassa syväoppimistutkimuksen kanssa, joka keskittyy vahvasti musiikin itsenäiseen generointiin (Frid ym., 2020). Vähäisestä syötteestä kokonaisia kappaleita generoivat työkalut ovat toki tärkeitä osoituksia äänipohjaisten tekoälyjärjestelmien potentiaalista, mutta ne eivät kuitenkaan sovi tyypillisiin studiossa tapahtuviin käyttötapauksiin (Deruty ym., 2022). Näissä käyttötapauksissa tekoälytyökalujen tuotokset ovat pieniä osasia, jotka sovitetaan yhteen suuremman rakenteen musiikillisten elementtien kanssa (Deruty ym., 2022). On tehtävä lisätutkimusta ihmisten ja tekoälysovelluksen välisestä suhteesta, jotta voidaan luoda viitekehys, jossa ihminen ja tietokone voivat hyödyntää parhaita ominaisuuksiaan (Civit ym., 2022).

4.4 Tekoälyn ja musiikin tulevaisuus

Musiikki on todella kaupallistettua ja isot, maailmanlaajuiset yhtiöt hallitsevat musiikkibisnestä. Civit ym. (2022) kertovat, että aiheita koskevat julkaisut tulevat sekä akateemisilta että yksityisiltä sektoreilta. Heidän mukaansa keskeinen kontribuutio tulee kuitenkin tekoälyalan pääpelaajilta, kuten Googlelta, OpenAI:lta ja Amazonilta, ja musiikkiryhtyksiltä, kuten Sonylta ja Spotifylta. Kaikkea ei kuitenkaan julkaista. Monista tämän hetken kaupallisista AI-järjestelmistä ei ole olemassa tieteellisiä julkaisuja (Civit ym., 2022). Koska kaikkea tutkimustietoa ei julkaista vapaaseen käyttöön, ei voi tarkasti tietää tekoälyn ja musiikin nykyisestä tilasta. Todellisuudessa tekoälyn ja musiikin tutkimus saattaa olla pidemmällä kuin mitä ulospäin näyttää. Tekoälyn ja musiikin yhdistäminen on nousussa, joten ei voi kieltää sitä, että alalla on kova kilpailu. On vain ajan kysymys kuka ehtii ensin ja kehittää merkittävän ja uutiskynnyksen ylittävän, musiikillisen tekoälysovelluksen. Tulevina vuosilta voi kuitenkin odottaa, että tekoälyn käytöstä musiikkituotannossa tulee valtavirtaa ja rutiinia ja sitä hyödynnetään populaarimusiikin lisäksi sosiaalisen median alustoilla (Avdeeff, 2019). Tavat, joilla nämä teknologiat voivat nopeuttaa musiikkituotannon eri vaiheita tekevät niistä korvaamattoman arvokkaita musiikkituotannolle, erityisesti popmusiikille ja sen kulttuurille, jossa vallitsee ohimenevyys ja musiikkiteosten nopea julkaisutahti (Avdeeff, 2019).

Nykyiset tekoälyyn perustuvat generaattorit voidaan integroida erittäin hyvin osaksi monenlaisia musiikkiluontirutiineja (Civit ym., 2022). Täytyy vain tutkia, miten tekoälysovellukset integroitaisiin parhaiten näihin rutiineihin ja tehtäisiin käyttäjäystävällisiksi. Koska monet näistä generaattoreista ovat uskomattoman nopeita, ne voivat myös vapauttaa muusikon rutiinitehtävistään, mikä antaa

muusikolle mahdollisuuden keskittyä musiikin ja äänien valitsemiseen sekä niiden yhteensovittamiseen (Civit ym., 2022). Tekoälyn vaikutus musiikkituotantoon on varmaa, mutta tulevat vuodet näyttävät missä määrin se tulee vaikuttamaan musiikkialaan.

Haluan nostaa vielä esiin tämän tutkielman kirjoittamisen aikana julkaistun artikkelin, joka käsittelee Googlen MusicLM-mallia. MusicLM on generoiva malli, joka kykenee tuottamaan korkealaatuista musiikkia 24 kHz:nä ollen konsistentti usean minuutin ajan ja uskollinen alkuperäiselle ehdollistamissignaalille (Agostinelli ym., 2023). Malli on hyvin mielenkiintoinen, sillä Agostinellin ym. (2023) mukaan se kykenee tuottamaan korkealaatuista musiikkia pelkästään tekstisyötteestä ja se päihittää aikaisemmat ratkaisut äänenlaadussa ja tekstisyötteen noudattamisen tarkkuudessa. MusicLM voi myös muuntaa hyräillyn tai vihelletyn melodian sellaiseksi, kuin sen on tekstisyötteessä kuvaillut (Agostinelli ym., 2023) Ääninäytteiden perusteella malli on erittäin lupaava, mutta voi tietenkin olla, että julkaistut musiikkiesimerkit on valikoitu todella tarkasti. Joka tapauksessa Google MusicLM:n kyvyt implikoivat tulevaisuuden tekoälymallien tekevän konseptoinnista ja ideoinnista todella helppoa musiikkituottajille. Toisaalta tämän kyseisen mallin kyky generoida pitkäkestoista, hyvää musiikkia luo hieman utopistisen ajatuksen musiikkituottajan roolista: tarvitaanko ihmismusiikkituottajia enää tulevaisuudessa?

5. Yhteenveto

Nykyaikainen musiikkituotanto on käynyt läpi suuria muutoksia digitaalisen teknologian kehityksen myötä. Tekoälytutkimuksen edistysaskelteet ovat tuomassa seuraavia merkittäviä muutoksia musiikkialalle. Tekoälymenetelmät, erityisesti generoivat menetelmät, ovat yksi merkittävä kehityssuunta musiikkituotantoon liittyvässä tutkimuksessa. Generatiivisia menetelmiä voidaan käyttää säveltämisessä ja soundien luomisessa. Generointimenetelmät eivät vielä kykene täysin itsenäiseen työhön; ne eivät ole täydellisiä eikä niiden toisaalta tarvitsekaan olla, sillä musiikkituottajat haluavat muokata niiden tuloksia ja jättää työhön myös oman jälkensä. He haluavat käyttää tekoälyä inspiroitumiseen sen sijaan että tekoäly tekisi heidän työnsä kokonaan itse (Alaeddine ja Tannoury, 2021). Transkribointimenetelmien kehitys implikoi helpotusta musiikin oppimiseen ja remixien tekoon.

Tekoäly voi auttaa musiikkituottajia löytämään uusia ideoita ja ilmaisemaan itseään tehokkaammin. Vaikka tekoäly ei voisikaan koskaan saavuttaa ihmismäistä luovuutta täysin, tekoälyn ja musiikin tutkimus näyttää, että se voi olla hyödyllinen työkalu musiikintuotannossa ja luovuuden edistämässä. Tekoälytyökalut ovat kehittyneitä, mutta niiden saatavuus tavalliselle musiikkituottajalle on vielä rajoitettua; niitä ei ole integroitu musiikkituotannon työnkulkuun hyvin. Tekoälytyökalujen käyttöönotto vaatii usein erityistä asiantuntemusta, joka voi olla haasteellista monille musiikkituottajille (Deruty ym., 2022). Joitain työkaluja, kuten LANDR, on toki saatavilla helppokäyttöisinä, web-pohjaisina tilauspalveluina. Tarvitaan käyttäjätutkimusta, jotta saataisiin selville, miten tekoälytyökalut voitaisiin integroida musiikkituottajien työprosessiin mahdollisimman luonnollisesti. Miksaamiseen liittyy paljon muuttujia, mikä tekee tekoälyn hyödyntämisestä tässä osa-alueessa hankalaa. Miksaamiseen ja tekoälyyn liittyvää tutkimusta täytyy tehdä enemmän. Äänen masteroimiseen liittyvä tutkimus on myös liian pintapuolista, ja aihe tarvitsee lisää tutkimusta.

Maailma muuttuu ja musiikkituotanto muuttuu. Tekoälystä tulee pian uusi työkalu, jota ilman muusikot ja säveltäjät eivät voi elää (Liu, 2022). Musiikkituottajien yleinen mielipide tekoälystä vaikuttaisi olevan myönteinen, mutta tämä olisi hyvä varmistaa kyselytutkimuksilla. Se on varmaa, että oli ala mikä tahansa, uudet työskentelytavat luovat jonkin verran vastustusta. Itse näen, että musiikkituottajan olisi kyettävä muuttumaan maailman mukana ja ottamaan kaikki ilo irti tekoälyn tuomista mahdollisuuksista luovalle työlleen ja musiikilleen.

Lähdeluettelo

Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N., & Frank, C. (2023). MusicLM: Generating Music From Text. arXiv preprint arXiv:2301.11325.

Alaeddine, M., & Tannoury, A. (2021). Artificial Intelligence in Music Composition. In *Artificial Intelligence Applications and Innovations: 17th IFIP WG 12.5 International Conference, AIAI 2021, Hersonissos, Crete, Greece, June 25–27, 2021, Proceedings 17* (pp. 387–397). Springer International Publishing.

ASA Acoustical Society of America (julkaisuaika tuntematon). Acoustical Terminology Database. Haettu osoitteesta <https://asastandards.org/working-groups-portal/asa-standard-term-database/>

Avdeeff, M. (2019). Artificial intelligence & popular music: SKYGGE, flow machines, and the audio uncanny valley. In *Arts* (Vol. 8, No. 4, p. 130). MDPI.

Birtchnell, T. (2018). Listening without ears: Artificial intelligence in audio mastering. *Big Data & Society*, 5(2), 2053951718808553.

Birtchnell, T., & Elliott, A. (2018). Automating the black art: Creative places for artificial intelligence in audio mastering. *Geoforum*, 96, 77–86.

Carnovalini, F., & Rodà, A. (2020). Computational creativity and music generation systems: An introduction to the state of the art. *Frontiers in Artificial Intelligence*, 3, 14.

Chu, H., Kim, J., Kim, S., Lim, H., Lee, H., Jin, S., Lee, J., Kim, T., & Ko, S. (2022). An Empirical Study on How People Perceive AI-generated Music. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management* (pp. 304-314).

Civit, M., Civit-Masot, J., Cuadrado, F., & Escalona, M. J. (2022). A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends. *Expert Systems with Applications*, 118190.

Deruty, E., Grachten, M., Lattner, S., Nistal, J., & Aouameur, C. (2022). On the development and practice of ai technology for contemporary popular music production. *Transactions of the International Society for Music Information Retrieval*, 5(1).

Fernández, J. D., & Vico, F. (2013). AI methods in algorithmic composition: A comprehensive survey. *Journal of Artificial Intelligence Research*, 48, 513-582.

Frid, E., Gomes, C., & Jin, Z. (2020, April). Music creation by example. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-13).

Grachten, M., Lattner, S., & Deruty, E. (2020). BassNet: A variational gated autoencoder for conditional generation of bass guitar tracks with learned interactive control. *Applied Sciences*, 10(18), 6627.

Jiang, M., Yang, Z., & Zhao, C. (2017, October). What to play next? A RNN-based music recommendation system. In *2017 51st Asilomar Conference on Signals, Systems, and Computers* (pp. 356-358). IEEE.

Liu, W. (2022). Literature survey of multi-track music generation model based on generative

confrontation network in intelligent composition. *The Journal of Supercomputing* 79, 6560–6582.

Moffat, D., & Sandler, M. B. (2019, September). Approaches in intelligent music production. In *Arts* (Vol. 8, No. 4, p. 125). MDPI.

Mukherjee, H., Dhar, A., Ghosh, M., Obaidullah, S. M., Santosh, K. C., Phadikar, S., & Roy, K. (2020). Music chord inversion shape identification with LSTM-RNN. *Procedia Computer Science*, 167, 607-615.

Roads, C. (1985). Research in music and artificial intelligence. *ACM Computing Surveys (CSUR)*, 17(2), 163–190.

Shahriar, S. (2022). GAN computers generate arts? a survey on visual arts, music, and literary text generation using generative adversarial network. *Displays*, 102237.

Smagulova, K., & James, A. P. (2019). A survey on LSTM memristive neural network architectures and applications. *The European Physical Journal Special Topics*, 228(10), 2313-2324.

Sterne, J., & Razlogova, E. (2021). Tuning sound for infrastructures: artificial intelligence, automation, and the cultural politics of audio mastering. *Cultural Studies*, 35(4-5), 750-770.

Sturm, B. L., Ben-Tal, O., Monaghan, Ú., Collins, N., Herremans, D., Chew, E., Hadjeres, G., Deruty, E. & Pachet, F. (2019). Machine learning research that matters for music creation: A case study. *Journal of New Music Research*, 48(1), 36–55.

Wu, C. W., Dittmar, C., Southall, C., Vogl, R., Widmer, G., Hockman, J., ... & Lerch, A. (2018). A review of automatic drum transcription. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(9), 1457-1483.