

TUNTEISTA TEKOÄLYYN JA MUSIIKIKSI : GENERATIIVISEN PELIMUSIIKIN TUTKIMUS 2020-LUVULLA

Peppi Salminen
Kandidaatintutkielma
Musiikkitiede
Musiikin, taiteen ja kulttuurin tutkimuksen laitos
Jyväskylän yliopisto
Kevät 2024

JYVÄSKYLÄN YLIOPISTO

Tiedekunta Humanistis-yhteiskuntatieteellinen	Laitos Musiikin, taiteen ja kulttuurin tutkimus
Tekijä Peppi Salminen	
Työn nimi Tunteista tekoälyyn ja musiikiksi : Generatiivisen pelimusiikin tutkimus 2020-luvulla	
Oppiaine Musiikkitiede	Työn laji Kandidaatintutkielma
Aika Kevät 2024	Sivumäärä 25
<p>Tiivistelmä</p> <p>Pelimusiikkia tutkitaan yhä enemmän, ja pelimusiikin sopivuutta peleihin ja pelikokemukseen halutaan parantaa niin peliteollisuudessa kuin akateemisella soveltavalla tutkimuksella. Generatiivinen musiikki muodostuu järjestelmällisesti reaaliajassa, ja sitä pidetään potentiaalisena työkaluna pelitilanteeseen mukautuvan pelimusiikin sävellyksessä ilman, että musiikin laatu muutoin heikkenisi.</p> <p>Tutkielman tarkoituksena on selvittää, millaisia generatiivisen pelimusiikin tuottamisen malleja ja järjestelmiä on tutkittu ja kehitetty 2020-luvulla sekä kuinka ne eroavat toisistaan. Lisäksi tutkielmassa pohditaan tekoälyavusteisesti sävelletyn musiikin ongelmallisuutta esimerkiksi tekijänoikeuksien osalta. Aineistona on käytetty 2020-luvulla aiheesta julkaistua kirjallisuutta, jota on haettu esimerkiksi pelimusiikin tutkimuksen ja informaatioteknologian julkaisuista.</p> <p>Tutkimuksen tuloksena löytyy seitsemän mallia tai järjestelmää, joita käytetään generatiivisen pelimusiikin luomiseen. Useimmat niistä pohjautuvat tekoälyyn, kuten neuroverkkoihin, sekä emotiomalleihin, kuten affektimalliin. Tutkimustulokset generatiivisen musiikin mallien ja järjestelmien toiminnasta ovat positiivisia, ja tekoälyllä sävellettyä pelimusiikkia aiotaan kehittää jatkossakin.</p>	
Asiasanat Generatiivinen musiikki, generatiivinen pelimusiikki, pelimusiikki, tekoälymusiikki, videopeli	
Säilytyspaikka Jyväskylän yliopisto	
Muita tietoja	

KUVIOT

Kuvio 1	AMS:n yksinkertaistettu toimintamalli (mukaillen Hutchings & McCormack, 2020, s. 276).....	11
Kuvio 2	PMGG:n yksinkertaistettu toimintamalli (mukaillen Washburn & Khosmood, 2020, s. 3).....	12
Kuvio 3	GMM:n yksinkertaistettu toimintamalli (mukaillen Xiang & Guo, 2021, s. 195).	13
Kuvio 4	Transformaattorin yksinkertaistettu toimintamalli (mukaillen Amaral, Baffa, Briot, Feijo & Furtado 2022).	14
Kuvio 5	LSTM RNN:n yksinkertaistettu toimintamalli (mukaillen Bhardwaj, Salim, Khan & Javadimasoudian 2022, s. 196).	15
Kuvio 6	PreGLAM-MMM:n yksinkertaistettu toimintamalli (mukaillen Plut, Pasquier, Ens & Tchemeube, 2022).	16
Kuvio 7	PAMG:n yksinkertaistettu toimintamalli (mukaillen Lopez Duarte, 2023, s. 43).	18

TAULUKOT

TAULUKKO 1	Yhteenvedo tutkielmassa esitellyistä 2020-luvun generatiivisen musiikin malleista ja järjestelmistä.....	20
------------	--	----

SISÄLLYS

1	JOHDANTO	1
2	TEOREETTINEN TAUSTA.....	3
2.1	Lyhyesti pelimusiikin teknologisesta kehittämisestä	3
2.2	Pelimusiikki ja interaktiivisuus.....	4
2.3	Generatiivinen musiikki.....	5
2.4	Aikaisempi generatiivisen pelimusiikin tutkimus	6
2.5	Generatiivisen musiikin käyttö peliteollisuudessa	8
3	GENERATIIVISEN PELIMUSIIKIN MALLIEN TUTKIMUS 2020-LUVULLA. 10	
3.1	Mallit ja järjestelmät	10
3.1.1	AMS (Adaptive Music System)	10
3.1.2	PMGG (Procedural Music Generation in Games)	12
3.1.3	GMM (Gaussian Mixture Model).....	13
3.1.4	Transformaattori (Transformer).....	14
3.1.5	LSTM RNN (Long Short-Term Memory Recurrent Neural Network).....	15
3.1.6	PreGLAM-MMM (Predictive Gameplay-based Layered Affect Model - Multi-Track Music Machine)	16
3.1.7	PAMG (Progressive Adaptive Music Generator)	17
3.1.8	Yhteenveto.....	18
3.2	Generatiivisen pelimusiikin jatkokehitys	21
3.3	Tekoälymusiikin ongelmallisuus.....	22
4	POHDINTA JA YHTEENVETO	23
	LÄHTEET	26

1 JOHDANTO

Digitaalinen pelaaminen, kuten video- ja mobiilipelit, ovat osa lähes jokaisen ihmisen arkea. Esimerkiksi *Pelaajabarometri 2022* -tutkimuksen mukaan suomalaisista 98 % pelaa jotakin peliä ainakin joskus (Kinnunen, Tuomela & Mäyrä, 2022, s. 1). Pelaamisen suosion kasvun myötä peliteollisuudesta on tullut myös taloudellisesti merkittävä. Esimerkiksi raportin mukaan vuonna 2022 globaalin videopelimarkkinoiden arvo oli yli 217 miljardia dollaria ja se tulee kasvamaan vuosittain 13,4 % seuraavan seitsemän vuoden aikana (Grand View Research, ei pvm.). Teknologian edistymisen myötä videopelit ovat kehittyneet viimeisen viidenkymmenen vuoden aikana 8-bittisestä pikseligrafiikasta ja piippauksista jopa suuriksi, tunteita herättäviksi elämyksiksi. Pelin äänimaailma musiikkeineen ja äänitehosteineen vaikuttaa osaltaan tunteikkaan ja merkityksellisen pelikokemuksen muodostumiseen.

Interaktiivisuus on olennainen osa videopelejä ja pelaamista. Pelaaja voi vuorovaikuttaa toiminnallaan pelinsisäisiin tapahtumiin, mikä taas vaikuttaa pelissä kuultaviin ääniin sekä musiikkiin. Jotta pelin musiikki ei kuulostaisi yksitoikkoiselta, käytetään sen muokkaukseen erilaisia tekniikoita, kuten generatiivista musiikkia. Se on musiikkia, joka muodostuu itsestään reaaliajassa järjestelmällisellä tavalla (Amaral, Baffa, Briot, Feijo, & Furtado, 2022; Plut & Pasquier, 2020). Generatiivisen pelimusiikin tutkimus on lisääntynyt viime vuosina (kts. esim. Clark, 2021; Lopez Duarte, 2023; Plut, 2022). Generatiivista musiikkia kuullaan peleissä yhä useammin, mutta sen käyttö on silti kahtiajakautunutta: yksinkertaisia ja tehokkaita järjestelmiä käytetään peliteollisuuden julkaisuissa, kun taas edistyneemmät, mutta vähemmän testatut järjestelmät ovat lähinnä tutkimukseen kehitettyjä (Plut & Pasquier, 2020).

Tekoälyn kehittyessä siitä on tullut myös olennainen osa generatiivisen musiikin luontia (Li, 2023, s. 1). Etenkin pelimusiikin sävellyksessä generatiivisen musiikin ajatteluaan toimivan säveltäjän työkaluna mukautuvan ja laadukkaan musiikin sävellyksessä (Plut, 2022, s. 8). Jo yli 30 vuotta sitten Helsingin Sanomien artikkelissa säveltäjä ja musiikkiteknologian lehtori Otto Romanowski pohti tekoälyn avulla säveltämisen ja ihmisen suhdetta: "Siirrämme ihmisyyttä koneisiin: taidon säveltää, taidon puhua,

taidon kirjoittaa ja lopulta myös taidon ajatella." (Siren, 1993). Innostus tekoälyyn ja sen hyödyntämiseen niin tutkimuksessa kuin arkielämän askareissakin on kasvanut etenkin sen yleistyttyä esimerkiksi generatiivisen kielimallin ChatGPT:n muodossa. Tämän esimerkin perusteella voidaan pohtia, tuleeko generatiivisesta musiikista tulevaisuudessa normalisoitunut musiikintekotapa. Tekoälyn käytön yleistyttyä keskustelua on kuitenkin syntynyt esimerkiksi tekijänoikeuksista ja tekoälyn käytön eettisyydestä eri taiteenalojen parissa (kts. Clancy, 2022). Ongelmallista tekoälyn käytössä on se, että sitä koulutetaan ihmisen tuottamalla taiteella, joka usein on tekijänoikeudella suojattua (Vesala, 2023). Akateemisessa maailmassa tällaisen materiaalin käyttöä perustellaan usein tutkimuskäytöllä (kts. Baio, 2022).

Tämä kandidaatintutkielma tarkastelee ajankohtaisen kirjallisuuskatsauksen avulla, kuinka generatiivista pelimusiikkia on tutkittu ja kehitetty videopeleihin 2020-luvulla. Tutkielman tavoitteena on selvittää, minkälaisia eroavaisuuksia ja yhtenevyyksiä generatiivisen pelimusiikin malleilla ja järjestelmillä. Tarkoituksena on yhtenäistää tietoa generatiivisen pelimusiikin malleista ja selittää niiden monimutkaista toimintaa yksinkertaistetusti. Kandidaatintutkielman tutkimuskysymyksinä ovat:

1. Millaisia generatiivista pelimusiikkia tuottavia malleja tai järjestelmiä on kehitetty ja tutkittu akateemisessa maailmassa 2020-luvulla?
2. Mitä yhtenevyyksiä ja eroavaisuuksia näillä malleilla ja järjestelmillä on?

Kandidaatintutkielman menetelmänä käytetään narratiivista kirjallisuuskatsausta. Sen avulla aiheesta saadaan ajankohtaista, yhtenäistettyä tutkimustietoa, joka on samalla helposti ymmärrettävää (Salminen, 2023, s. 8–9). Kandidaatintutkielman aineisto perustuu aiempaan englanninkieliseen kirjallisuuteen. Aineistoa on haettu JYKDOKin kautta tietokantoja selaten sekä Google Scholar -sivuston kautta. Käytetyt hakusanoja ovat *generative music*, *video game music*, *adaptive game music*, *procedural music* sekä *algorithmic music*, jotka suomeksi ovat generatiivinen musiikki, pelimusiikki, adaptiivinen pelimusiikki, prosessuaalinen musiikki ja algoritminen musiikki. Myös eri tietokantoja, kuten IEEE Xplore sekä Ebook Central, on hyödynnetty aineistonhaussa. Valittu aineisto koostuu lähinnä informaatiotieteen tutkimuksesta, johon yhdistyy myös musiikkitieteellinen näkökulma etenkin musiikkiteknologian ja -psykologian kautta. Haku tutkimusartikkeleista, jotka käsittelevät generatiivisen musiikin käyttöä peleissä, on rajattu alkavaksi vuodesta 2020, sillä tutkielman tarkoituksena on toimia ajankohtaiskatsauksena aiheeseen. Muuta tutkimusaineistoa on 1990-luvulta alkaen nykyhetkeen. Generatiivista pelimusiikkia on tutkittu vähemmän verrattuna yleisesti generatiivista musiikkia luoviin malleihin. Valittu aineisto on englanninkielistä, sillä generatiivisen pelimusiikin tutkimusta ei ole tehty suomeksi.

2 TEOREETTINEN TAUSTA

2.1 Lyhyesti pelimusiikin teknologisesta kehittämisestä

Kun puhutaan pelimusiikista, monelle saattaa tulla mieleen esimerkiksi Super Mario-sarjan iloinen, ”kasibittinen” musiikki. Nykyään pelimusiikki voi olla tyyllilajiltaan mitä tahansa musiikkia, jota on mahdotonta määritellä vain yhdeksi genreksi. Tässä tutkielmassa pelimusiikilla tarkoitetaan musiikkia, joka on sävelletty videopeliä varten. Pelissä äänimaailma koostuu musiikista, tunnelmasta (engl. *ambience*), dialogista, ääniefekteistä ja käyttöliittymän eli esimerkiksi pelivalikon äänistä, joista puhutaan yleensä yhteisesti audiona (Collins, 2009, s. 1). Alun perin ensimmäisissä 1950-luvulla tehdyissä videopeleissä ei ollut ääntä tai musiikkia, ja vasta vuonna 1972 julkaistu *Pong* toimi aikansa suunnannäyttäjänä peliäänten käytössä (Collins, 2008, s. 8). Pelissä itsessään soiva musiikki sai alkunsa siis vasta vuosia pelinkehityksen alkamisen jälkeen. Tähän aikaan tietokoneiden ja konsolien muisti oli rajallista, minkä vuoksi musiikin piti olla yksinkertaista, ja 1980-luvun alussa peleissä kuultiinkin ensimmäisen kerran toistuvista luupeista (engl. *loop*) koottua musiikkia (Collins, 2008, s. 12).

8-bittinen äänimaailma luotiin ensimmäisiin arcade- ja pinball-pelikoneisiin esimerkiksi ohjelmoitavan äänigeneraattorin (engl. *Programmable Sound Generator, PSG*) avulla analogisesti (Collins, 2005, s. 7). Äänikorttien käyttö tietokoneissa lisäsi MIDI-musiikin (engl. *Musical Instrument Digital Interface*) käytön mahdollisuuden peleihin, kunnes 1990-luvulla siirryttiin CD-ROM -teknologiaan (Collins, 2005, s. 12–13). Peli-audio on kehittynyt tämän jälkeen nopeaa vauhtia. Sampleja eli tallennettuja ääniä on käytetty peleissä etenkin 1990- ja 2000-luvuilla, ja ääntä on voitu varioida erilaisilla äänisuodattimilla pelin tapahtumiin sopiviksi (Farnell, 2010, s. 318). Nykyään peliteollisuudessa käytetään äänelle erilaisia väliohjelmistoja eli äänimoottoreita (Plut & Pasquier, 2020, s. 16).

2.2 Pelimusiikki ja interaktiivisuus

Pelimusiikkia ja sen merkitystä pelaajan pelikokemukseen on tutkittu paljon (kts. esim. Collins, 2008; Fritsch & Summers, 2021), minkä takia laadukkaaseen musiikkiin halutaan panostaa ja investoida peliteollisuudessa yhä enemmän. Elokuva-alalla affektivistia eli tunnepitoista musiikkia on sävelletty jo kauan, ja lineaarisesti alusta loppuun eteneviä elokuvia verrataan usein peleihin. Pelin interaktiivisuuteen sisältyvä epälineaarisuus (engl. *nonlinearity*) erottaa pelit kuitenkin elokuvista, sillä jokainen pelikerta voi olla erilainen (Collins, 2008, s. 4).

Musiikki ei pelkästään luo tunnelmaa tai kuvasta pelimaailman tapahtumia pelaajalle, vaan sillä on videopeleissä muitakin funktioita. Energisellä tai rauhallisella musiikilla voidaan vaikuttaa pelin tahtiin, tai musiikki voi toimia ”yleisönä” pelaajan toiminnalle (Phillips, 2014, s. 106–109). Musiikilla voidaan myös parantaa pelaajan immersiota peliin (Huiberts, 2010, s. 101). Pine II:n & Gilmoren (1999, s. 31) määritelmän mukaan immersio tarkoittaa sitä, kun ihminen pelatessaan uppoutuu virtuaalisen maailman mahdollistamaan kokemukseen. Tärkeä osa immersiota on musiikin ja toiminnan luoma välitön vuorovaikutus (van Elferen, 2016, s. 39). Kun pelaajalle halutaan luoda tunnetasolla merkittävä pelikokemus, musiikin vaihtelu ja mukautuvuus ovat olennaisia tekijöitä. Immersioon liittyviä pelimusiikintutkimuksessa esiintyviä termejä ovat dynaaminen, adaptiivinen tai interaktiivinen musiikki tai audio, joita käsitellään seuraavaksi.

Collins (2009, s. 5) määrittelee dynaamisen audion niin, että se reagoi pelimaailman tapahtumiin ja pelaajan toimintaan jakautuen adaptiiviseen ja interaktiiviseen audioon. Interaktiivinen audio tarkoittaa pelin ääniä, jotka toistetaan pelissä välittömästi pelaajan toiminnan takia, joten esimerkiksi kävellessä hahmon kävelyaskeleet ovat osa interaktiivista audiota. Adaptiivinen audio taas tarkoittaa toistettavia ääniä, jotka eivät suoraan johdu pelaajan toimista, vaan ne toistetaan tietystä kohtaa peliä: esimerkiksi musiikki voi muuttua, kun pelissä päivä vaihtuu yöksi. (Collins, 2009, s. 5–6.) Phillips (2014, s. 187) taas puhuu dynaamisesta, adaptiivisesta ja interaktiivisesta musiikista lähes synonyymeina, ja Sweetkin (2015, s. 36) esittää interaktiivisen ja adaptiivisen musiikin tarkoittavan lähes samaa asiaa. Eron tekee se, että pelaajalla on interaktiivisessa musiikissa suurempi kontrolli kuultavaan musiikkiin, mutta tämän toiminnalla on epäsuorempi vaikutus adaptiivisen musiikin soimiseen (Sweet, 2015, s. 36). Tuoreemmassa tutkimusaineistossa adaptiivinen musiikki tai audio onkin terminä korvannut myös dynaamisen ja interaktiivisen musiikin tai audion termien käytön. Nykyään adaptiivisella musiikilla tarkoitetaan lähinnä lineaarista, jo ennalta sävellettyä ja tallennettua musiikkia, joka mukautuu pelin tapahtumiin (Plut & Pasquier, 2020, s. 2).

2.3 Generatiivinen musiikki

Kun puhutaan itsestään säveltyvästä musiikista, lähes poikkeuksetta sen esimuodoksi mainitaan Mozartin musikaalinen noppapeli (saks. *Ein musikalisches Würfelspiel*), jossa sävellys muodostuu valmiiksi sävelletyistä sektioista noppaa heittämällä (Collins, 2008, s. 156, 2009, s. 9; Phillips, 2014, s. 186; Plut & Pasquier, 2020, s. 2). Ensimmäisen kerran generatiivisesta musiikista on puhuttu vuonna 1996 itsestään muodostuvana, ennalta arvaamattomana musiikkina, joka järjestäytyy sääntöjen mukaisesti (Eno, 1996). Pelimusiikintutkimuksessa generatiivisen musiikin synonyymeinä käytetään joskus myös algoritmista musiikkia (engl. *algorithmic music*), prosessuaalista musiikkia (engl. *procedural music*) tai musiikillista metaluomusta (engl. *musical metacreation*) (Plut & Pasquier, 2020, s. 2). Tutkielman aineistossa musiikillista metaluomusta ei esiinny, joten sitä ei määritellä.

Algoritminen musiikki tarkoittaa musiikkia, joka perustuu tietokoneella käytäviin sääntöihin ja prosesseihin. Se pohjaa algoritmeihin, jotka ovat tietokoneella tarkasti määritellyjä toimintoja tai sääntöjä. Algoritminen musiikki on ikään kuin tapa luoda musiikkia ihmisen ja algoritmin vuorovaikutuksessa, sillä algoritmit voivat toimia nuottien tapaan ihmisen säveltämänä tai itsenäisesti musiikkia luoden. (McLean & Dean, 2018, s. 4.) Woollerin, Brownin, Mirandan, Berryn sekä Diederichin (2005) viitekehysten mukaan algoritmisella musiikilla on erilaisia funktioita. Jokin algoritmi voi esimerkiksi transponoida yksittäisiä nuotteja tai kääntää fraaseja. Algoritmi on generatiivinen, jos se tuottaa musiikkia, joka ei ole alkuperäisen toisintoa, vaan se sisältää enemmän vaihtelua ja uusia musiikillisia ideoita. (Wooller ym., 2005, s. 8–9.)

Collins (2009, s. 8) käyttää Woollerin ja kumppaneiden (2005) viitekehystä prosessuaalisen musiikin määrittelyyn: se on sävellys, joka kehittyy reaaliajassa tiettyjen sääntöjen perusteella, kun taas Clarkin (2021, s. 1) mukaan prosessuaalinen ääni luodaan reaaliajassa säveltäjän määrittelemien pelitapahtumien muuttujien mukaisesti, eikä se vaadi suurta äänidataa. Farnellin (2010, s. 1) määritelmän mukaan prosessuaalinen ääni on ”elävä ääniefekti”, joka toimii tietokonekoodilla ja mukautuu ennalta-arvaamattomiin tapahtumiin reaaliajassa.

Generatiivisen musiikin tuottamistavat ovat teknologian kehittyessä muuttuneet nopeasti. Edeltävien määritelmien ja 2020-kirjallisuuden mukaan generatiivinen musiikki luodaan reaaliaikaisesti jollakin järjestelmällisellä tavalla (Plut & Pasquier, 2020, s. 2). Amaral ja muut (2022, s. 2) määrittelevät lisäksi, että se ei ole ennalta ole-massa tai dynaamisesti sovitettua, kuten adaptiivinen musiikki. Teknisiltä vaatimuksiltaan generatiivisen musiikin kontrollointi on haastavampaa kuin adaptiivisen musiikin hallinta, sillä se vaatii tietokoneelta suurempaa laskennallista kapasiteettia (Amaral ym., 2022, s. 3).

Kuten aiemmin jo mainittiin, nykytutkimuksessa adaptiivinen musiikki yhdistetään enemmän ennalta sävellettyyn musiikkiin. Peliteollisuudessa käytettävät äänimootorit mahdollistavat interaktiivisen ja adaptiivisen musiikin toistamisen pelin tapahtumien mukaan (Phillips, 2014, s. 227; Plut & Pasquier, 2020, s. 16). Tallennettu musiikki kuitenkin rajoittaa adaptiivisen musiikin sujuvuutta, sillä isokokoiset äänitiedostot eivät ole toivottavia pelin toimivuuden kannalta ja tiedoston yhden kohdan muuttaminen vaikuttaa lopulta koko tiedostoon (Clark, 2021, s. 1). Modernista pelimusiikista puhuttassa adaptiivisen ja generatiivisen pelimusiikin välinen ero ei ole selkeä, koska generatiiviksi luokiteltavat musiikkijärjestelmät sisältävät usein adaptiivisia osatekijöitä (Amaral ym., 2022, s. 3). Musiikki voi olla siis samanaikaisesti adaptiivista ja generatiivista, mutta enemmän sen luokitteluun vaikuttaa luontitapa ja täten luodun musiikin määrä. Hyvä esimerkki tästä on se, että vaikka tämä tutkielma on otsikoitu käsittelemään generatiivista pelimusiikkia, on aineistoon valittu myös adaptiivisen pelimusiikin järjestelmiä, joista ei välttämättä puhuta generatiivisena musiikkina ollenkaan (kts. esim. Hutchings & McCormack, 2020). Tästä voidaan huomata, että samoista asioista puhutaan monella eri termillä, mikä hankaloittaa myös yhtenäisen tiedon löytämistä.

2.4 Aikaisempi generatiivisen pelimusiikin tutkimus

Generatiivisen musiikin luomisessa ongelmana on ollut sen laatu: onko mielekästä kuunnella musiikkia, joka ei herätä kuuntelijassa mitään tunteita? Yli viisitoista vuotta sitten Collins (2009, s. 12) arvioi, ettei generatiivista musiikkia ole käytetty peleissä, koska pelimusiikin pitää olla merkityksellistä ja herättää tunteita. Tähän aikaan generatiivisen musiikin tuottaminen on ollut vielä alkeellisempaa. Collins (2009, s. 12) toteaa myös, että MIDI-musiikin ollessa toimiva ratkaisu monipuolisuutensa ansiosta, sitä kuitenkin pidetään huonolaatuisena ja ei niin luonnollisen kuuloisena musiikkina verrattuna ihmisen soittamaan, tallennettuun musiikkiin. Viidessätoista vuodessa teknologia on kehittynyt niin, että generatiivista musiikkia tutkitaan yhä enemmän, ja MIDI on kuitenkin todettu toimivaksi ratkaisuksi. Generatiivisen musiikin tutkimuksessa ja kehityksessä malleissa on usein uudenlaisia generatiivisia algoritmeja, jotka järjestelmän avulla luovat reaaliajassa symbolista musiikkia MIDI-muodossa (Plut, Pasquier, Ens & Tchemeube, 2022, s. 1). Symbolinen musiikki tarkoittaa esimerkiksi MIDI-tallennusmuodossa olevaa nuotinnosta, joka sisältää tietoa sävelkorkeudesta eri instrumenttiraidoilla tai nuottien soivasta ajasta. Siinä ei ole ääntä, kuten esimerkiksi digitaalisesti tallennetussa musiikkitiedostossa. (Lidy & Rauber, 2009, s. 456.)

Plut ja Pasquier (2020) esittelevät 34 eri generatiivisen musiikin järjestelmää tai mallia, joita on kehitetty niin peliteollisuudessa kuin tutkimuksessakin (kts. esim.

Engels, Tong & Chan, 2015; López Ibáñez, Álvarez & Peinado, 2018; Precht, 2015). Aikasempaan tutkimukseen perehdytään tässä tutkielmassa kuitenkin vain lyhyesti kahden prosessuaalista pelimusiikkia tuottavan mallin avulla. Ensimmäinen on Lopesin, Liapiksen ja Yannakakiksen (2015) kehittämä malli, joka muodostaa musiikkia kauhupeliin. Kyseinen järjestelmä valitsee ihmisen äänittämistä sampleista tiettyjen arvojen mukaiset äänet, jotka yhdistetään instrumentin, nuotin, oktaavin ja äänenvoimakkuuden mukaan. Järjestelmä toimii reaaliaikaisesti miksaavan algoritmin avulla. (Lopes ym., 2015, s. 4.)

Useimmissa viime vuosina kehitetyissä malleissa ja järjestelmissä käytetään jonkinlaista emootiomallinnusta, joka muokkaa generatiivisen musiikin affektiivisuutta (Plut ym., 2022, s. 1). Emootiomallinnus tarkoittaa sitä, että pelitilanteessa koettuja tunteita, kuten iloa tai surua, välitetään pelaajalle musiikin kautta. Lopesin ja kollegoiden (2015) malliin verrattuna esimerkiksi Scirean (2017) malli on paljon kehittyneempi, sillä siinä tekoäly reagoi reaaliajassa ärsykkeisiin sekä tuottaa samanaikaisesti affektiivista musiikkia. Scirean (2017, s. 3) mukaan pelikokemuksessa erilaiset tunnelmat ja tunteet (engl. *feelings*) ovat merkittäviä, koska musiikin ilmaistessa mielialaa (engl. *mood*), myös emootioiden (engl. *emotion*) pitäisi kuulua musiikista pelitapahtumien etenemisen myötä. Esimerkiksi Eerola & Saarikallio (2010) määrittelevät edellä mainittuja termejä seuraavasti:

Tunteet ovat emootioiden yksityistä kokemista, kun taas emootioilla tarkoitetaan sitä väestien joukkoa, joista useimmat voidaan julkisesti havaita ja mitata. Emootiot ja tunteet eroavat puolestaan mielialoista siten, että ne ovat lyhytkestoisia (sekunneista minuutteihin) ja intensiivisiä prosesseja, joilla yleensä on selkeä syy ja kohde. Mielialat puolestaan ovat pitkäkestoisempia taustalla vaikuttavia tiloja, joilla ei ole selkeää syytä. (Eerola & Saarikallio, 2010, s. 260.)

Scirean (2017) mallissa emootiomallinnus pohjautuu Russellin (1980) luomaan kaksiulotteiseen valenssin ja vireyden affektimalliin. Affektilla (engl. *affect*) viitataan yleisesti tunteiden, emootioiden ja mielialojen ilmiöihin (Eerola & Saarikallio, 2010, s. 260). Mallissa vaaka-akselin ulottuvuudella on valenssi (engl. *valence*) eli miellyttävyyden tai epämiellyttävyyden tunne, kun taas poikkiakselilla määritetään vireystila (engl. *arousal*) kiihkeästä rentoutuneeseen (Russell, 1980, s. 1163). Musiikintutkimuksessa käytetään usein edellä mainitun mallin lisäksi myös kategorista mallia, jossa tunteet luokitellaan tavanomaisimpiin tunteisiin, kuten ilo, surullisuus tai pelko (Eerola & Saarikallio, 2010, s. 262). Scirean (2017) mallissa jokainen musiikillinen ominaisuus (tempo, äänenvoimakkuus, sointi, rytmin vahvuus, säännöllisyys ja dissonanssit) yhdistyy joko valenssin tai vireystilan ulottuvuuteen (Scirea, 2017, s. 3).

Scirean (2017) mallissa käytetään tekoälyä niin uusien algoritmien kuin neuroverkkojenkin avulla. Keinotekoiset neuroverkot (engl. *Artificial Neural Network*) mallintavat matemaattisesti biologisia neuroverkkoja. Ne toimivat samaan tapaan kuin oikeat biologiset neuroverkot simuloiden niiden rakennetta. Neuroverkot koostuvat

erilaisista neuroneista, jotka ovat yksinkertaisimmillaan matemaattisia funktioita. (Graupe, 2013, s. 1; Krenker ym., 2011, s. 3.) Neuroverkot toimivat niin, että niitä koulutetaan tuottamaan haluttua dataa, kuten musiikkia, antamalla niille ensin esimerkkejä. Tätä esimerkkipohjaista koulutustapaa kutsutaan koneoppimiseksi. (Rebala, Ravi & Churiwala, 2019, s. 2.)

Tekoälyn yhteydessä puhutaan joskus myös agenteista tai moniagenttisuudesta. Agentilla tarkoitetaan toimijaa, joka havaitsee ympäristöään ja toimii sen mukaan. Esimerkiksi ihmisagentti havaitsee aisteillansa, kun taas tekoälyagentti saa havaintonsa datasyötteenä. (Russell & Norvig, 1995, s. 31.) Agentit voivat myös toimia yhdessä moniagenttisesti, jolloin samassa ympäristössä olevat agentit toimivat autonomisesti esimerkiksi joko yhteistyössä tai kilpaillen yhteisen tavoitteen eteen (Balaji & Srinivasan, 2010, s. 3).

Vaikka Scirea (2017, s. 97) toteaa mallissa olevan ongelmia, se kuitenkin pystyy tuottamaan luotettavasti ja havaittavasti affektiivista musiikkia. Esimerkiksi Lopesin ja kumppaneiden (2015, s. 5) mallissa Russellin (1980) affektimallia tai neuroverkkoja ei ole käytetty, mutta jatkokehittämisen kannalta he pohtivat niiden lisäämistä mallia parantaakseen. Aiemman tutkimuksen perusteella emootiomallin ja neuroverkkojen käytöstä on tullut yksi generatiivisen pelimusiikin kehityssuunta, jota käytetään myös 2020-luvun tutkimuksessa.

2.5 Generatiivisen musiikin käyttö peliteollisuudessa

Generatiivinen musiikki ei toistaiseksi ole saavuttanut suurta suosiota peliteollisuudessa, sillä peleissä käytetään valmiiksi äänitettyjä, yhteen sulautuvia äänilähteitä tai ennalta tehtyjä musiikkisampleja, jotka järjestellään satunnaisesti (Plut, 2022, s. 6). Farnell (2010, s. 318) kuitenkin kertoo peliaudion teknologian olleen alun perin prosessuaalista, jossa hyödynnettiin konsoleiden ja tietokoneiden syntetisaattorisirua. Teknologian kehittyessä äänisamplen käyttö vei kuitenkin voiton realistisuudellaan (Farnell, 2010, s. 318). Aikana, jolloin kaikki tämä oli vielä uutta, mahdollisimman realistisen pelikokemuksen ajateltiin myös lisäävään pelimyyntiä (Collins, 2013, s. 36).

Ensimmäinen isompi pelijulkaisu, jossa käytettiin generatiivista musiikkia uudella tavalla, on *Spore* vuodelta 2008 (Collins, 2009, s. 11). Merkittävänä esimerkkinä peli nousee esille tutkimuskirjallisuudessa edelleen useasti (Clark, 2021; Collins, 2009; Harkiolakis, 2023; Lopez Duarte, 2023; Plut & Pasquier, 2020). Sporessa pelimusiikki luodaan moniagenttisesti muodostamalla fraaseja pelitilanteen mukaan (Lopez Duarte, 2023). Sporen musiikki koostuu monista pienistä musiikkisampleista sääntöjen mukaisesti reaaliajassa, ja pelaaja voi myös tehdä tai muokata omaa musiikkiaan eri rytmisekvenssien ja nuottien perusteella (Collins, 2009, 11). Valinnanvapautta on

kuitenkin rajoitettu, jotta musiikki kuulostaisi edelleen miellyttävältä (Clark, 2021, s. 67; Collins, 2009, s. 11). Plut ja Pasquier (2020, s. 9) määrittelevät Sporen musiikin lineaarisesti generatiiviseksi, sillä kentän latautuessa luodaan samanaikaisesti musiikki, joka soi luuppina kentän loppuun asti.

Sporen jälkeen generatiivista musiikkia ei kuitenkaan ole erityisemmin kuultu isommissa pelijulkaisuissa muutamia pelejä lukuun ottamatta. Clarkin (2021, s. 70–71) mukaan Sporen prosessuaalinen musiikki ei yllä suuren pelituotannon tasolle, eikä se onnistunut vakuuttamaan peliteollisuutta tällaisen pelimusiikin käytöstä. Harkiolakis (2023, s. 25) taas toteaa, että Sporen myötä monimutkaista generatiivisen musiikin järjestelmää on kuitenkin käytetty suuressa kaupallisessa pelijulkaisussa, mikä kertoo sen käytön olevan toteutettavissa. Generatiivista pelimusiikkia on tutkittu kuitenkin yhä enemmän viimeisen 15 vuoden ajan, mikä osoittaa, että etenkin teknologian kehittymisen myötä siitä haluttaisiin uusi, käyttökelpoinen malli peliteollisuuden käytettäväksi.

3 GENERATIIVISEN PELIMUSIIKIN MALLIEN TUTKIMUS 2020-LUVULLA

3.1 Mallit ja järjestelmät

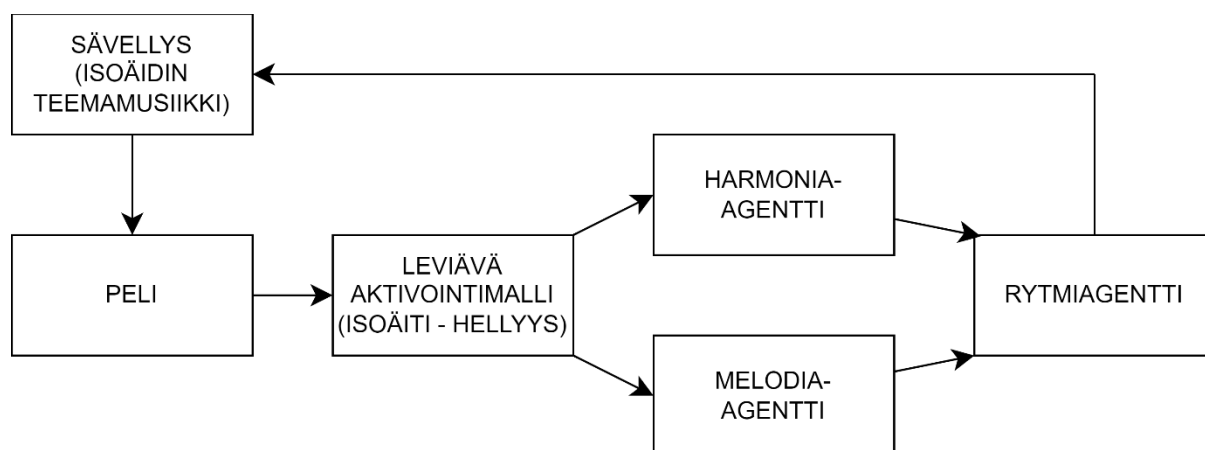
Tässä osiossa käydään läpi eri generatiivisen musiikin malleja ja järjestelmiä, joita on julkaistu 2020-luvulla. Tutkimukset esitellään julkaisuvuoden mukaan.

3.1.1 AMS (Adaptive Music System)

Hutchings ja McCormack (2020) haluavat adaptiivisella musiikkijärjestelmällään (engl. *Adaptive Music System, AMS*) ratkaista generatiivisen pelimusiikin tuottamisen ongelmia, kuten algoritmien tai pelimaailman mallinnuksen rajoittuneisuutta. Kun videopeleihin tuotetaan affektiivista musiikkia, olennainen osa järjestelmää on malli, joka viestii pelin tapahtumista, sisällöstä ja tunnelmasta siirtäen näiden antaman tiedon musiikin affektiivisuuteen. AMS koostuu kolmesta pääosasta: (1) kognitiotieteissä käytetystä leviävästä aktivointimallista (engl. *spreading activation model*), joka käsittelee pelin sisältöä, (2) kuuden kategorian emootiomallista, joka kuvailee pelin sisältöä sekä (3) adaptiivisesta järjestelmästä, joka säveltää ja sovittaa ammattilaisen säveltämiä musiikkiosia käyttäen edellä mainittuja malleja. (Hutchings & McCormack, 2020, s. 270.)

Musiikki muodostuu järjestelmässä seuraavalla tavalla (kts. Kuvio 1). Leviävä aktivointimalli mallintaa peliä käyttäen kuutta affektikategoriaa, jotka ovat suru, ilo, uhka, viha, hellyys ja innostuneisuus (Hutchings & McCormack, 2020, s. 273). Kyseinen emootiomalli pohjautuu Juslinin (2013, s. 11) emootiotutkimukseen. Kategoriat päivittyvät pelin sisällön mukaan. Esimerkiksi jos pelissä esiintyy isoäitihahmo, tällöin affektikategoriasta valitaan hellyys, jonka mukaan musiikki muokkautuu.

(Hutchings & McCormack, 2020, s. 274.) Musiikin luominen tapahtuu AMS:ssä moniagenttisen dynaamisen sävellysjärjestelmän avulla, jossa ohjelmistoagentteja käytetään musiikillisten sovitusten luomiseen. Järjestelmä luo musiikkia kaksi tahtia kerrallaan ja se koostuu kolmesta agentista, jotka ovat harmonia, melodia ja rytmi. (Hutchings & McCormack, 2020, s. 274.)



Kuvio 1 AMS:n yksinkertaistettu toimintamalli (mukaiillen Hutchings & McCormack, 2020, s. 276).

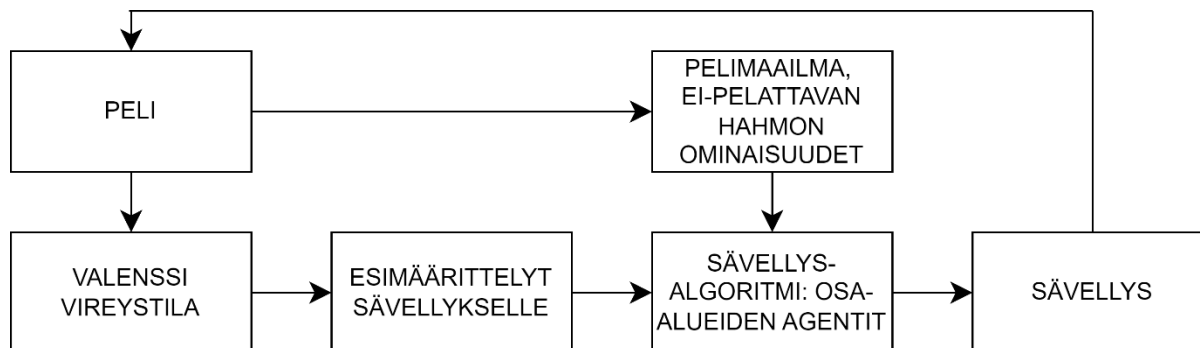
Hutchingsin ja McCormackin (2020, s. 274) AMS:n harmonia-agentin mallina on takaisinkytketty neuroverkko (engl. *recurrent neural network*, RNN). Se on neuroverkkomalli, joka säilyttää aikaisemmin saatua tietoa samalla, kun se vastaanottaa uutta tietoa (Plut & Pasquier, 2020, s. 14). Myös rytmin lisäämisessä käytetään RNN:ää ja se toimii yhdessä melodia-agentin kanssa (Hutchings & McCormack, 2020, s. 276). Malli on koulutettu käyttäen lähes 6800 kappaletta eri genreistä, kuten folk-, pop-, rock- tai jazzmusiikkia. Harmonia-agentti ei lisää sävellykseen nuotteja, vaan se tuottaa pelin musiikkiin harmonista sisältöä. (Hutchings & McCormack, 2020, s. 274.)

Melodia-agentit taas saavat tietoa pelintilanteesta leviävän aktivointimallin kautta. Tämän perusteella pelissä olevan objektin, kuten edellä mainitussa esimerkissä isoäidin, melodista teemaa soitetaan tarvittaessa. Pelissä kuullaan myös ihmisen säveltämää musiikkia, jota ohjataan melodia-agenteilla. Jokaiselle instrumentille on oma agenttinsa ja niitä on muokattu toimimaan paremmin reaaliajassa. Ihmisen ennalta säveltämät melodiaosat ovat pituudeltaan yhdestä neljään tahtia ja ne on sävelletty pelin eri konsepteille. Sävellystekniikoina on käytetty esimerkiksi sointujen vähennystä, lisäämistä tai kääntämistä. Muodostettava musiikki on symbolista. (Hutchings & McCormack, 2020, s. 273–275.)

Adaptiivista musiikkijärjestelmää on tutkittu kahden eri pelin avulla. Osallistujat ovat pelanneet pelejä ja arvioineet pelimusiikin toimivuutta kyselytutkimuksessa. Tutkimustulosten mukaan AMS:n avulla pelaaja voi kokea immersoituvansa enemmän pelimaailmaan, ja että pelimusiikki sopii pelitilanteeseen alkuperäistä pelimusiikkia paremmin. (Hutchings & McCormack, 2020, s. 277.)

3.1.2 PMGG (Procedural Music Generation in Games)

Washburnin & Khosmoodin (2020) kehittämän prosessuaalisen musiikkijärjestelmän (engl. *Procedural Music Generation in Games, PMGG*) tarkoituksena on luoda teematista musiikkia ei-pelattaville hahmoille (engl. *non-playable character, NPC*). PMGG pohjautuu Scirean (2017) esittelemään malliin. AMS:n kaltaisesti PMGG on moniagenttinen järjestelmä, jonka osasten toiminnan mukaan musiikki koostuu (kts. Kuvio 2). Järjestelmässä musiikin affektiivisuuden määrittely pohjautuu Russellin (1980) affektiteoriaan. Aluksi PMGG-järjestelmä saa syötteenä esimääritettyjä tietoja. Tämän jälkeen ei-pelattavan hahmon toiminnan mukaisesti sävellysalgoritmi muodostaa pelissä reaaliajassa soivaa MIDI-musiikkia. Algoritmi vaikuttaa sävellyksen kuuteen eri osa-alueeseen, jotka esitellään seuraavaksi. (Washburn & Khosmood, 2020, s. 1-2.)



Kuvio 2 PMGG:n yksinkertaistettu toimintamalli (mukaillen Washburn & Khosmood, 2020, s. 3).

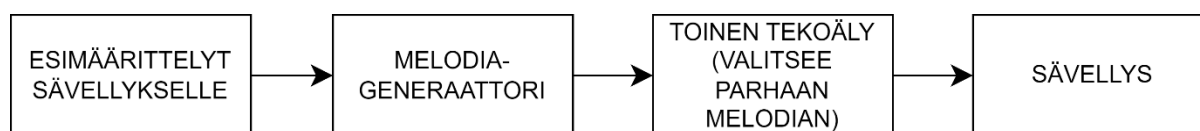
Ensimmäisessä osa-alueessa harmoninen rakenne määritellään valenssin mukaan "valoisaksi tai tummaksi" (Washburn & Khosmood, 2020, s. 2). Toisin sanoen musiikkia voi luonnehtia iloiseksi tai pelottavaksi. Seuraavassa osa-alueessa soinnut muodostetaan ABAB-sointukulun mukaan (Washburn & Khosmood, 2020, s. 2). Melodiaa luovassa osa-alueessa määritetään tahti ja rytmi, jonka jälkeen sävelkorkeuskäyrän generaattori muodostaa sävellyksen tiettyjen sääntöjen mukaan. Esimerkiksi jos pelaaja hyppää pitkälle, sävellyksestä poistetaan kolmas toistuva nuotti, jonka tilalle tulee intervallisuus. Lisäksi pitkälle hyppääminen lisää dissonanssia ja vain

tietyin sävelasteikon käyttämistä. (Washburn & Khosmood, 2020, s. 2.) Oletusarvoisena rytminä musiikissa on 4/4-tahti, mutta kevyempää ja eloisampaa tunnelmaa tavoitellessa tahtina on kolmijakoinen valssi. Lopuksi ei-pelattavan hahmon ominaisuudet vaikuttavat esimerkiksi tempoon ja instrumentaatioon. (Washburn & Khosmood, 2020, s. 2–3.)

PMGG-järjestelmän toimimista pelissä on tutkittu, ja kyselytutkimuksessa osallistujat arvioivat musiikin laatua ja sitä, mitä musiikilla halutaan kertoa. Tutkimustulosten mukaan PMGG:n tuottama musiikki on useimmiten miellyttävää ja musiikillisesti harmonista. Pelaajat pystyvät musiikin avulla ymmärtämään pelihahmonsia ja ei-pelattavien hahmojen välisiä suhteita, vaikka musiikkiin liittyviä yksityiskohtia ei täsmällisesti pystytä nimeämään. Tällä tarkoitetaan sitä, että esimerkiksi melodian tarkoituksena on havainnollistaa vihollisen vaikeustasoa, mutta kyselytutkimuksessa vastaajat ymmärtävät sen enimmäkseen kuvaavan ympäristöä, riskiä tai ei-pelattavaa hahmoa. (Washburn & Khosmood, 2020, s. 3–4.)

3.1.3 GMM (Gaussian Mixture Model)

Xiangin ja Guon (2021) tutkivat tekoölyavusteisen pelimusiikin säveltämistä tavoitteenaan kehittää malli, joka on kustannustehokas säveltämistyökalu. Heidän mukaansa tämänhetkisten generatiivisen musiikin mallien ongelmana on vain lyhyen sävellyksen luominen, kun taas hyvä pelimusiikki vaatii pidempiaikaista musiikillista jatkuvuutta. He esittävät viitekehysten, jonka melodiageneraattorissa käytetään Gaussin sekoitemallia (engl. *Gaussian mixture model*, GMM), joka on todennäköisyyteen perustuva koneoppimismalli. (Xiang & Guo, 2021, s. 193–196.)



Kuvio 3 GMM:n yksinkertaistettu toimintamalli (mukaillen Xiang & Guo, 2021, s. 195).

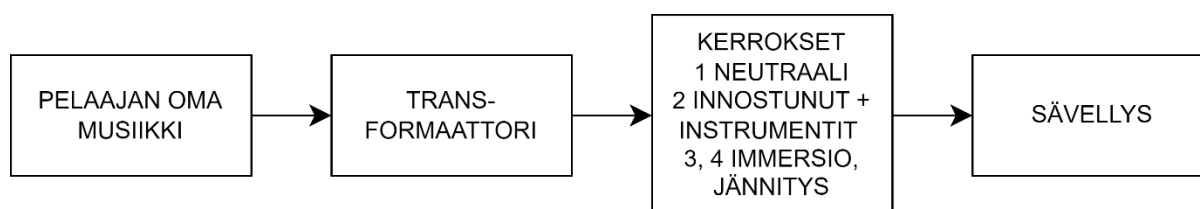
Mallin perustoimintaa on esitetty Kuviossa 3. Mallissa käyttäjä valitsee sävelkorkeuden tai nuottien määrän, jolloin tekoöly luo niihin todennäköisimmin sopivia musiikkivaihtoehtoja. Sen jälkeen toinen tekoöly, joka on koulutettu halutulla musiikilla, valitsee sopivimman vaihtoehdon. Täten käyttäjältä ei vaadita erityistä säveltämistäittoa. Melodiageneraattori pystyy noudattamaan samanaikaisesti niin ihmisen kuin tekoölyn antamia ohjeita. (Xiang & Guo, 2021, s. 193.) Viitekehyksessä melodia esitetään äänenkorkeuden ja äänentaajuuden muutoksina. Esimerkiksi rytmit

yksinkertaistetaan nuottiarvoiksi, jotka muunnetaan koodiksi, joka sisältää melodioiden pituudet. (Xiang & Guo, 2021, s. 196.)

Xiang ja Guo (2021) väittävät GMM:n olevan esimerkiksi RNN:nä parempi ratkaisu musiikin tuottamiseen. GMM:ää on helpompi käsitellä, kun taas RNN vaatii sen käyttäjältä suurta osaamista joko säveltämisestä tai tekoälystä (Xiang & Guo, 2021, s. 195–196). Kyselytutkimuksessa vastaajat arvioivat ja vertailevat eri tekoälyillä tuotettua generatiivista pelimusiikkia. Musiikkia ei kuitenkaan testata pelin avulla, vaan pelkästään kuunnellen. Tulosten mukaan GMM:llä voidaan automatisoida sävellystapaa, joka seuraa ihmisen antamia ohjeita. Sävelletty musiikki on polyfonista MIDI-musiikkia. Kyselytutkimuksessa GMM:n tuottama musiikki arvioidaan paremmaksi, kuin RNN:n tuottama musiikki. (Xiang & Guo, 2021, s. 201–203.)

3.1.4 Transformaattori (Transformer)

Amaralin ja kollegoiden (2022) tavoitteena on luoda pelimusiikin luomiseen soveltuva muokattava malli, jota pelin pelaaja pystyy kouluttamaan haluamallansa musiikilla. Mallin arkkitehtuuri pohjautuu transformaattori- eli muuntajamalliin (engl. *Transformer*), ja sen toimintamalli on esitetty Kuviossa 4. Transformaattori on tekoäly, jonka muisti on pitkäaikaisempi kuin RNN:n (Shih, Wu, Zalkow, Müller & Yang, 2023, s. 3495). Muistin ansiosta transformaattori pystyy luomaan harmonista, tyyliään jatkuvaa musiikkia (Amaral ym., 2022, s. 1; Shih ym., 2023, s. 3495).



Kuvio 4 Transformaattorin yksinkertaistettu toimintamalli (mukaillen Amaral, Baffa, Briot, Feijo & Furtado 2022).

Amaral ja muut (2022, s. 4) vertaavat mallia AMS:ään, sillä molemmissa käytetään neuroverkkoja ja emotiomallinnusta. Tässä mallissa emotiomallinnus tapahtuu kuitenkin kategorioiden sijasta Russellin (1980) affektimallin pohjalta. Kehittäjien mukaan transformaattori ilmaisee vahvemmin musiikin affektiivisuutta sekä onnistuu paremmin kerrostamaan musiikkia (Amaral ym., 2022, s. 4). Esimerkiksi Ferreiran, Limongin ja Fáveron (2023, s. 16) tutkimuksessa, jossa verrataan erilaisia malleja symbolisen musiikin luomiselle, transformaattorimalli menestyy sokkotestauksessa jopa ihmisen säveltämää musiikkia paremmin. Tutkimukseen osallistuneista vastaajista

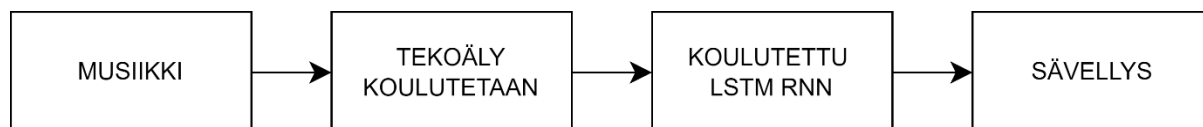
etenkin klassista musiikkia harrastaneet erottavat paremmin tekoälyllä sävelletyn musiikin ihmisen säveltämästä musiikista (Ferreira ym., 2023, s. 16).

Mallin koulutuksessa on käytetty ambient-musiikin soittolistaa, jonka 165 kappaletta on muunnettu polyfonisiksi MIDI-tiedostoiksi. Eri tyyliä lisätään musiikkiin kerroksittain: ensimmäinen tuottaa tavallista ja neutraalin kuuloista musiikkia, toinen lisää musiikkiin innostuneisuutta täydentäen musiikkia esimerkiksi instrumentein ja loput kaksi kerrosta lisäävät immersiota ja jännitystä. (Amaral ym., 2022, s. 5.) Mallista ei ole julkaistu vielä tutkimusta, joka arvioisi musiikin laatua tai toimivuutta pelissä.

3.1.5 LSTM RNN (Long Short-Term Memory Recurrent Neural Network)

Bhardwaj'n ja kumppaneiden (2022) tutkimuksen tavoitteena on luoda tekoälyllä nuotintettua musiikkia. Bhardwaj ja muut (2022) käyttävät mallissaan RNN:ää ennustamaan nuottisekvenssisistä seuraavan nuotin. Tässä tutkimuksessa RNN:ään on lisätty pitkä lyhytkestoinen muisti (engl. *Long Short-Term Memory, LSTM*). Tämän lisäyksen avulla malli voi tunnistaa ja koodata paremmin tiettyjä musiikillisia rakenteita ja kaavoja, joita sävellyksessä yleisesti käytetään tuottamaan yhtenäisemmän kuuloista musiikkia. (Bhardwaj ym., 2022, s. 194.)

LSTM RNN-malli on koulutettu *Final Fantasy* -pelisarjan pianolla soitettulla musiikilla. Mallin perustoiminta on esitetty Kuviossa 5. Musiikki muunnetaan neuroverkkojen avulla MIDI-tiedostoiksi, joista erotellaan nuotit ja soinnut. Jokainen nuotti sisältää tietoa esimerkiksi sävelkorkeudesta ja oktaavista. Tämän jälkeen tiedot muunnetaan LSTM:lle sopiviksi syötteiksi. Seuraavassa vaiheessa luodaan neuroverkko-malli musiikkidatan ennustamiselle. Kun malli on koulutettu onnistuneesti, voidaan luoda generatiivista musiikkia ja nuotinnos siitä. (Bhardwaj ym., 2022, s. 195–197.)



Kuvio 5 LSTM RNN:n yksinkertaistettu toimintamalli (mukaiillen Bhardwaj, Salim, Khan & Javadimasoudian 2022, s. 196).

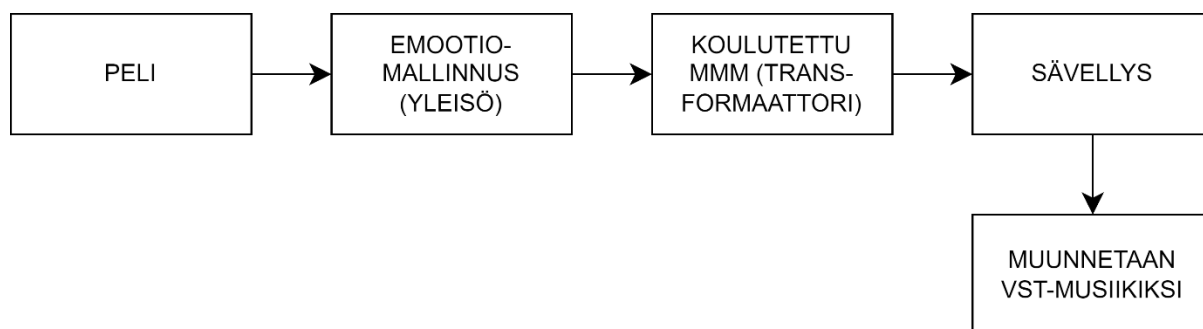
Bhardwajin ja kumppaneiden (2022) mukaan LSTM RNN-malli vaikuttaa lupaavalta, sillä se pystyy luomaan monofonista MIDI-musiikkia ilman, että mallin käyttäjältä vaaditaan musiikillista osaamista. Tulevaisuudessa mallia halutaan kehittää niin, että se tuottaa sekä polyfonista että yksityiskohtaisempaa musiikkia, mikä vaatii suu-remman koulutusdatan käyttämistä ja parempaa suorituskykyä tietokoneelta.

(Bhardwaj ym., 2022, s. 198.) Mallin toimivuutta ei ole arvioitu esimerkiksi kyselytutkimuksen avulla.

3.1.6 PreGLAM-MMM (Predictive Gameplay-based Layerd Affect Model - Multi-Track Music Machine)

Plut ja kollegat (2022) ovat kehittäneet kaksiosaisen pelitilannetta ennustavan mallin (engl. *Predictive Gameplay-based Layerd Affect Model - Multi-Track Music Machine, PreGLAM-MMM*), joka luo affektiivista musiikkia pelaajan voitoille ja häviöille kannustaen pelaajaa. Mallissa pohjataan Phillipsin (2014, s. 109) ideaan musiikin funktiosta toimia yleisönä pelaajan toiminnalle. PreGLAM-mallin emootiomallinnus perustuu Schimackin & Grobin (2000) kolmiulotteiseen VAT-malliin (engl. *Valence-Arousal-Tension, VAT*). Plutin, Pasquierin, Ensin ja Bouguengin (2023, s. 3) mallissa se tuottaa pelin aikana reaaliajassa tietoa ”yleisön” valenssin ja vireystilan lisäksi jännittyneisyydestä (engl. *tension*). PreGLAM:in yhdistetään moniraitainen musiikkikone (engl. *Multi-Track Music Machine, MMM*), jolloin voidaan tuottaa transformaattoritekoälyn avulla vaihtelevaa, monista instrumenteista kerrostettua symbolista musiikkia (Plut ym., 2022, s. 2).

Aikaisempi tutkimus on keskittynyt ratkomaan generatiivisen musiikin teknologisia rajoituksia, kun taas Plut ja kollegat haluavat soveltaa sen käyttöä pelien ominaispiirteiden tukemisessa. Malli ei luo täysin generatiivista musiikkia, vaan laajentaa valmiiksi sävellettyä musiikkia. (Plut ym., 2022, s. 2.) Se toimii siis tältä osin samantyyllisesti AMS:n kanssa.



Kuvio 6 PreGLAM-MMM:n yksinkertaistettu toimintamalli (mukaiillen Plut, Pasquier, Ens & Tchemeube, 2022).

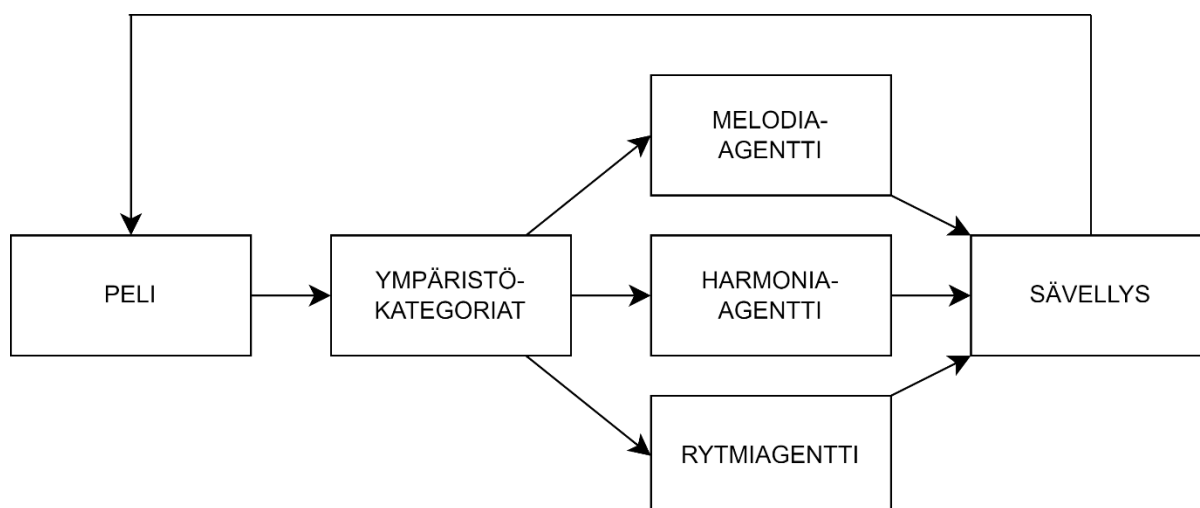
MMM-malli on koulutettu tuhansilla MIDI-tiedostoilla ja se toimii viidessä vaiheessa (kts. Kuvio 6). Ensiksi musiikkiraitaa täydennetään, minkä jälkeen valitaan tahti. Seuraavaksi säädetään instrumenttien ominaisuudet ja nuottien sointitiheys, jonka jälkeen musiikin generointi voi tapahtua. (Ens & Pasquier, 2020, s. 6–8.)

PreGLAM-MMM:ssa melodia ja rytmi voidaan sovittaa ryhmäksi, joten melodiaa voidaan luoda joko olemassa olevan rytmitahtien perusteella tai toisin päin. Musiikin tuottamisprosessi on tietokoneelle kuitenkin niin raskas, ettei musiikkia voida toistaiseksi tuottaa peliin kanssa reaaliajassa. (Plut ym., 2022, s. 7.) PreGLAM-MMM -mallissa käytetään VST-instrumentteja (engl. *Virtual Studio Technology*) MIDI-instrumenttien sijaan, jotta musiikki olisi laadukkaampaa (Plut ym., 2022, s. 2). VST-instrumentteja käytetään elektronisen musiikin tekemisessä ja ne toimivat samalla tavalla kuin MIDI-instrumentit, mutta vain virtuaalisesti. VST-instrumenttien käyttö on lähempänä peliteollisuuden julkaisujen äänimaailmaa, mutta MIDI-äänten käyttö on helpompaa ja halvempaa etenkin tutkimuskäytössä (Plut ym., 2022, s. 1–2).

Plutin ja kumppaneiden (2022, s. 5) mallia varten on kehitetty oma peli. Mallin toimivuutta arvioidaan tutkimuksessa, jossa pelivideoita varten on sävelletty lineaarista sekä adaptiivista pelimusiikkia, jotka toimivat tutkimuksessa verrokkina PreGLAM-MMM:n generatiiviselle musiikille (Plut ym., 2022, s. 2). Kyselytutkimuksen tuloksissa lineaarinen sävellys arvioidaan parhaimmaksi, mutta generatiivinen musiikki ei jää pisteytyksessä kauas etenkään immersion, jatkuvuuden ja musiikillisen miellyttävyyden kohdalla. Tutkijat perustelevat lineaarisen musiikin suosioita esimerkiksi sen sujuvilla transiatioilla. (Plut ym., 2022, s. 8.)

3.1.7 PAMG (Progressive Adaptive Music Generator)

Lopez Duarte (2023, s. 6–7) nostaa esille tekijänoikeusongelmat tekoälyä käyttävissä generatiivisen musiikin malleilla ja käyttää sen sijaan algoritmeja, joiden mukaan musiikki syntyy. Progressiivinen adaptiivisen musiikin musiikkigeneraattori (engl. *Progressive Adaptive Music Generator, PAMG*) toimii moniagenttisesti ja sen käyttämät algoritmit ovat stokastisia ja sääntöpohjaisia (Lopez Duarte, 2023, s. 13, 55). Stokastinen tarkoittaa sitä, että järjestelmä generoi syötteen satunnaisesti. Sääntöpohjainen algoritmi taas käyttää opittuja tai ohjelmoituja sääntöjä, joiden mukaan luominen tapahtuu. (Plut & Pasquier, 2020, s. 10.) Esimerkiksi peliteollisuudessa lyhyitä musiikkisampleja järjestellään stokastisesti (Plut ym., 2022, s. 2). Sweetin (2015, s. 243) mukaan sääntöpohjaista, tonaalista musiikkia sävelletessä musiikki ei välttämättä kuulosta aidolta, mutta atonaalisen musiikin luonti stokastisesti voisi kuulostaa ”oikeammalta” musiikilta. Juuri näiden algoritmien yhteiskäytöllä tarkoituksena on ehkä saada MIDI-musiikki kuulostamaan paremmalta.



Kuvio 7 PAMG:n yksinkertaistettu toimintamalli (mukaillen Lopez Duarte, 2023, s. 43).

PAMG:ssa on rakenteellisia samankaltaisuuksia Hutchingsin ja McCormackin (2020) AMS:n kanssa, sillä järjestelmä toimii moniagenttisesti (kts. Kuvio 7). PAMG:ssa ei kuitenkaan käytetä emootiomallinnusta. Järjestelmässä on melodia-agentti, harmonia-agentti sekä rytmiagentti (Lopez Duarte, 2023, s. 52). Se luo symbolista musiikkia, jossa esimerkiksi sävelkorkeus ja pituus ovat numeerisena tietona. Esimerkiksi harmonia-agentti luo ainutlaatuista musiikkia ja se toimii mallin luoneen käyttäjän musiikillisen tiedon ja osaamisen pohjalta. (Lopez Duarte, 2023, s. 55–56.) Mallissa on myös dissonanssia lisäävä osa sekä instrumenttien yleistä äänenvoimakkuutta muokkaava osa, jonka tarkoituksena on tehdä musiikista ikään kuin ihmisen soittaman kuuloista (Lopez Duarte, 2023, s. 63–64).

PAMG-järjestelmää varten on luotu peli, jossa pelimaailman alue on jaettu näkymättömästi neljään kategoriaan (kirkas, mysteerinen, jännittävä ja eoppinen). Pelaajan lähestyessä aluetta PAMG muokkaa arvoja esiasetusten mukaisesti, jolloin musiikki muuttuu. (Lopez Duarte, 2023, s. 71–72.) Järjestelmän toimivuutta on tutkittu pelaamalla ja vastaamalla kyselytutkimukseen. Tutkimuksessa verrataan järjestelmän tuottamaa musiikkia peliteollisuudesta tuttuun tapaan yhdistellä ennalta tehtyä musiikkia kokonaisuudeksi. Tutkimustulosten mukaan pelaajat kokevat PAMG:n luoman musiikin hieman peliin sopivampana verrattuna valmiiksi tehtyyn musiikkiin. (Lopez Duarte, 2023, s. 91.)

3.1.8 Yhteenveto

Jokainen malli toimii tekoälyn avulla, mutta sen käyttötavat vaihtelevat. Esitellyt mallit käyttävät generatiivisen pelimusiikin luomiseksi erilaisia menetelmiä, kuten neuroverkkvoja, transformaattoreita, ja tavallisia sääntöpohjaisia algoritmeja. Tämän

lisäksi muutama malli myös soveltaa jonkinlaista emootiomallinnusta parantaakseen musiikin adaptiivisuutta. Eroja löytyy myös tuotetun musiikin muodossa, järjestelmän soveltuvuudessa reaaliaikaiseen musiikin tuottamiseen sekä järjestelmän validoinnissa. Yhtenevyyksiä ja eroavaisuuksia on esitetty tarkemmin Taulukossa 1.

TAULUKKO 1 Yhteenveto tutkielmassa esitellyistä 2020-luvun generatiivisen musiikin malleista ja järjestelmistä.

Malli tai järjestelmä	Tekijä	Vuosi	Algoritmi	Emootiomalli	Musiikillinen tieto	Musiikin esitysmuoto	Reaaliaikaisuus	Validointi
AMS	Hutchings & McCormack	2020	RNN	kategorinen	koulutettu ulkoinen	MIDI	kyllä	kyselytutkimus
PMGG	Washburn & Khosmood	2020	sääntöpohjainen algoritmi	valenssi-vireys	ulkoinen	MIDI	kyllä	kyselytutkimus
GMM	Xiang & Guo	2021	GMM	-	koulutettu	MIDI	ei	kyselytutkimus
Transformaattori	Amaral, Baffa, Briot, Feijo & Furtado	2022	transformaattori	valenssi-vireys	koulutettu	MIDI	-	-
LSTM RNN	Bhardwaj, Salim, Khan & Javadimasoudian	2022	LSTM-RNN	-	koulutettu	MIDI	ei	-
PreGLAM-MMM	Plut, Pasquier, Ens & Tchemeube	2022	transformaattori	VAT	koulutettu ulkoinen	VST	ei	kyselytutkimus
PAMG	Lopez Duarte	2023	sääntöpohjainen ja stokastinen algoritmi	-	ulkoinen	MIDI	kyllä	kyselytutkimus

Taulukon sarakkeiden selitys:

Malli tai järjestelmä: Tutkielmassa esiteltyjen generatiivisen musiikin mallin tai järjestelmän nimi tai lyhenne.

Tekijä: Generatiivisen musiikin mallin tai järjestelmän kehittäjät.

Vuosi: Tutkimusartikkelin julkaisuvuosi.

Algoritmi: Mallissa tai järjestelmässä käytetty tekoälymalli.

Emootiomalli: Mallissa tai järjestelmässä käytettyyn emootiomalliin pohjaava teoria.

Musiikillinen tieto: Mistä mallin tai järjestelmän musiikillinen tieto on peräisin. Ulkoinen tarkoittaa mallin kehittäjää.

Musiikin esitysmuoto: Missä muodossa malli tai järjestelmä muodostaa musiikin.

Reaaliaikaisuus: Tuottaako malli tai järjestelmä musiikkia reaaliaikaisesti.

Validointi: Millä tavalla mallin tai järjestelmän generatiivisen musiikin tuottaminen on tutkimuksellisesti validoitu.

3.2 Generatiivisen pelimusiikin jatkokehitys

Kuten Harkiolakis (2023) esittää, on generatiivista pelimusiikkia käytetty jo Sporessa, joka on isompi pelijulkaisu. Generatiivisen musiikin onnistunut toteutus toimii kannustavana esimerkkinä tutkimuksen kehittämiseksi. Esimerkiksi Hutchings ja McCormack (2020, s. 278) toteavat AMS:n positiivisten tutkimustulosten kannustavan kehittämään järjestelmää eteenpäin. Monia malleja ja järjestelmiä varten on kehitetty myös oma peli. Esimerkiksi Lopez Duarte (2023, s. 116–121) esittää, että PAMG-mallia voisi kokeilla monissa tyyliään erilaisissa peleissä, kun taas Plut ja kollegat (2022, s. 10) pohtivat, että vaikka PreGLAM-MMM toimii testattuun yksinpeliin, ei toteutus toimi välttämättä toisenlaisissa peleissä. Myös Hutchings ja McCormack (2020, s. 278) pohtivat, että AMS voisi toimia muissakin peleissä, etenkin jos sitä kehittää Scirean (2017) mallin tyyliisesti ennakoimaan musiikkia.

Washburn ja Khosmood (2020) eivät käsittele PMGG-mallinsa jatkotutkimusta, kun taas Bhardwaj ja muut (2022, s. 198) pohtivat LSTM RNN -mallin kehittämistä esimerkiksi käyttämällä suurempaa koulutusdataa sekä muokkaamalla sitä niin, että se tuottaisi monofonisen melodian sijaan polyfonista musiikkia. Uusien neuroverkkojen lisääminen vaatisi kuitenkin suurempaa prosessointikykyä (Bhardwaj ym., 2022, s. 198). Xiang ja Guo (2021, s. 203) toteavat mallin jatkokehityksestä, että MIDI muunnettaisiin symbolisesta musiikista tallennetuksi ääneksi. He pitävät tekoälyä lupavana reaaliaikaisen musiikin sävellystekniikkana, mikä parantaa pelaajan immersiota (Xiang & Guo, 2021, s. 203).

Amaralin ja kumppaneiden (2022, s. 6–7) tavoitteena on kehittää transformatorimallia esimerkiksi niin, että musiikin sovittamista voidaan hienosäätää pelikokemuksen mukaisesti, riippumatta valitusta, personoidusta musiikista. Amaral ja kollegat (2022, s. 7) puhuvat generatiivisen pelimusiikin mallin kehittämisestä myös ”seuraavan sukupolven pelimusiikkina”. Toisaalta adaptiivisen pelimusiikin toteuttamiseen keksitään varmasti toisenlaisiakin ratkaisuja. Esimerkiksi Harkiolakis (2023, s. 24) esittää, että teknologian kehittymisen myötä generatiivisen musiikin rinnalle nousee jo tutkittu vaihtoehtoinen tapa, jossa musiikki mukautuu pelissä esimerkiksi pelaajan sydämenlyönnin perusteella. Vaikka tämä ajatus tuntuu toistaiseksi kaukaiselta, voi se olla peliteollisuuden normi jo muutaman vuoden kuluttua.

3.3 Tekoälymusiikin ongelmallisuus

Tekoälyn käyttö musiikin säveltämisessä nostaa esille kysymyksiä tekijänoikeuksista. Esimerkiksi Lopez Duarten (2023) PAMG-mallissa tekoäly käyttää vain sitä musiikkilista tietoa, jonka sen kehittäjäkin omaa. Musiikkidatalla koulutetut mallit, kuten Bhardwaj'n ja kumppaneiden (2022) LSTM RNN tai Amaralin ja kollegoiden (2022) transformaattori, tuottavat musiikkia kopioimalla jo olemassa olevan musiikin tyylin. Jos jotakin tekoälyavusteisesta sävellysohjelmasta pyytäisi luomaan vaikkapa latinorytmin, joka soi tietyssä tempossa, se ei tuntuisi ongelmalliselta, sillä musiikkiteollisuuden sävellysohjelmistoissa on jo tällaisia toimintoja. Musiikkia voidaan tehdä käyttämällä erilaisia luuppeja rytmeistä melodioihin. Esimerkiksi LSTM RNN on kuitenkin koulutettu pelkästään Final Fantasy -pelisarjan musiikilla. Tunnistaisiko sarjan pelejä pelannut henkilö LSTM RNN:n tuottaman musiikin pelisarjan jatkumoksi? Kysymykseksi nouseekin, onko eettisesti oikein ottaa musiikillinen data yhdeltä säveltäjältä ja luoda sen pohjalta "uutta" musiikkia.

Puolustuksena tekoälyllä tuotetulle musiikille voidaan myös pohtia perinteisempää ihmisen säveltämää musiikkia ja sen sävellystapaa. Spiegelin (2018, s. 105) mukaan generatiivisen, logiikkaan perustuvan tai algoritmisen musiikin ja ihmisen säveltämän musiikin ero on epämääräinen. Ennen tietokoneavusteista musiikinluontia sävellyksessä vallitsi vuosisatoja jonkinlaista säännönmukaisuutta, kuten kontrapunktin käyttö tai sonaattimuoto (Spiegel, 2018, s. 105). Mozartin noppapeli toimii myös hyvänä esimerkkinä tästä. Säveltäessään ihminen joko tahallisesti tai tiedostamattaan käyttää aikaisemmin kuulemaansa musiikkia tuotoksensa pohjana (Lopez Duarte, 2023, s. 3). Tämä väittämä pitää varmasti ainakin osittain paikkansa. Pystymme helposti luokittelemaan musiikkia esimerkiksi klassisen tai rockmusiikin välillä ja osaamme nimetä kullekin musiikkilajille tyypillisiä ominaisuuksia. Toisaalta myös nämä musiikkigenret ovat muotoutuneet säveltäjien työn tuloksina.

Perinteisesti on ajateltu, että on osattava laaja-alaisesti musiikinteoriaa ja hallita instrumentti taidokkaasti, jos haluaa tehdä musiikkia. Kun säveltämiseen tarjotaan "oikotie" tekoälyn avulla, generatiivista musiikkia ei arvosteta "musiikkina". Esimerkiksi Xiangin ja Guon (2021) tutkimuksen yhtenä tavoitteena on kuitenkin kehittää generatiivisesta musiikista työkalu säveltämiseen, sillä peliproduktiossa musiikin teettäminen ammattilaissäveltäjällä saattaa olla kallista. Jos generatiivisesta musiikista saadaan kehitettyä toimiva ratkaisu musiikin säveltämiseen, voi se nopeuttaa ja parantaa pienten peliyritysten työprosessia huomattavasti.

4 POHDINTA JA YHTEENVETO

Tässä tutkielmassa käydään läpi seitsemän eri generatiivisen pelimusiikin mallia tai järjestelmää. Jokainen malli toimii tekoälyn avulla, mutta sen käyttötavat vaihtelevat. Luvussa kolme esitellyt mallit käyttävät erilaisia tekoälyjä ja algoritmeja, jotka ovat RNN, sääntöpohjainen algoritmi, GMM, transformaattori, LSTM RNN sekä stokastinen algoritmi. Mallien emootiomallinnus pohjaa Russellin (1980) affektimalliin, Juslinin (2013) kategoriseen emootiomalliin tai Schimmackin ja Grobin (2000) VAT-malliin. Kolme malleista ei käytä emootiomallinnusta. Kahdessa mallissa käyttäjältä vaaditaan musiikkiteoreettista osaamista generatiivisen musiikin luomiseen. Kolmessa mallissa musiikillinen tieto saadaan koulutusdatasta. Kahdessa mallissa käytetään sekä koulutettua dataa että käyttäjän musiikillista osaamista. Kuudessa mallissa musiikki on symbolista MIDI-muodossa olevaa musiikkia ja yhdessä käytetään VST-instrumentaatiota, joka kuitenkin toistaiseksi estää reaaliaikaisen musiikin toistamisen pelitilanteessa. Malleista viittä on tutkittu suorittamalla kyselytutkimus, jolla arvioidaan joko mallin toimivuutta tai sen tuottaman musiikin laatua.

Tutkimusaineiston perusteella generatiivisen pelimusiikin kehittämisen keskiössä on etenkin reaaliaikaisen, affektiivisen musiikin tuottaminen. Musiikki saa ihmisissä aikaan tunnereaktioita, ja pelaaminen on oivallinen media eri tunteiden kokemiseen. Peliteollisuus on tulonlähteenä maailmanlaajuisesti merkittävä, joten teknologian kehittyessä on ymmärrettävää, että myös pelejä halutaan parantaa jatkuvasti. Ei silti ole yllättävää, että Plutin ja muiden (2022) tutkimuksessa lineaarinen musiikki arvioidaan parhaaksi. Elokuvan tarjoamaan audiovisuaaliseen saumattomuuteen on vuosien saatossa totuttu. Pelien ollessa interaktiivisia ja epälineaarisia, ei generatiivinen pelimusiikki ole vielä yltänyt samalle tasolle. Videopelit ovat kuitenkin kehittyneet vuosi vuodelta realistisemmiksi niin toiminnallisuudeltaan, grafiikoiltaan kuin musiikiltaan. Luvussa 3 esiteltyjen mallien pelit ja niiden toteutus eivät kuitenkaan ole verrattavissa peliteollisuuden tuotoksiin, joihin käytetään paljon rahaa ja aikaa.

2020-luvun generatiivinen pelimusiikin tutkimus seuraa selkeästi 2010-luvulla alkanutta suuntausta, jossa käytetään neuroverkkoja ja emootiomallinnusta. Jälkimmäistä ei kuitenkaan käytetä kolmessa mallissa. Näistä kaksi, Bhardwaj'n ja kumppaneiden (2022) sekä Xiangin ja Guon (2021) mallit eivät myöskään luo pelimusiikkia reaaliaikaisesti pelissä, vaan tavoitteena on käyttää generatiivista musiikkia sävellystyökaluna tai muodostaa nuotinnettua musiikkia koulutusdatan pohjalta. Lopez Duarte (2023) siirtää affektisuuden suoraan malliin oman musiikillisen tietonsa kautta, sillä pelimaailman alueet jaetaan esimerkiksi eepiseksi tai mysteeriseksi, jonka myötä musiikki adaptoituu. Tämän perusteella muissa järjestelmissä käytetty emootiomallinnus antaa suurempaa vapautta pelinsisäisille tapahtumille, kun se ei ole rajoittuna tiettyihin alueisiin. Lopez Duarten (2023) ratkaisu on kuitenkin tekijänoikeudellisesta näkökulmasta kestävämpi. Jos taas tutkimuksen tavoitteena on nopeuttaa ja helpottaa sävellysprosessia, on se kestävyuden näkökulman kanssa toistaiseksi ristiriidassa.

Kandidaatintutkielma on pituudeltaan rajattu, minkä vuoksi aiheen monipuolinen käsittely on myös rajallista. Esimerkiksi tutkimuksia varten kehitettyjä pelejä ei avata, vaikka se auttaisi lukijaa ymmärtämään paremmin mallien toimintaa. Narratiivinen kirjallisuuskatsaus menetelmänä ei myöskään sisällä kaikkia mahdollisia lähteitä, sillä aineistoa ei ole haettu systemaattisesti. Kerätty tutkimustieto on vain englanninkielistä, mikä sulkee pois mahdollisen muunkielisen tutkimuksen analysoinnin, mutta antaa myös tilaisuuden tuottaa ja koostaa tietoa suomeksi. Pienempi aineisto mahdollistaa toisaalta myös sen tarkemman ja monipuolisemman tutkimisen.

Generatiivinen pelimusiikki aiheena on moniulotteinen ja vaikeaselkoinen, sillä tutkimus pohjaa enimmäkseen informaatiotieteeseen. Tämän tieteenalan termien kokonaisvaltainen ymmärtäminen tutkimusaineiston pohjalta jää osittain vajavaiseksi. Toisaalta musiikkitieteellinen lähestymistapa tuo aiheeseen täysin erilaisen näkökulman, sillä tutkielmassa pohditaan esimerkiksi tekoälyn käytön ongelmallisuutta verrattuna positivistisempaan näkökulmaan. Valittu aineisto ei välttämättä olisi tarpeeksi suuri informaatiotieteen kandidaatintutkielmaan, sillä tekoälyyn liittyviä termejä tulisi selittää monipuolisemmin. Molempien tieteenalojen näkökulmissa kuitenkin esimerkiksi emootiotutkimuksen teoriat tai affektiivinen musiikki nähdään valituissa tutkimuksissa faktoina, vaikka ne perustuvat hyvin länsimaalaisiin arvoihin ja tapoihin ymmärtää maailmaa.

Murroksessa oleva tekoälymusiikki jakaa varmasti mielipiteitä, mutta teknologian kehittyminen on vaikuttanut aina myös musiikkiin. Esimerkiksi bluesmusiikki on muuttunut vuosikymmenien saatossa elektronisten instrumenttien lisääntymisen myötä, mutta genrenä se on vaikuttanut rockmusiikin ja jazzin syntyyn. Nykypäivän esimerkkinä viime vuosikymmenien teknologisesta muutoksesta toimii autotune, jota käytetään puhtaamman lauluäänen tavoittelussa yli genererajojen. Valtaosa

kuuntelijoista ei edes kiinnitä huomiota sen käyttöön. Teknologia-avusteiseen musiikkiin ollaan jo niin tottuneita, joten tekoälyllä sävelletty musiikki on varmasti myös täysin tavallista musiikkia tulevaisuudessa. Tekoälymusiikin vähäisempi arvostus tuntuu myös paradoksilta, sillä uutta luodessaan ihmiset inspiroituvat jatkuvasti vanhasta.

Kuten elokuvat ja musiikki, myös pelit toimivat ajankuvana. Nykypäivän teknologia ei enää rajoita samalla tavoin kuin vuosikymmeniä sitten, ja peleissä pitää olla jokaisella pelikerralla aina uudenlaista sisältöä, jotta pelaaja saadaan koukuttettua mukaan. Pelien avulla halutaan saavuttaa myös virtuaalinen todellisuus, jolloin elämä lähentelee jo elokuvaa. Generatiivinen musiikki on yksi keino luoda monipuolisuutta, mutta toisaalta monissa vanhemmissa, lineaarisen musiikin peleissä on oma nostalginen viehätöksensä, minkä takia niiden pariin saatetaan palata yhä uudelleen. Silti Amaralin ja kumppaneiden (2022) ajatus generatiivisesta pelimusiikista seuraavan sukupolven pelimusiikkina pitänee hyvin pitkälti paikkaansa. Tulevaisuudessa säveltäjältä ei enää välttämättä vaadita musiikkiteoreettista tietämystä, sillä tieto on siirretty säveltävälle järjestelmälle. Nähtäväksi jää, tuleeko yhä paremman mallin etsinnästä eri tekoälyjen kilpailu vai löydetäänkö generatiivisen musiikin tuottamiseen aivan toisenlainen – ja jopa kestävämpi ratkaisu.

LÄHTEET

- Amaral, G., Baffa, A., Briot, J. P., Feijo, B., & Furtado, A. (2022). An adaptive music generation architecture for games based on the deep learning Transformer model. *Brazilian Symposium on Games and Digital Entertainment, SBGAMES, 2022-October*, 1–6.
<https://doi.org/10.1109/SBGAMES56371.2022.9961081>
- Baio, A. (30.9.2022). *AI Data Laundering: How Academic and Nonprofit Researchers Shield Tech Companies from Accountability*. Waxy.org.
<https://waxy.org/2022/09/ai-data-laundering-how-academic-and-nonprofit-researchers-shield-tech-companies-from-accountability/>
- Balaji, P. G., & Srinivasan, D. (2010). An introduction to multi-agent systems. *Studies in Computational Intelligence*, 310, 1–27.
<https://doi.org/10.1007/978-3-642-14435-6>
- Bhardwaj, S., Salim, S. M., Khan, T. A., & Javadimasoudian, S. (2022). Automated Music Generation using Deep Learning. *International Conference Automatics and Informatics, ICAI 2022 - Proceedings*, 193–198.
<https://doi.org/10.1109/ICAI55857.2022.9960063>
- Clancy, M. (2022). Whose ethics? Approaches to an equitable and sustainable music ecosystem. *Teoksessa Artificial Intelligence and Music Ecosystem*, 109–125. <https://doi.org/10.4324/9780429356797>
- Clark, C. A. (2021). *Gameplay as Discrete Form: Leveraging Procedural Audio for Greater Adaptability in Video Game Music*.
<https://escholarship.org/uc/item/5wj1q3j2>
- Collins, K. (2005). From Bits to Hits : Video Games Music Changes its Tune. *Film International*, 3(1), 4–18.
- Collins, K. (2008). *Game sound: An introduction to the history, theory, and practice of video game music and sound design*. MIT Press.
- Collins, K. (2009). An Introduction to Procedural Music in Video Games. *Contemporary Music Review*, 28(1), 5–15.
<https://doi.org/10.1080/07494460802663983>
- Collins, K. (2013). *Playing with sound : a theory of interacting with sound and music in video games*. MIT Press.
- Eerola, T., & Saarikallio, S. (2010). Musiikki ja tunteet. Teoksessa J. Louhivuori & S. Saarikallio (Toim.), *Musiikkipsykologia* (s. 259–275).
- Engels, S., Tong, T., & Chan, F. (2015). Automatic Real-Time Music Generation for Games. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 11(1), 220–222.
<https://doi.org/10.1609/AIIDE.V11I1.12775>
- Eno, B. (7.7.1996). Generative Music. *In Motion Magazine*.
<https://www.inmotionmagazine.com/eno1.html>
- Ens, J., & Pasquier, P. (2020). MMM : Exploring Conditional Multi-Track Music Generation with the Transformer. <https://arxiv.org/abs/2008.06048v2>
- Farnell, A. (2010). *Designing Sound*. MIT Press.

- Ferreira, P., Limongi, R., & Fávero, L. P. (2023). Generating Music with Data: Application of Deep Learning Models for Symbolic Music Composition. *Applied Sciences (Switzerland)*, 13(7). <https://doi.org/10.3390/APP13074543>
- Fritsch, M., & Summers, T. (2021). *The Cambridge Companion to Video Game Music*. Cambridge University Press. <https://doi.org/10.1017/9781108670289>
- Grand View Research. (ei pvm.). *Video Game Market Size, Share and Growth Report, 2030*. Luettu 20. marraskuuta 2023, osoitteesta <https://www.grandviewresearch.com/industry-analysis/video-game-market>
- Graupe, D. (2013). *Principles of Artificial Neural Networks: 3rd Edition (Advanced Series in Circuits & Systems)* (3rd edition, Vol. 7). World Scientific.
- Harkiolakis, A. (2023). The use of generative music in nonlinear music composition for video games. *International Journal of Technology Enhanced Learning*, 1(1), 1. <https://doi.org/10.1504/ijtel.2023.10055011>
- Huiberts, S. (2010). *Captivating Sound - The Role of Audio for Immersion in Computer Games*.
- Hutchings, P. E., & McCormack, J. (2020). Adaptive Music Composition for Games. *IEEE Transactions on Games*, 12(3), 270–280. <https://doi.org/10.1109/TG.2019.2921979>
- Juslin, P. N. (2013). What does music express? Basic emotions and beyond. *Frontiers in Psychology*, 4(SEP), 60433. <https://doi.org/10.3389/FPSYG.2013.00596>
- Kinnunen, J., Tuomela, M., & Mäyrä, F. (2022). *Pelaajabarometri 2022 : Kohti uutta normaalia*. <https://urn.fi/URN:ISBN:978-952-03-2732-3>
- Krenker, A., Bešter, J., Kos, A., Krenker, A., Bešter, J., & Kos, A. (2011). Introduction to the Artificial Neural Networks. *Artificial Neural Networks - Methodological Advances and Biomedical Applications*. <https://doi.org/10.5772/15751>
- Li, F. (2023). Chord-based music generation using long short-term memory neural networks in the context of artificial intelligence. *Journal of Supercomputing*, 1–25. <https://doi.org/10.1007/S11227-023-05704-3>
- Lidy, T., & Rauber, A. (2009). Music information retrieval. *Handbook of Research on Digital Libraries: Design, Development, and Impact*, 448–456. <https://doi.org/10.4018/978-1-59904-879-6.CH046>
- Lopes, P., Liapis, A., & Yannakakis, G. N. (2015). Sonancia: Sonification of Procedurally Generated Game Levels. *Proceedings of the 1st Computational Creativity and Games Workshop*.
- Lopez Duarte, A. E. (2023). *A Progressive-Adaptive Music Generator for Videogames (PAMG): an Approach to Real-Time Algorithmic Composition*.
- López Ibáñez, M., Álvarez, N., & Peinado, F. (2018). Towards an emotion-driven adaptive system for video game music. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10714 LNCS, 360–367. https://doi.org/10.1007/978-3-319-76270-8_25

- McLean, A., & Dean, R. T. (2018). Musical Algorithms as Tools, Languages, and Partners: A Perspective. Teoksessa *The Oxford Handbook of Algorithmic Music* (s. 3–16). Oxford University Press.
<https://doi.org/10.1093/OXFORDHB/9780190226992.013.24>
- Phillips, W. (2014). *A Composer's Guide to Game Music*. The MIT Press.
- Pine II, B. J., & Gilmore, J. (1999). *The Experience Economy: Work Is Theater & Every Business a Stage*. Harvard Business School Press.
- Plut, C. (2022). *Application and evaluation of affective adaptive generative music for video games*.
- Plut, C., & Pasquier, P. (2020). *Generative music in video games: State of the art, challenges, and prospects*. <https://doi.org/10.1016/j.entcom.2019.100337>
- Plut, C., Pasquier, P., Ens, J., & Bougueng, R. (2023). PreGLAM: A Predictive, Gameplay-based Layered Affect Model. *IEEE Transactions on Games*.
<https://doi.org/10.1109/TG.2023.3287732>
- Plut, C., Pasquier, P., Ens, J., & Tchemeube, R. (2022). PreGLAM-MMM Application and evaluation of affective adaptive generative music in video games. *ACM International Conference Proceeding Series*.
<https://doi.org/10.1145/3555858.3555947>
- Precht, A. (2015). *Adaptive Music Generation for Computer Games*. The Open University, UK.
- Rebala, G., Ravi, A., & Churiwala, S. (2019). *An Introduction to Machine Learning*.
<https://doi.org/10.1007/978-3-030-15729-6>
- Russell, J. A. (1980). A Circumplex Model of Affect. Teoksessa *Journal of Personality and Social Psychology* (Vol 39, Numero 6). Ryman.
- Russell, S. J., & Norvig, Peter. (1995). *Artificial intelligence : a modern approach*. Prentice Hall.
- Salminen, A. (2023). *Mikä kirjallisuuskatsaus? Johdatus kirjallisuuskatsauksen tyyppeihin ja joihinkin hallintotieteellisiin sovelluksiin*.
<https://urn.fi/URN:ISBN:978-952-395-082-5>
- Schimmack, U., & Grob, A. (2000). Dimensional models of core affect: A quantitative comparison by means of structural equation modeling. *European journal of personality*, 14(4), 325–345.
- Scirea, M. (2017). *Affective Music Generation and its effect on player experience*.
- Shih, Y. J., Wu, S. L., Zalkow, F., Muller, M., & Yang, Y. H. (2023). Theme Transformer: Symbolic Music Generation With Theme-Conditioned Transformer. *IEEE Transactions on Multimedia*, 25, 3495–3508.
<https://doi.org/10.1109/TMM.2022.3161851>
- Siren, V. (21.11.1993). Tietokonemusiikin alussa oli Mozart Kolmetoista miljoonaa pikkukappaletta "säveltävä" noppapeli edelsi nykyisiä säveltäjän työvälineitä Koneet pystyvät kartoittamaan tuntemattomia sointimaailmoja - tai jäljittelemään ennestään tunnettuja. *Helsingin Sanomat*. <https://www.hs.fi/kulttuuri/art-2000003285415.html>
- Spiegel, L. (2018). Thoughts on Composing with Algorithms. Teoksessa *The Oxford Handbook of Algorithmic Music* (s. 105–112). Oxford University Press.
<https://doi.org/10.1093/OXFORDHB/9780190226992.013.26>

- Sweet, M. (2015). *Writing Interactive Music for Video Games: A Composer's Guide*.
- van Elferen, I. (2016). Analyzing game musical immersion: the ALI model. Teoksessa M. Kamp, T. Summers, M. Sweeney, & J. B. Masse (Toim.), *Ludomusicology: Approaches to video game music* (s. 32–52).
- Vesala, J. (2023). Developing Artificial Intelligence-Based Content Creation: Are EU Copyright and Antitrust Law Fit for Purpose? *IIC*, 54, 351–380.
<https://doi.org/10.1007/s40319-023-01301-2>
- Washburn, M., & Khosmood, F. (2020). Dynamic Procedural Music Generation from NPC Attributes. *ACM International Conference Proceeding Series*.
<https://doi.org/10.1145/3402942.3409785>
- Wooller, R., Brown, A. R., Miranda, E., Berry, R., & Diederich, J. (2005). A framework for comparison of processes in algorithmic music systems. *Generative Arts Practice*, 109–124.
- Xiang, Z., & Guo, Y. (2021). Controlling Melody Structures in Automatic Game Soundtrack Compositions with Adversarial Learning Guided Gaussian Mixture Models. *IEEE Transactions on Games*, 13(2), 193–204.
<https://doi.org/10.1109/TG.2020.3035593>