

Konsta Torniainen

Oppimisanalytiikan ilmentymistä ja haasteista

Tietotekniikan kandidaatintutkielma

17. joulukuuta 2023

Jyväskylän yliopisto

Informaatioteknologian tiedekunta

Tekijä: Konsta Torniainen

Yhteystiedot: torniake@jyu.fi

Ohjaaja: Tytti Saksa

Työn nimi: Oppimisanalytiikan ilmentymistä ja haasteista

Title in English: Manifestations and challenges of learning analytics

Työ: Kandidaatintutkielma

Sivumäärä: 25+0

Tiivistelmä: Tutkielmassa tarkastellaan oppimisanalytiikan ilmentymiä ja haasteita pedagogisessa ja teknisessä mielessä. Tutkittuja ilmentymiä ovat erilaiset oppimisanalytiikan menetelmät, toimijoiden roolit, erilaiset tietojärjestelmät ja tietotekninen perusta. Tutkittuja haasteita on oppimisanalytiikan tarkoituksenmukainen hyödyntäminen, tietojärjestelmien standardointiin liittyvät haasteet ja suunnitelmallisuuden haasteet.

Avainsanat: oppimisanalytiikka, koulutuksessa syntyvän datan louhinta, big data, big data koulutuksessa

Abstract: This thesis examines the manifestations and challenges of learning analytics from both pedagogical and technical perspectives. Explored manifestations include various methods of learning analytics, the roles of different stakeholders, roles of information systems, and the technological foundation. Investigated challenges encompass the effective utilization of learning analytics, issues related to the standardization of information systems, and challenges associated with systematic planning.

Keywords: learning analytics, educational data mining, big data, big data in education

Kuviot

Kuvio 1. EDM ja oppimisanalytiikan ulottuvuudet (Romero ja Ventura 2020).	3
Kuvio 2. Graafi, joka on piirretty Moodle-opetusjärjestelmään asennetun liitännäisen avulla (Schmitt 2015).	9

Sisällys

1	JOHDANTO	1
2	TEOREETTINEN VIITEKEHYS: DATASTA ANALYTIikkaAN, OPETUS- JÄRJESTELMIIN JA PEDAGOGIIKKAAN	2
3	PEDAGOGISET MAHDOLLISUUDET JA HAASTEET	6
3.1	Oppimisanalytiikan jako toimijoiden suhteen	6
3.2	Oppimisanalytiikan yleisempiä metodeja ja niiden pedagogisia käyttökohteita	7
3.3	Oppimisanalytiikan pedagogisia haasteita	9
4	TEKNISET MAHDOLLISUUDET JA HAASTEET	11
4.1	Teknilliset ilmentymät	11
4.2	Teknillisiä haasteita	13
5	TEKNOLOGIAN, ANALYYSIN JA PEDAGOGIIKAN KOHTAAMINEN VAA- TII KÄYTTÄJIEN HUOMIOIMISTA JA SUUNNITELMALLISUUTTA.....	15
6	JOHTOPÄÄTÖKSET.....	17
	LÄHTEET	19

1 Johdanto

Tämän tutkielman aiheena on oppimisanalytiikka, joka on tutkimusalueena kasvanut suuresti viimeisen vuosikymmenen aikana (Romero ja Ventura 2020, 12). Nyky-yhteiskunnassa ihmiset jättävät runsaasti digitaalista jälkeä, ja kasvatustieteissä on havaittu, että tätä syntyvää dataa voidaan hyödyntää eri tavoin opetuksessa ja oppimisessa (Auvinen 2017, 3; Romero ja Ventura 2013, 12). Oppimisanalytiikalla tarkoitetaan opetuksen järjestämisen seurauksena syntyneen datan prosessointia ja hyödyntämistä opetuksessa (Opetus- ja kulttuuriministeriö 2021).

Oppimisanalytiikka mahdollistaa datan hyödyntämisen sekä oppilaiden että opettajien näkökulmasta (Opetus- ja kulttuuriministeriö 2021, 27). Oppimisanalytiikan avulla oppilaat voivat seurata omaa suoriutumistaan ja käyttää kerättyä dataa apunaan. Tämä voi auttaa heitä ymmärtämään vahvuuksiaan ja heikkouksiaan sekä tarjota mahdollisuuden parantaa oppimistuloksiaan. Opettajien näkökulmasta data voi olla arvokas työkalu ohjauksessa ja opetustyön suunnittelussa. Se voi auttaa opettajia tunnistamaan oppilaiden mahdolliset ongelmatkohdat ja tarjoaa mahdollisuuden räätälöidä opetusta paremmin oppilaiden tarpeisiin. Tästä jaottelusta löytyy lisää tietoa tämän tutkielman luvusta 3, jossa jaottelua käydään läpi opetus- ja kulttuuriministeriön luoman viitekehyksen pohjalta. Tutkielman teoriaosuuden alussa oppimisanalytiikka jaetaan myös tietotekniseen, pedagogiseen ja tilastotieteelliseen jakoon Romeron ja Venturan teoksen pohjalta (Romero ja Ventura 2020, 2). Tutkielmassa kiinnitetään huomiota pedagogisen ja tietoteknisen näkökulman ilmentymiin ja haasteisiin.

Oppimisanalytiikka valikoitui kirjallisuuskartoituksen aiheeksi, sillä aihe on ajankohtainen ja kasvavan kiinnostuksen kohteena (Romero ja Ventura 2020, 12). Lisäksi oppimisanalytiikalla on yhteiskunnallisen merkityksen kannalta potentiaalia, mikä on tärkeää. Tutkielman teoriaosuus esittelee oppimisanalytiikan ulottuvuuksia ja kontekstia. Kolmannessa luvussa kiinnitetään huomiota ilmiön pedagogiseen puoleen toimijoiden, metodien ja haasteiden kautta. Neljännessä luvussa tarkastellaan teknisiä ilmentymiä esittelemällä yksi oppimisanalytiikan sovellusesimerkki ja pohjustetaan teknistä puolta ohjelmistostandardien avulla. Viidennessä luvussa tehdään huomioita, kuinka aiemmin tarkastellut ilmentymät ja haasteet ovat riippuvaisia riittävästä käyttöönoton suunnittelusta ja käyttäjien huomioimisesta.

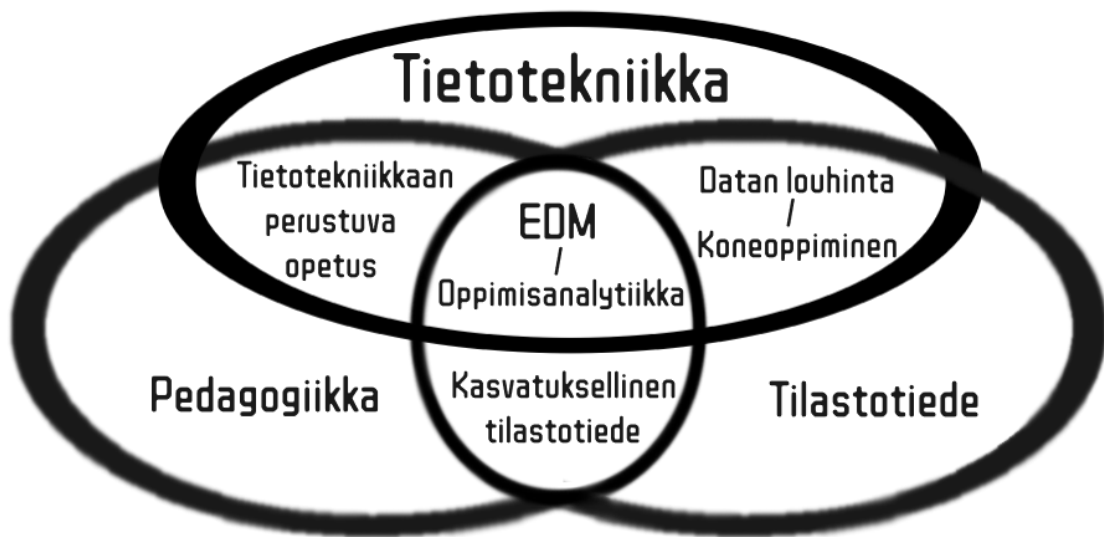
2 Teorettinen viitekehys: Datasta analytiikkaan, opetusjärjestelmiin ja pedagogiikkaan

Tarkemmin oppimisanalytiikalla (engl. *learning analytics*, LA) tarkoitetaan oppijoiden ja heidän oppimisympäristöjen tietojen mittaamista, analysoimista ja raportointia. Tämän toiminnan tarkoituksena on kehittää oppimista ja sen ympäristöjä (Lang ym. 2017, 9). On hyvä huomata, että oppimisanalytiikkaan ja sen tutkimukseen liittyy tutkimussuuntauksen tuoreuden kannalta paljon termejä, joille ei vielä ole kunnolla vakiintunut suomenkielisiä termejä, joten tässä tekstissä tulee ilmenemään myös englanninkielisiä termejä.

Oppimisanalytiikka ja siihen liittyvä käsite *educational data mining* (EDM), eli koulutuksessa syntyvän datanlouhinta jakavat ilmiön kolmen tieteenalan pohjalta: tietotekninen, kasvatuksellinen/pedagoginen ja tilastotieteellinen (Romero ja Ventura 2020, 2). (Katso kuvio 1). Oppimisanalytiikka ja EDM hyödyntää näiden eri tieteenalojen konsepteja kehittääkseen ja tutkiakseen keinoja, joilla suuresta määrästä dataa voidaan löytää toistuvia opetuksen kehityksen kannalta hyödyllisiä malleja. Ilman näitä EDM tarjoamia keinoja olisi näin suurista tietomassoista mahdotonta löytää ja analysoida vastaavaa tietoa (Romero ja Ventura 2020, 2).

Romero ja Ventura (2020) kuvaavat artikkelissaan aiemmin mainittujen kolmen tieteenalan risteyskohtaa ja EDM:n muodostumista (pedagoginen, tietotekninen ja tilastotieteellinen jako). Pedagogisen ja tilastotieteellisen risteyskohdan muodostaa kasvatuksellinen tilastotiede (engl. *educational statistics*). Tilastotieteen ja tietotekniikan leikkauksessa on datan louhinta ja koneoppiminen. Tietoteknisen puolen ja kasvatustieteen risteyksessä on tietotekniikkaan perustuva pedagogiikka. Kuten jo johdannossa mainittu, niin tässä kirjallisuuskartoituksessa keskitytään lähinnä pedagogiseen ja tietotekniseen näkökulmaan, eikä tilastotieteen metodeihin ja kaavoihin mennä syvälle.

Oppimisanalytiikan taustalla on niin kutsuttu *big data*-ilmiö. Big datalla viitataan nykyiseen datavetoiseen yhteiskuntaan, jossa yhteiskunnan jokaiselle sektorille on tyypillistä kertyvän ja käsiteltävän datan suuri ja kasaantuva määrä (Auvinen 2017, 5). Tästä termistä on koulutussektorille oma versio big data koulutuksessa (engl. *big data in education*, BDE), jol-



Kuvio 1. EDM ja oppimisanalytiikan ulottuvuudet (Romero ja Ventura 2020).

la tarkoitetaan tämän big datan käsittelyn tekniikoiden siirtämistä koulutuksen ympäristöön (Daniel 2019). Tämän kanssa saman kaltainen termi datavetoinen opetus (engl. *data driven education*, DDE) tarkoittaa systemaattista koulutusdatan keräämistä ja analysoimista, jotta voitaisiin neuvoa oppilaita sekä koululaitoksia (Romero ja Ventura 2020, 2). BDE ja DDE ilmenee kuviossa 1 juurikin tietotekniikan ja pedagogiikan leikkauksessa.

Sen lisäksi, että termeille ei ole kunnolla vakiintunut suomenkielisiä muotoja, on hyvä huomata, että oppimisanalytiikan isommassa viitekehyksessä on useampia termejä, jotka muistuttavat toisiaan ja joiden merkityksellä on jonkin verran päällekkäisyyttä keskenään. Esimerkiksi termit akateeminen analytiikka (engl. *academic analytics*, AA), instituutioanalytiikka (engl. *institutional analytics*, IA) ja opetusanalytiikka (engl. *teaching analytics*, TA) voivat ensi silmäykseltä vaikuttaa hyvin samalta kuin oppimisanalytiikka (engl. *learning analytics*). Näillä termeillä onkin päällekkäisyyttä, mutta myös eroavaisuuksia esimerkiksi siinä minkä toimijan kautta ilmiötä tarkastellaan. Akateeminen analytiikka ja instituutioanalytiikka tarkastelevat ilmiötä koulutuksen pedagogisen tarkastelun sijasta enemmän ekonominen haasteen kautta. Tälle tarkastelun tavalle on tyypillistä kiinnittää huomiota opiskelijamaksuihin, kurssiarviointeihin, tutkinto-ohjelmiin, resurssien jakoon/hallintaan ja näiden eri teemojen datan keräämiseen, analysointiin sekä raportointiin oppilaitoksen kehittämisen

vuoksi. Opetusanalytiikka sen sijaan tarkastelee ilmiötä ohjaajan tai opettajan näkökulmasta. Tässä näkökulmassa kiinnostaa opetustoimintojen suunnittelu ja kehittäminen (Romero ja Ventura 2020, 2). Oppimisanalytiikkaa voidaan pitää kattavana käsitteenä, joka yhdistää näitä eri näkökulmia. Tekstissä tullaan kuitenkin huomaamaan, että tarkastelemalla ilmiötä eri toimijoiden näkökulmasta saadaan myös erilaisia haasteita oppimisanalytiikan hyödyntämiselle.

Perinteisen luokkahuoneopetuksen keräämät tiedot kuten osallistuminen, arvosanat ja opintojen eteneminen ovat suhteellisen perinteisiä mittareita opetuksen tehokkuudesta. Viime vuosikymmeninä siirtyminen enemmän verkkopohjaiseen opetukseen (engl. *computer based education*, CBE) on avannut uusia mahdollisuuksia oppimisanalytiikalle (Romero ja Ventura 2020, 10). Verkkopohjaisista oppimisympäristöistä voidaan kerätä runsaasti tietoa, esimerkiksi sovellusten lokitietojen kautta, jotka kattaa lähes jokaisen verkossa tapahtuvan interaktion, kuten napin painallukset. Tämä automaattinen tiedonkeruu on tehokkaampaa ja mahdollistaa laajemman datan hyödyntämisen, mikä tarjoaa syvällisemmän kuvan opiskelijoiden oppimisesta ja käyttäytymisestä.

Romero ja Ventura (2020) ovat teoksessaan listanneet erilaisia verkkopohjaisten opetusjärjestelmien tyyppisiä. Näiden eri tyyppien hahmottaminen on tärkeää, sillä ne tarjoavat erilaisia mahdollisuuksia sekä opetuksen että oppimisanalytiikan kannalta. Opetushallintajärjestelmä (engl. *learning management systems*, LMS) tarkoittaa ympäristöä, joka kerää käyttäjistä tietoa eri aktiviteeteista kuten kirjoittamisesta, tehtävien teosta ja testien tekemisestä (Romero ja Ventura 2020, 11). Suomessakin käytössä oleva Moodle-järjestelmä on esimerkki tämän kaltaisesta järjestelmästä. Myös Suomessa suhteellisen yleiset MOOC-kurssit (engl. *massive open online course*) muodostavat toisenlaisen CBE-järjestelmän tyyppin. Näiden ympäristöjen tavoitteena on tarjota lähes rajoittamattomalle määrälle osallistujia oppimisympäristö, jossa nämä järjestelmät keräävät samankaltaista tietoa kuin LMS-järjestelmät. Tämä ”lähes rajoittamaton määrä osallistujia” on juuri se, miksi oppimisanalytiikkaan olisi hyvä suunnata huomiota – se voi todistautua erittäin arvokkaaksi työkaluksi opetuksen ja arvioinnin tueksi, kun halutaan tutkia suurta määrää osallistujia ja heidän luomaa dataa.

Hieman erilaisia järjestelmiä ovat adaptiiviset ja älykkäät hypermedian järjestelmät (engl. *adaptive and intelligent hypermedia system*, AIHS) ja älykkäät opetusjärjestelmät (engl. *in-*

telligent tutoring system, ITS). AIHS-järjestelmät pyrkivät luomaan opiskelijoilleen henkilökohtaisen mallin heidän tavoitteistaan ja tiedoistaan, jonka jälkeen mallia käytetään vuorovaikutuksessa opiskelijan kanssa, jotta järjestelmä soveltuisi tämän tarpeisiin (Romero ja Ventura 2020, 11). ITS-järjestelmät pyrkivät tarjoamaan käyttäjilleen räätälöityä opetusta tarkastelemalla opiskelijan käyttäytymistä verkkoympäristössä, kuten hiiren klikkausten ja kirjoittamisen kuuntelun avulla (Romero ja Ventura 2020, 11). Nämä järjestelmät eroavat aiemman kappaleen LMS-järjestelmistä siinä, että nämä ovat järjestelmiä, jotka osaavat tehdä kerätyllä datalla esimerkiksi oppilaan kohdalla henkilökohtaisia toimenpiteitä datan pohjalta. Tämän tekstin kirjoitushetkellä ei kirjoittajan tiedossa ole, että onko tällaisia järjestelmiä käytössä Suomessa.

3 Pedagogiset mahdollisuudet ja haasteet

Tässä osioissa tarkastellaan oppimisanalytiikkaa sen pedagogisten tavoitteiden pohjalta. Ensin jaetaan oppimisanalytiikka sen toimijoiden mukaan. Tämän jälkeen tarkastellaan erilaisia metodeja, joita oppimisanalytiikka hyödyntää tällä hetkellä. Lopuksi tarkastellaan oppimisanalytiikan pedagogisia haasteita.

3.1 Oppimisanalytiikan jako toimijoiden suhteen

Oppimisanalytiikalla tarkoitetaan koulutuksessa syntyvän tiedon keräämistä sekä sen analyysi ja raportointi työtä, jonka tarkoituksena on tuoda analysoitava ilmiö saavutettavammaksi. Pedagogisessa, eli kasvatuksellisessa mielessä tähän ilmiöön liittyy useampia eri tasoisia toimijoita, kuten itse oppijat, opettajat ja opetusta tarjoavat organisaatiot (Opetus- ja kulttuuriministeriö 2021, 26). Tämän kirjallisuuskatsauksen kiinnostus keskittyy oppijoiden ja opettajien näkökulmaan, sillä tältä tasolta on mielekästä tarkastella oppimisanalytiikan pedagogisia ja opetukseen käytettäviä teknillisiä ratkaisuja.

Opetus- ja kulttuuriministeriön (2021) laatimassa viitekehyksessä korostetaan oppijoiden näkökulmasta seuraavia asioita: Oppimisanalytiikan avulla luodut yhteenvedot ja omasta toiminnasta tehdyt visuaalisatiot voivat auttaa oppijaa tarkastelemaan ja kehittämään omaa toimintaa. Lisäksi oppimisanalytiikka voi auttaa opintojen persoonallistamisessa. Analyysin avulla kerätystä tiedosta voidaan arvioida oppijan osaamistasoa ja tätä kautta kohdentaa hänelle sopivia materiaaleja, tehtäviä ja kurssikokonaisuuksia. Tämä voi parhaassa tapauksessa parantaa oppimisen tuloksia ja tätä kautta vähentää opintojen kesken jättämisen riskiä.

Opettajan näkökulmasta oppimisanalytiikan luoma tieto voi tuoda esille oppijoiden kohtamia ongelmakohtia ja opettajan oman opetuksen toimivuutta. Tämä voi edistää opetuksen suunnittelua ja kehitystyötä. Tämä voi myös ohjata opettajaa suuntaamaan ohjausta eniten sitä tarvitseville ja tätä kautta kehittää parempaa ajan käyttöä (Opetus- ja kulttuuriministeriö 2021, 27).

3.2 Oppimisanalytiikan yleisempiä metodeja ja niiden pedagogisia käyttökohteita

Aiemmassa alaluvussa esiteltiin millaista hyötyä oppimisanalytiikka voisi tarjota oppilaan ja opettajan näkökulmasta. Tämä näkökulman asettaminen on tämän tekstin kannalta tärkeää, sillä ilmiötä voisi tarkastella esimerkiksi oppilaitos- ja organisaatiotasolla, mutta tämän kirjallisuuskartoituksen tarkoitus on tarkastella ilmiötä käytännönläheisemmin, jotta kuva oppimisanalytiikan käytännön ilmentymisestä olisi selkeämpi. Seuraavissa kappaleissa avataan tarkemmin oppimisanalytiikan olemassa olevia keinoja/metodeja, joilla kerätää tietoa ja edistää näitä haluttuja tavoitteita. Näitä tietoja voidaan kerätä verkossa olevien opetusjärjestelmien kautta. Yleisimpiä keinoja ovat esimerkiksi luokittelu ja ennustaminen, ryhmittely, oudokin havainnointi, tapahtumien tarkastelu ja datan tislaukset ihmisen tarkasteltavaksi (Romero ja Ventura 2020, 13).

Oppimisessa syntyvän datan luokittelulla ja ennustamisella viitataan tiedonlouhinnan menetelmien joukkoon, jonka tavoitteena on tunnistaa suhde riippumattomien muuttujien (kohde- ja riippuvaisten muuttujien välillä tarkasteltavan datan aikaisempien arvojen perusteella). Luokittelumenetelmät määrittelevät ennalta määrätyn luokkien määrän ja ennustavat, mihin luokkaan kohteet sijoittuvat koulutetun mallin perusteella. Koulutuksessa näitä menetelmiä on sovellettu esimerkiksi opiskelijoiden suoritusarvojen ennustamiseen verkko-oppimisympäristössä heidän käyttäytymisensä perusteella (Ray ja Saeed 2018, 138).

Jos olisi esimerkiksi tilanne, jossa opettaja haluaisi jakaa opiskelijat ryhmiin ”hyvä suoritus”, ”keskinkertainen suoritus” ja ”heikko suoritus” heidän aiempien tehtävien ja kokeiden perusteella, tämä olisi esimerkki luokittelumenetelmästä. Opettaja voisi kehittää mallin, joka perustuu esimerkiksi opiskelijan läsnäoloon, tehtävien palautusaikatauluun ja osallistumiseen keskusteluissa. Tämä malli voisi sitten luokitella opiskelijat eri suoritusluokkiin auttaen opettajaa tunnistamaan ne, jotka saattavat tarvita lisätukea tai haastetta. Toisaalta, jos opettaja haluaisi ennustaa opiskelijoiden tulevia arvosanoja tietämättä tulevia tehtäviä tai kokeita, hän voisi käyttää ennustusmenetelmää. Tämä menetelmä voisi perustua opiskelijan aikaisempiin suorituksiin ja muihin tekijöihin, kuten opiskelijan osallistumisen aktiivisuuteen ja opintojen edistymiseen. Näin opettaja voisi saada arvion siitä, miten opiskelija todennäköi-

sesti suoriutuu tulevissa tehtävissä tai kokeissa.

Ryhmittelyllä (engl. *clustering*) tarkoitetaan ison data-altaan ryhmittelyä pienempiin osiin jonkin tietyn datapisteen avulla. Koulutuksellisessa mielessä tätä metodia on käytetty esimerkiksi ryhmittelemällä samankaltaisia kurssimateriaaleja tai oppilaiden ryhmittelyä heidän oppimis ja kanssakäymis tottumusten kannalta (Romero ja Ventura 2020, 13). Tällainen ryhmittely voi olla hyödyllistä, jotta samankaltaisille opiskelijaryhmille olisi helpompaa antaa yhtenevää ohjausta (Ray ja Saeed 2018, 138). Ryhmittelyä voi pitää opettajan työkaluna. Ray ja Saeed (2018) ottavat julkaisussaan esille, että verkossa olevien oppimisympäristöjen monimuotoisuus tarjoaa hyvän lähtökohdan tälle metodille, sillä näissä ympäristöissä on esimerkiksi helppo seurata opiskelijoiden keskusteluja ja niiden määrää.

Oudokin havainnoilla (engl. *outlier detection*) tarkoitetaan oudokkien eli muusta datasta poikkeavan havainnon löytämistä. Oudokin löytäminen voi kertoa esimerkiksi oppilaan tai ohjaajan yllättävästä suoriutumisen tason vaihtelusta (Ray ja Saeed 2018, 138). Jos oppilasta tai ohjaajaa mitataan pitkään ja suoriutumisesta saadaan tasalaatuisia tuloksia ja dataa, niin nopea poikkeus datassa voi kertoa oppilaan/ohjaajan poikkeuksellisesta heikosta suoriutumisesta. Tämän avulla voidaan myös löytää esimerkiksi erittäin hyvin tai heikosti suoriutuvia oppilaita datan perusteella (Romero ja Ventura 2020, 13). Tämä voi auttaa siirtämään huomiota heidän ohjaamiseen. Oudokkien havainnoinnissa tulee kuitenkin tilastotieteen sääntöjen mukaan huomioida mahdollisuus siihen, että joku testiaineistosta johtumaton syy voi johtaa oudokin ilmestymiseen (esimerkiksi virhe mittauksessa).

Tapahtumien tarkastelu (engl. *process mining*) on metodi joka perustuu verkkoympäristössä tehtyjen tapahtumien tarkasteluun. Ohjaaja voi esimerkiksi tarkastella kuinka paljon vastauksia on annettu mihinkin verkossa olevaan tehtävään ja tämän avulla havainnoida opiskelijan aktiivisuutta tai visualisoida oppilaan menestystä kurssilla (Ray ja Saeed 2018, 139). Tätä voidaan pitää yhtenä keskeisimmistä metodeista teknisen näkökulman pohjalta, sillä tähän perustuu monen teknisen ratkaisun varsinainen toiminta ja tiedon keruu. Esimerkiksi LMS-järjestelmät kuten Moodle keräävät tietoa tällä keinolla.

Datan tislauk (engl. *distillation of data for human judgement*) tarkoittaa metodia, jonka tarkoituksena on ”tislata” käsitelty data ihmiselle ymmärrettävämpään muotoon. Tätä voidaan

tehdä esimerkiksi visuaalisaation tai interaktiivisten ympäristöjen avulla. Datan tislauks on hyvä metodi, sillä se auttaa ohjaajaa tulkitsemaan suuren määrän dataa nopeasti (Ray ja Saeed 2018, 139). Käytännössä tämä metodi perustuu tiedon keräämiseen ja esimerkiksi graafin piirtämiseen sen pohjalta. Kuviossa 2 on graafi, joka on piirretty tarkkailemalla aktiivisten käyttäjien määrää (engl. number of active students) ja heidän aktiivisuuttaan kelloaikaan nähden. Alkuperäisessä kuvassa aikajana on ympäri vuorokautinen.



Kuvio 2. Graafi, joka on piirretty Moodle-opetusjärjestelmään asennetun liitännäisen avulla (Schmitt 2015).

3.3 Oppimisanalytiikan pedagogisia haasteita

Oppimisanalytiikan haasteita voidaan jaotella hallinnollisiin, tutkimuksellisiin sekä opetuksen liittyviin haasteisiin (El Alfyn, Gómez ja Dani 2018, 5). El Alfyn, Gómezin ja Dandin (2018) jaottelussa hallinnollisiksi haasteiksi lasketaan organisaatioiden resurssien saatavuuden ja datankäytön yksityisyyden suojan ongelmat. Oppimisanalytiikan uuden tutkimussuuntauksen nuoresta iästä johtuen ei vielä ole tarpeeksi varsinaista näyttöä oppimisanalytiikan

lytiikan tehokkuudesta ja tästä syntyy jaottelu tutkimukseen liittyviin haasteisiin. Opetuksen haasteet voivat liittyä esimerkiksi juuri resurssien puutteeseen, riittämättömään tutkimustietoon ja käytännön ongelmiin tekniikan hyödyntämisessä.

Ensimmäinen El Alfyn, Gómezin ja Danin (2018) esiin ottama haaste liittyy osittain sekä tutkimukselliseen että pedagogiseen ongelmaan. El Alfyn, Gómezin ja Danin (2018) viittauksessa McNamaran ym. (2014) teokseen otetaan esille, että oppimisanalytiikan tutkimuksessa ei keskitetä tarpeeksi huomiota oppimisstrategioihin ja oppimisen tuotoksien analyysiin. Tällä tarkoitetaan sitä, että pedagogisessa mielessä esimerkiksi kurssitenttien ja kurssiarvosanojen arviointi ei luo kokonaista kuvaa oppimisprosessista eikä myöskään kuvaa oppimisprosessista mahdollisesti muissa opinnäytteissä, kuten blogeissa tai verkossa tapahtuvassa keskustelussa (Gašević ym. 2016, 18). Täten huomiota tulisi siirtää myös muunlaisten muuttujien mittaamiseen analytiikan erilaisia työkaluja (kuten esimerkiksi tekstinlouhinta - työkaluja) hyödyntäen, jotta voitaisiin saada kokonaisvaltaisempi kuva oppimisesta (Gašević ym. 2016, 18).

Toinen El Alfyn, Gómezin ja Danin (2018) esiinottama haaste heijastuu samankaltaisesta ongelmasta – yhtenä oppimisanalytiikan vahvuutena nähdään se, että se tuo automatisoidusti yhteenvetotietoja ihmisille helposti ymmärrettävässä muodossa esimerkiksi datan tislauksen avulla (Ray ja Saeed 2018, 139). El Alfyn, Gómez ja Dani (2018) kuitenkin viittaavat Rogersin (2016) teokseen ja nostavat esille ajatuksen, että nämä yhteenvetotiedot eivät kuitenkaan kaikissa tapauksissa riitä antamaan oikeaa kuvaa oppijoiden oppimisprosesseista ja kehityksestä. Sen sijaan tulisi erottaa ajatus palautteesta, joka kuvailee oppimista ja kognitiota palautteesta, joka kuvailee saavutuksia. Esimerkiksi palautettujen osatehtävien määrä voi kuvata opiskelijan etenemistä kurssilla saavutuksen omaisesti, mutta se ei välttämättä kerro siitä millainen oppimisen käyrä opiskelijalla on ollut.

Tästä herää ajatus siitä, että vaikka oppimisanalytiikka on omaksinut useita erilaisia metodeja tiedon käsittelemiseen, kuten luvussa 3.1 suppeasti listataan, niin metodien sovellettu käyttö ei välttämättä vastaa maalia, johon oppimisanalytiikan avulla ultimaattisesti tähdätään. Alfyn, Gómezin ja Danin (2018) listaamista haasteista voi huomioda, että oppimisanalytiikan käytännön toteuttaminen ei aina hyödynnä eri metodeja siinä määrin, että oppimisesta saataisiin laajempaa ymmärrystä opetuksen kehittämistyön kannalta.

4 Tekniset mahdollisuudet ja haasteet

Tässä luvussa tarkastellaan ensin oppimisanalytiikan teknillisiä ilmentymiä Moodle-järjestelmän analytiikka rajapinnan kautta ja tarkastellaan millaisia rajapintoja/standardeja sovelluksien kehitykselle on yleisesti asetettu. Tämän jälkeen tarkastellaan teknisiin standardeihin ja tietopankkeihin liittyviä haasteita, joiden kautta voidaan tarkastella aiemmin keskusteltuja pedagogiikkaan liittyviä haasteita.

4.1 Tekniset ilmentymät

Kuten tämän tekstin teoriaosiossa todettiin, oppimisanalytiikka on ilmiö, joka koostuu tietotekniikan, tilastotieteen ja kasvatustieteen yhteisestä panoksesta (Romero ja Ventura 2020). Tässä jaottelussa ikään kuin haetaan oppimisanalytiikalle tavoitteita pedagogisessa mielessä. Hyödynnetään tilastotieteen tekniikoita tiedon käsittelyyn ja tämän mahdollistamiseksi käytetään tietotekniikkaa. Tietotekniikka ja teknologia mahdollistavat oppimisanalytiikan siten, että nämä verkossa olevat ympäristöt keräävät käsiteltävän datan tietokantaan, jota siten myöhemmin käsitellään tietoteknisin keinoin tilastotieteen teorian avulla. Näin syntyy tarkoituksenmukaista tietoa, jota voi hyödyntää pedagogisessa mielessä. On myös hyvä huomioda tietotekniikan tärkeä asema siinä mielessä, että perinteisin data-analytiikan keinoin tätä tietoa olisi todella työlästä tai jopa mahdotonta analysoida sen suuren määrän vuoksi. Tässä luvussa käsitellään teknisiä ilmentymiä, kuten yhtä sovellusesimerkkiä ja tietoteknistä perustaa, joiden avulla sovelluksia suunnitellaan.

Tämän teoksen kirjoitushetkellä kirjoittajan oli vaikea arvioida, mitkä oppimisanalytiikan työkalut ovat laajasti otettu käyttöön. Täten tekstissä avataan oppimisanalytiikan työkalujen toimintaa Moodle-järjestelmän analytiikka ohjelmointirajapinnan kautta (Nicols 2022a). Moodle on laajasti opetuksessa hyödynnettävä tietojärjestelmä, joka luo perustaa valinnalle. Engl. *Moodle analytics API* oli alunperin moodleen asennettava laajennus, mutta Moodle version 3.4 jälkeen se on integroitu osaksi Moodle-järjestelmää. Tämän rajapinnan tarkoitus on tarjota ohjaajille työkalu, jonka avulla on mahdollista luoda ennustamisen malleja (engl. *predictive models*) kohteen ja mittarin perusteella – tässä kohde on päämäärä, jota haluaan

mitata ja mittarit ovat muuttujia, joiden uskotaan johtavan oikeaan ennustukseen kohteen kannalta (Nicols 2022a). Jos näiden mallien ennustetarkkuus todetaan tarpeeksi korkeaksi, niin Moodle toteuttaa koneoppimisalgoritmien avulla laskelmia, jotka perustuvat sivuston dataan määritettyihin indikaattoreihin (Nicols 2022b). Täten kun uutta mallin määritelmiin sopivaa dataa on tarjolla, niin Moodle alkaa ennustamaan todennäköisyyttä, jolla kohdeta-
pahtuma tapahtuu. Tätä käyttäjät, kuten opettajat voivat hyödyntää erilaisten tehtävien suorittamiseen, kuten raporttien tai viestien lähettämiseen.

Analytiikka rajapinnan dokumentaatiossa mainitaan, että liitännäisen vakioversiossa on vain malleja, jotka on todettu tarkoiksi ennustajiksi. Vakioversiossa kuitenkin ainoa oppimisanalytiikan työkalu on malli, jonka perusteella voidaan arvioida opiskelijan riskiä lopettaa kurssi kesken. Tämän mallin mittareina voidaan pitää opiskelijan aiempaa suoriutumista muilla kursseilla ja ennustuksen kohteena voidaan pitää arviota siitä pääseekö opiskelija kyseisen kurssin läpi vai ei. Tämän ainoan valmiin mallin lisäksi ohjelmointirajapinta kuitenkin tarjoaa käyttäjälle mahdollisuuden tuoda malleja tai dataa muista järjestelmistä ja kouluttaa näitä malleja Moodle-järjestelmän rajapinnan avulla (Nicols 2022a). Tämän mahdollistaa erilaiset ohjelmistokehityksessä sovitut standardit, joiden tavoitteena on juuri tämän kaltainen toiminnallisuus ja tiedon sujuva siirtäminen eri ohjelmistojen välillä. Seuraavassa kappaleessa avaan tarkemmin oppimisanalytiikan työkaluissa hyödynnettyjä standardeja.

Standardien noudattaminen on ohjelmistokehityksessä yhteisten toimintatapojen laatimista ja yksi yhteentoimivuuden perusta (Opetus- ja kulttuuriministeriö 2021, 31). Tämä korostuu myös oppimisanalytiikan ohjelmistosuunnittelussa, sillä oppimisanalytiikan käytännön hyödyntäminen perustuu usein heterogeeniseen, eri digitaalisista lähteistä kaivettuun tietoon. Opetus- ja kulttuuriministeriö (2021) huomauttaa teoksessaan, että eri toimijoiden epäyh-
tenäiset datan keruun, käsittelyn ja tallentamisen keinot voivat muodostua esteeksi oppimisanalytiikan harjoittamiselle. Täten osa kansainvälisistä standardointitoimijoista (IEEE, ISO/IEC) onkin ottanut oppimisanalytiikan huomioon standardoinnissa. Opetus- ja kulttuuri-
ministeriö (2021) listaa seuraavia esimerkkejä oppimisanalytiikan standardoinnista:

1. ”ISO/IEC TR 20748-1:2016 Referenssimalli: sis vaatimukset IT-systeemeille oppimisanalytiikan yhteentoimivuuteen liittyen sekä oleellisen termistön, käyttäjien vaatimukset ja järjestelmien arkkitehtuurin”

2. ”ISO/IEC TR 20748-2:2017 Järjestelmien vaatimukset: sis. vaatimukset oppimisanalytiikkaan liittyville järjestelmille niiden välisen kommunikaation ja toimintojen tehokkuuden optimoimiseksi”
3. ”ISO/IEC TR 20748-3 Ohjeistus datan yhteentoimivuudelle”
4. ”ISO/IEC TR 20748-4 Yksityisyys ja datan suojaaminen”

Tästä listauksesta näkee, että näillä standardoinneilla on jo koitettu selvittää isoja oppimisanalytiikan haasteita, kuten käyttäjän yksityisyydensuojaan liittyviä haasteita, sovellusten välisen kommunikaation tehostamisen ja datan yhteentoimivuuden haasteita.

4.2 Teknillisiä haasteita

2010-luvun alkupuolella oppimisanalytiikan teknillisenä haasteena on nähty aiemmin mainitut haasteet erilaisten oppimisjärjestelmien tavasta kerätä, analysoida ja tallentaa tietoa (Verbert ym. 2012). Näitä ongelmia onkin koitettu ratkoa 2010-luvun loppupuolella aiemmin mainitulla ISO/IEC-standardoinnilla, jotta eri tietojärjestelmien välillä olisi helpompi siirtää tietoa ja metodeja. Tämän teoksen tekohetkellä kirjoittajalle oli kuitenkin epäselvää, millaisella tasolla standardoinnin hyödyntäminen on nykyhetken sovelluksissa ja käytännöissä. Wolf ym. (2018) ovat teoksessaan tutkinut oppimisanalytiikan hyödyntämisen tasoa Yhdysvalloissa ja listanneet myös kehitysehdotuksia (s. 209). He huomasivat, että osavaltio tasolla on panostettu näiden sovelluksien kehitykseen ja käyttöönottoon tulevaisuus mielessä pitäen, mutta kuitenkin oppimisanalytiikan päivittäinen hyödyntäminen ei ole todellisuutta useimmille opettajille ja opiskelijoille. Tämä johtuu heidän mukaansa osittain siitä, että opetuksen kulttuuri ja tavat eivät ole vielä sopeutuneet oppimisanalytiikan käyttöönottoon, mutta myös siitä, ettei teknillinen yhtenäisyys toteudu tyydyttävällä tasolla. Näistä ongelmista on lisää tietoa luvussa 5, jossa ongelmia tarkkaillaan käyttäjän näkökulmasta.

Standardien lisäksi toinen vakava teknillinen haaste liittyy oppimisanalytiikassa käytettyihin datapankkeihin, niiden turvallisuuteen ja toimijoiden yksityisyydensuojaan (Wang 2016, 383). Wang (2016) nostaa esille tekstissään, että suurien koulutuksessa kasvavien tietomäärien käsittelyssä on tärkeää huomioida se, kuinka niitä varastoidaan oikealla tavalla. Tällä hetkellä tietoa tallentuu erilaisiin paikkoihin, kuten koulujen toimistoihin, verkkopohjaisiin

järjestelmiin (esimerkiksi LMS) ja erillisiin laitteisiin. Usein nämä laitteet tai tietokannat eivät ole yhteydessä toisiinsa, mikä on hyvä siinä mielessä että paikallinen tietomurto ei johda koko tietomäärän paljastumiseen (Wang 2016, 383). Tässä on kuitenkin oppimisanalytiikan kannalta se huono puoli, että tietokannat on erityisen tärkeitä metodien teknisen toteutuksen kannalta, koska näistä tietokannoista löytyy hyödynnettävät mallit ja verrattava tieto, jonka pohjalta voidaan suorittaa vertailuja ja laskuja uuden tiedon luomiseksi.

Tämä tilanne jossa oppimisanalytiikan työkaluille ja tietokannoille ei ole vakiintuneita käytäntöjä ja standardeja voi selittää sitä, miksi aiemmin luvussa 3.3 keskustellut pedagogiset haasteet ovat vielä yleisiä. Se että eri analytiikan metodeita käytetään vielä puutteellisesti tai mitattavat asiat eivät kerro totuutta kognitiosta voi selittyä sillä, että yhteisten käytäntöjen puute voi hiertää oppimisanalytiikan hyödyntämistä arjessa.

5 Teknologian, analyysin ja pedagogiikan kohtaaminen vaatii käyttäjien huomioimista ja suunnitelmallisuutta

Aiemmissa kappaleissa on käyty läpi oppimisanalytiikan ilmentymistä teknillisessä ja pedagogisessa mielessä. On tarkasteltu metodeja, toimijoita, teknologiaa ja näiden eri näkökulmien haasteita. On huomattu, että vaikka oppimisanalytiikka on ollut tutkimuksien kohteena, niin ilmiön nuori ikä, tutkimuksen monimuodottomuus ja vakiintumattomat käytännöt luovat suuren osan näistä ongelmista. Tässä luvussa käydään vielä läpi ihmisen roolia oppimisanalytiikan hyödyntäjänä.

Klein ja Hess (2018) ottavat teoksessaan esille ajatuksen siitä, että vaikka oppimisanalytiikan hyödyntäminen arviointiin on monen oppilaitoksen mielenkiinnon kohteena, niin organisaatiotason haasteet ja toimijoiden henkilökohtaiset käyttöönoton haasteet rajoittavat sen hyödyntämistä. Avataan tätä ajatusta tutkimalla Kleinin ja Hessin huomioita aiheesta. Ensimmäinen huomio liittyy inklusion ja luoton tärkeyteen oppimisanalytiikkaa käyttöönotossa. Teoksessa otetaan esille tapaus, jossa oppimisanalytiikan avulla oppilaille lähetetty varoitus liian heikosta suoriutumisesta kurssilla sai osan oppilaista lopettamaan kurssin sen sijaan, että olisi tehnyt asian korjaamisen eteen työtä. Klein ja Hess toteavat, että tähän tilanteeseen on saattanut johtaa se, että tämänkaltainen viesti on voinut vaikuttaa oppilaiden mielestä yllättävältä, eli oppilaat eivät saata luottaa heistä kerätyn tiedon läpinäkymättömyyteen ja yksityisyydensuojaan, josta voi syntyä epäluottamusta systeemiä kohtaan. Tämän esimerkin pohjalta he nostavat esille inklusion merkityksen oppimisanalytiikan käyttöönotossa (Klein ja Hess 2018, 150). Tällä inklusiolla tarkoitetaan ihanteellista tilannetta, jossa oppimisanalytiikan kehityksessä, hankinnassa ja toteuttamisessa otetaan mukaan sitä käyttävä kohderyhmä/toimijat, kuten oppilaat ja opettajat.

Wise ja Vytasek (2017) tukevat ajatusta ilmaisemalla huolensa oppimisanalytiikan tilanteesta ”tarkoituksellinen toteutuksen suunnittelu on olennainen, ei valinnainen vaihe, oppimisanalytiikan käyttöönotossa. Jos haluamme välttää liian monien lupaavien teknologioiden kohdalon, jotka eivät koskaan vaikuttaneet todellisesti koulutukseen, tutkimus ihmisen ja teknologisten elementtien vuorovaikutuksesta, jotka vaikuttavat analytiikan käyttöön, on kriit-

minen huomion kohde alalla jatkossa”. He ottavat kantaa siihen, että tulevaisuudessa tulee kiinnittää enemmän huomiota suunnitelmallisuuteen, jos oppimisanalytiikkaa halutaan hyödyntää merkityksellisesti. Klein ja Hess (2018) toteavat vielä, että kun erilaisia työkaluja otetaan käyttöön, niin tämänkaltainen inkluusio lisää käyttäjien luottoa sekä työkaluun, että toimijoiden yhteistyöhön jatkossa. On hyvä huomioida se, että suhteellisen uutena työkaluna oppimisanalytiikka voi olla kohderyhmälle vieras käsite ja adaptaatio tämänkaltaiseen teknologiaan voi vaatia totuttelua.

Toinen Kleinin ja Hessin huomio liittyy oppimisanalytiikan onnistumisen riippuvuudesta kontekstiin, tulkintaan ja viestimiseen. He toteavat, että luoton ja inkluusion lisäksi onnistumisen kannalta on tärkeää opettajien ja ohjaajien huomion siirtäminen näihin osa alueisiin. Jotta opettajat voisivat adaptoida onnistuneesti näitä tekniikoita opetukseen ja arviointiin, niin heidän tulisi ymmärtää kontekstista se, mitä kerätty data heille kertoo ja kuinka dataa tulkitaan. Heillä tulisi myös olla käsitys siitä, mitkä muuttujat syötetään analyysialgoritmeille ja millaisia tuloksia tämän pohjalta generoidaan. Lisäksi tulisi tietää, että millaisia syitä näille tuloksille on ja kuinka oppimisanalytiikka ylipäänsä istuu heidän opetukseensa (Klein ja Hess 2018, 151). Tämän kaltainen kontekstin ja syvän ymmärryksen luominen on tärkeää tekniikan lähestyttävyyden ja syvemmän ymmärryksen kannalta.

Klein ja Hess (2018) huomauttavat myös, että ohjaajien tulisi osata opastaa oppilaille riittävällä tasolla kuinka lukea oppimisanalytiikan luomaa tietoa – jos oppilaille tarjotaan esimerkiksi kuvaajia heidän suoriutumisestaan, mutta näitä ei selitetä tarpeeksi, niin oppilaalle voi jäädä hyvin pintapuolinen kuva siitä, miten kuvaajat korreloivat heidän arvosanojen tai oppimisen kanssa. Klein ja Hess huomauttavat, että usein tätä kontekstualisoimista ja näiden visualisatioiden auki selittämistä tapahtuu hyvin vähän tai joskus ei ollenkaan, mikä voi tehdä kyseisestä tekniikasta täysin turhaa. Kohderyhmän tulisi saavuttaa tämä tieto ja hyötyä siitä, joten riittävä kommunikaatio ja koulutus laitoksen eri toimijoiden välillä on erittäin tärkeää, jos oppimisanalytiikasta halutaan hyötyä. Uutena teknologiana oppimisanalytiikan metodien ymmärtäminen on tärkeää myös siksi, että sitä hyödyntävät toimijat pystyisivät arvioimaan tuotettua tietoa kriittisesti ja tarkastella täten sen tarkoituksenmukaisuutta ja mahdollisia vikoja.

6 Johtopäätökset

Tämän kirjallisuuskartoituksen tarkoituksena oli analysoida kirjallisuutta oppimisanalytiikan ilmentymistä ja haasteista. Tarkempaan tutkimuskysymyksenä tarkasteltiin oppimisanalytiikan pedagogisia ja teknillisiä näkökulmia. Teoksessa tuotiin esille metodeja, joilla oppimisanalytiikka pyrkii tehostamaan opetusta oppilaan ja opettajan näkökulmasta. Teknillistä ilmentymää pohjustettiin Moodle ohjelmistorajapinnan kautta ja oppimisanalytiikan teknisen puolen ohjelmistokehityksen kysymyksiä esiteltiin standardoinnin kautta.

Katsauksessa esiin nousseista haasteista yksi merkittävä oli oppimisanalytiikan suoriutumisen heikkous sen potentiaaliin nähden. Vaikka oppimisanalytiikka on omaksunut erilaisia tekniikoita ja metodeita, niiden tehokas hyödyntäminen ei aina toteudu. Tämä voi osittain johtua oppimisanalytiikan suhteellisen nuoresta iästä. Vaikka ilmiöstä on keskusteltu jo 90-luvulla, nopea teknologinen kehitys globaalilla tasolla on johtanut jatkuvaan teknologisten ratkaisujen kehitykseen ilman vakiintuneita standardeja. Nuori ikä voi myös poikia ongelmia siksi, että vaikka oppimisanalytiikka on ollut tutkimusten mielenkiinnon kohteena, niin tutkimukset eivät ole tarpeeksi monimuotoisia ja varsinkin empiiriset tutkimukset olisivat hyödyllisiä kehityksen kannalta.

Lisäksi oppimisanalytiikan käyttöönotossa on olennaista huomioida varsinaiset käyttäjät ja varmistaa suunnitelmallisuus. Koska oppimisanalytiikka on vielä nuori ilmiö, sen integroiminen opetuskäytäntöihin edellyttää erityistä huomioimista käyttäjäryhmän näkökulmasta, kuten inklusion ja luottamuksen tunteen vahvistamisen kautta. Menestyksellinen käyttöönotto vaatii harkittua suunnitelmaa, jota ihanteellisesti hiotaan eri toimijoiden kanssa yhdessä.

Tulevaa tutkimusta ajatellen voisi olla hyvä siirtää huomiota empiiriseen tutkimukseen. Kirjallisuudessa on keskusteltu oppimisanalytiikan mahdollisuuksista ja keinoista, mutta olisi arvokasta saada laajempaa tutkimustietoa oppimisanalytiikan arjen ilmentymistä ja mahdollisista epäonnistumisista tai onnistumisista. Tämä voisi auttaa tutkimaan sitä, mikä oppimisanalytiikassa toimii ja mikä ei, jolloin saataisiin tietoa jatkuvan kehityksen kannalta.

Tätä kirjallisuuskartoitusta tehdessä ei löytynyt paljoa Suomessa tuotettua tutkimustietoa ja kartoitustyössä keskityttiin lähinnä ulkomaisiin teoksiin. Suomessa toteutettu tutkimustieto

voisi avata oppimisanalytiikan alueellisia ja kulttuurillisia eroja verrattuna ulkomaisiin teoksiin.

Lähteet

- Auvinen, Ari-Matti. 2017. “Oppimisanalytiikka tulee - oletko valmis?” Artikkelin on eoppimiskeskuksen julkaisu. Viitattu 8.10.2023. <https://poluttamo.fi/2017/08/02/oppimisanalytiikka-tulee-oletko-valmis/>.
- Daniel, Ben Kei. 2019. “Big Data and data science: A critical review of issues for educational research”. *British Journal of Educational Technology* 50 (1): 101–113. <https://doi.org/10.1111/bjet.12595>. <https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12595>.
- El Alfy, Shahira, Jorge Gómez ja Anita Dani. 2018. “Exploring the benefits and challenges of learning analytics in higher education institutions: A systematic literature review”. *Information Discovery and Delivery* 47. <https://doi.org/10.1108/IDD-06-2018-0018>.
- Gašević, Dragan, Shane Dawson, Tim Rogers ja Danijela Gasevic. 2016. “Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success”. *Internet High. Educ.* 28:68–84. <https://doi.org/10.1016/j.iheduc.2015.10.002>. <https://api.semanticscholar.org/CorpusID:4661661>.
- Klein, Carrie ja Richard M. Hess. 2018. “Using Learning Analytics to Improve Student Learning Outcomes Assessment”. *Learning Analytics in Higher Education*, 140–159. <https://doi.org/10.4324/9780203731864>.
- Lang, Charles, Alyssa Wise, George Siemens ja Dragan Gasevic. 2017. *Handbook of Learning Analytics*. <https://doi.org/10.18608/hla17>.
- McNamara, Danielle S., Arthur C. Graesser, Philip M. McCarthy ja Zhiqiang Cai. 2014. *Automated Evaluation of Text and Discourse with Coh-Matrix*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511894664>.
- Nicols, Andrew. 2022a. “Moodle Analytics Api”. Viitattu 27.11.2023, kesäkuu. <https://moodledev.io/docs/apis/subsystems/analytics>.
- . 2022b. “Moodle machine learning backend”. Viitattu 27.11.2023, toukokuu. <https://moodledev.io/docs/apis/plugin/types/mlbackend>.

Opetus- ja kulttuuriministeriö, Oppimisanalytiikkajaosto, toimittanut. 2021. *Oppimisanalytiikan viitekehys : Hyvät käytännöt oppimisanalytiikan käyttöönotossa ja hyödyntämisessä*. Opetus- ja kulttuuriministeriön julkaisuja. Opetus- ja kulttuuriministeriö. <http://urn.fi/URN:ISBN:978-952-263-842-7>.

Ray, Santosh ja Mohammed Saeed. 2018. “Applications of educational data mining and learning analytics tools in handling big data in higher education”, 135–160. https://doi.org/10.1007/978-3-319-76472-6_7.

Rogers, Tim. 2016. “Learning analytics and the imperative for theory-driven research”. Teoksessa *The Sage Handbook of E-learning Research, 2nd edition*, 232–250. ISBN: 978-1-4739-0232-9.

Romero, Cristobal ja Sebastian Ventura. 2013. “Data mining in education”. *WIREs Data Mining and Knowledge Discovery* 3 (1): 12–27. <https://doi.org/10.1002/widm.1075>.

———. 2020. “Educational data mining and learning analytics: An updated survey”. *WIREs Data Mining and Knowledge Discovery* 10 (3). <https://doi.org/10.1002/widm.1355>.

Schmitt, Marcelo. 2015. “Analytics graphs”. Viitattu 12. marraskuuta 2023. https://moodle.org/plugins/block_analytics_graphs.

Wang, Yinying. 2016. “Big Opportunities and Big Concerns of Big Data in Education”. *TechTrends* 60 (4): 381–384. <https://doi.org/10.1007/s11528-016-0072-1>. <https://www.proquest.com/scholarly-journals/big-opportunities-concerns-data-education/docview/1793530015/se-2>.

Verbert, K, N Manouselis, H Drachsler ja E Duval. 2012. “Dataset-driven research to support learning and knowledge analytics”. *Journal of Educational Technology Society* 15 (3): 133–148. ISSN: 1176-3647. <https://pure.tue.nl/ws/portalfiles/portal/3424470/2272055620921.pdf>.

Wise, Alyssa Friend ja Jovita Vytasek. 2017. “Learning Analytics Implementation Design”. Teoksessa *Handbook of Learning Analytics*, 151–160. <https://doi.org/10.18608/hla17>.

Wolf, Mary Ann, Rachel Jones, Sara Hall ja Governor Bob Wise. 2018. *Policies and Capacity: Enablers and Barriers for Learning Analytics*, 175–214. ISBN: 9781641133692. <https://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=1852538&site=ehost-live>.