

Mikko Toitturi

**BIG DATAN HYÖDYT MODERNEISSA VERKKO-
OPPIMISYMPÄRISTÖISSÄ**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2018

TIIVISTELMÄ

Toitturi, Mikko

Big datan hyödyt moderneissa verkko-oppimisympäristöissä

Jyväskylä: Jyväskylän yliopisto, 2018, 34 s.

Tietojärjestelmätiede, kandidaatintutkielma

Ohjaaja(t): Clements, Kati

Opiskelu verkko-oppimisympäristöissä on lisääntynyt yhä enemmän viime vuosina. Samaan aikaan opiskelumuodon murros on esittänyt uusia haasteita oppilaitoksille: kuinka integroida tehokkaasti tätä uutta teknologiaa osaksi opettamista? Eräänä ratkaisuna ongelmaan on esitelty oppimisen analytiikka, jossa verkko-oppimisympäristöt käyttävät big dataa osana oppimisen ja opettamisen tehostamista. Tutkielmassa esitellään eri tapoja hyödyntää big dataa osana opetusta ja oppimista, arvioidaan niiden hyötyjä ja verrataan niitä perinteisen verkko-oppimisen esittämiin haasteisiin. Tuloksina ilmaantui kolme päätapaa käyttää big dataa osana verkko-oppimisympäristöjä: datan visualisaatio ja ennusteet, suositusjärjestelmät sekä oppimisympäristöjen personointi. Näiden päähyötyinä esiteltiin mm. paranneltu päätöksenteko sekä oppimisen ja opettamisen tehostaminen kerätyn datan pohjalta. Perinteisiin verkko-oppimisen haasteisiin se vastasi mm. monipuolisemmalla ympäristön käytöllä ja jatkuvalla ympäristön kehityksellä. Tutkielma toteutettiin kirjallisuuskatsauksena.

Asiasanat: Oppimisen analytiikka, big data, verkko-oppimisympäristöt, verkko-oppiminen

ABSTRACT

Toitturi, Mikko

The benefits of big data in modern e-learning environments

Jyväskylä: University of Jyväskylä, 2018, 34 p.

Information Systems, Bachelor's Thesis

Supervisor(s): Clements, Kati

Nowadays, education is shifting more and more into different e-learning environments. This poses a new challenge to the learning institutions: how to effectively integrate these new technologies into teaching? As one solution to the problem learning analytics is presented. In learning analytics, e-learning environments utilize big data to reinforce both teaching and learning. In the thesis, different ways of utilizing big data in e-learning environments are presented and their benefits are assessed and compared to the traditional challenges of e-learning. As result, three ways to utilize big data were discovered: data visualization and prediction, recommendation systems and personalized learning environments. The main benefits of these technologies were more effective decision-making, learning and teaching based on the data collected. To the traditional challenges of e-learning learning analytics offered i.a. more versatile usage of the environment and the continuous development of the environment. The thesis was conducted as a literature review.

Keywords: Learning analytics, big data, learning environments, e-learning

KUVIOT

KUVIO 1: Big datan kolme ominaisuutta	16
KUVIO 2: Oppimisen analytiikka prosessina (Chatti ym., 2013)	19
KUVIO 3: Wakoopa-ympäristön käyttöliittymä (Duval, 2011)	24
KUVIO 4: Big datan päämahdollisuudet koulutuksen kolmelle toimijalle (Daniel, 2015) (muokattu).....	27

TAULUKOT

TAULUKKO 1: Eri verkko-oppimisympäristöjen ominaisuudet (Udupi, Sharma & Jha, 2016)	10
TAULUKKO 2: Perinteisen oppimisen ja verkko-oppimisen edut ja haitat (Zhang ym., 2004)	13

SISÄLLYS

TIIVISTELMÄ	2
ABSTRACT	3
KUVIOT	4
TAULUKOT	4
SISÄLLYS.....	5
1 JOHDANTO.....	6
2 VERKKO-OPPIMINEN & VERKKO-OPPIMISYMPÄRISTÖT.....	8
2.1 Verkko-oppimisen määritelmä	8
2.2 Erilaiset verkko-oppimisympäristöt	9
2.3 Verkko-oppimisen edut ja haitat	11
2.3.1 Edut	12
2.3.2 Haitat.....	13
3 BIG DATA	15
3.1 Big datan määritelmä	15
3.1.1 Määrä	16
3.1.2 Nopeus.....	17
3.1.3 Monipuolisuus.....	17
3.2 Big datan hyödyntäminen osana oppimista	18
3.3 Big datan nykyisiä käyttökohteita.....	20
4 BIG DATA OSANA VERKKO-OPPIMISYMPÄRISTÖJÄ	22
4.1 Big datan hyödyntämistavat verkko-oppimisympäristöissä	22
4.1.1 Datan visualisaatio ja ennusteet	22
4.1.2 Suositusjärjestelmät.....	24
4.1.3 Personointi.....	25
4.2 Hyödyt.....	26
5 YHTEENVETO	30
LÄHTEET	31

1 Johdanto

Oppiminen kehittyy jatkuvasti: jo internetiä hyödyntävä tietokoneavusteinen opetus on muuttanut radikaalisesti perinteistä opetusta (Alonso, López, Manrique & Viñes, 2005), mutta samaan aikaan nykyinen tietoyhteiskunta on osoittanut yhä kasvavaa kysyntää uusille, innovatiivisille tavoille luoda koulutusta (Zhang, Zhao, Zhou & Nunamaker, 2004). Verkko-oppiminen on osa tätä uutta dynamiikkaa, joka kuvaa oppijärjestelmiä 2000-luvun alussa (Sangra, Vlachopoulos & Cabrera, 2012). Vaikka näiden innovaatioiden tuloksena verkko-opiskelu on antanut opiskelijoille uusia mahdollisuuksia mm. opetuksen joustavuuden ja muiden teknologian mahdollistamien etujen ansiosta, samalla sen kehitys on muodostanut uusia haasteita oppilaitoksille: kuinka integroida tätä uutta teknologiaa osaksi sekä organisaatiota että itse opetusta? (Sangra ym., 2012). Vastaukseksi on muodostunut useita erityyppisiä verkko-oppimisympäristöjä, joista uusimpien joukossa on syntynyt ympäristöjä, jotka pystyvät käyttämään big dataa toiminnoissaan.

Verkko-oppiminen on tuonut paljon etuja perinteiseen oppimiseen: kehitys internetin, langattoman teknologian, datansiirron ja tietoverkkojen parissa on varmistanut verkkomateriaalin tehokkaan jakamisen käyttäjien välillä maantieteellisistä rajoista huolimatta (Udupi, Sharma & Jha, 2016). Samalla älylaitteiden yleistyessä oppilaat ovat alkaneet käyttää älypuhelimiaan päästäkseen käsiksi eri oppimateriaaleihin. Tämä tuottaa valtavia määriä dataa, jota ei hyödynnetä mitenkään, sillä perinteiset analyysimetodit eivät pysty prosessoimaan näitä tietoja. (Sin & Muthu, 2015.) Tämä data tunnetaan big datana, ja tutkielmassa otetaan selvää, miten sitä voidaan hyödyntää osana verkko-oppimisympäristöjä. Varsinkin verkko-oppimisessa tietokoneen avulla tehdyt interaktiot tarjoavat paljon dataa esim. siitä, miten kauan sisältöä on luettu oppimisympäristössä (Elias, 2011). Tällaisen oppimisympäristöistä kerätyn datan analysointi onkin keskiössä tekniikoissa, jotka käyttävät big dataa oppimisen tehostamiseksi tutkimusalueella, jota kutsutaan oppimisen analytiikaksi (Fournier, Kop & Sitlia, 2011).

Vaikka toisaalta mm. yrityssektori on jo ottanut huomioon tällaisen big datan hyödyntämismahdollisuudet organisaation tuottavuuden ja

kilpailukyvyen tehostamiseksi, ovat oppilaitokset hyödyntäneet tätä dataa vain vähän parantakseen opetusta ja oppimista (Siemens, 2013), ja yleensä ottaen nykyiset oppimisympäristöjen sisäänrakennetut oppilaiden seurantaominaisuudet ovat kaukana tyydyttävästä (Hijon & Carlos, 2006, Elias, 2011, mukaan). Tämän johdosta ne eivät tuota tarpeeksi tietoa oppimisesta, mitä opettajat voisivat käyttää tehokkaasti räätälöimään opetussuunnitelmaa oppilaiden tarpeiden mukaiseksi (Zhang ym., 2007, Elias, 2011, mukaan). Mahdollisuuksia siihen kuitenkin olisi: ne tietokoneet ja erilaiset älylaitteet, joilla ympäristöjä käytetään, mahdollistavat hyvin monipuolisen datan keräämisen (Sin & Muthu, 2016), ja varsinkin ylemmillä oppilaitoksilla on jo valmiiksi suuria datavarastoja, joita etäkurssit kartuttavat edelleen. Tätä dataa analysoimalla on potentiaalia mm. luoda ennusteita tulevista haasteista sekä oppilaitokselle että oppilaille (Elias, 2011.), ja tutkimuksen tavoitteena on ottaa selvää näistä eri tavoista hyödyntää big dataa osana verkko-oppimisympäristöjä sekä kartoittaa sen mukanaan tuomia hyötyjä. Tutkimuskysymys on:

- *"Mitä hyötyjä big datan käyttö osana verkko-oppimisympäristöjä voi tuottaa modernien verkko-oppimisympäristöjen käyttäjille?"*

Tutkimus on toteutettu kirjallisuuskatsauksena. Tutkimusmateriaali kerättiin käyttämällä mm. hakusanoja *"big data"*, *"learning analytics"*, *"e-learning"*, *"educational data mining"* ja *"learning environment"* palveluissa Google Scholar sekä Jykdok.

2 VERKKO-OPPIMINEN & VERKKO-OPPIMISYMPÄRISTÖT

Viime vuosikymmeninä IT:n käyttö opetuksellisiin tarkoituksiin on kasvanut, ja tietoverkkoteknologioiden leviäminen on aiheuttanut näiden verkko-oppimiskäytäntöjen merkittävän kehityksen (Kahiigi ym., 2008). Samaan aikaan verkko-oppiminen ja yleinen oppimisen avoimuus on saanut yhä kasvavaa huomiota osana paitsi koulutusta myös liiketoimintaa (Stoffragen, Pawlowski & Pirkkalainen, 2015.) Siksi olisi hyödyllistä ymmärtää, mitä verkko-oppimisympäristöt ovat ja mikä niistä tekee tehokkaan välineen oppimiselle. Tässä luvussa määrittellään sekä verkko-oppimisen että verkko-oppimisympäristöjen käsitteitä ja tutkitaan niiden etuja ja haittoja perinteiseen oppimiseen verrattuna.

2.1 Verkko-oppimisen määritelmä

Verkko-oppimisesta ja verkko-oppimisympäristöistä puhuttaessa epäselvyyttä voi aiheuttaa se, mikä käytännössä tekee oppimisesta verkko-oppimista, ja mitkä ominaisuudet puolestaan määrittävät tiettyä verkko-oppimisympäristöä. Yleisesti ottaen verkko-oppimiseksi mielletään kaikki verkon välitse tapahtuva oppiminen, ja verkko-oppimisympäristö on täten se ympäristö - esimerkiksi verkkosivu - jolla kyseinen oppiminen tapahtuu. Kuitenkin verkko-oppimisen tyhjentävä määrittely on vaikeaa, sillä sen käsite muuttuu kaiken aikaa, ja valtaosaa tiedeyhteisöstä on vaikea vakuuttaa yhdellä määrittelyllä (Sangra ym., 2012). Epäselvyyttä voi aiheuttaa myös se, kuinka verkko-oppiminen usein sekoitetaan aivan erillisiin käsitteisiin, kuten tietokone- tai teknologiapohjaiseen oppimiseen ja harjoitteluun, tai sitten käsitteen luullaan tarkoittavan vain esimerkiksi internetissä suoritettavia kursseja. Nämä voivat toisaalta olla osa verkko-oppimista, mutteivät yksin määrittele sitä tarpeeksi tarkasti (Sangra ym., 2012.) Myös verkko-oppimisympäristöt voivat erota suuresti ominaisuuksiltaan,

joten myös niille on kehitelty omia määrittelyjä. Näitä määrittelyjä käydään läpi seuraavassa luvussa.

Keskustelu verkko-oppimisen määrittelystä ja käytänteistä keskittyy koulutuksen, opetuksen ja oppimisen risteykseen ICT:tä käyttäen. (Friesen, 2009). Sangra ym. (2012) ovat tutkimuksessaan keränneet eri julkaisuista määrittelyjä verkko-oppimiselle, ja päätyneet neljään eri luokitteluryhmään, joiden mukaan verkko-oppimista määritellään: teknologialähtöiseen, alustalähtöiseen, kommunikaatiopainotteiseen sekä oppimispainotteiseen määrittelyyn. Tutkimuksessaan Sangra ym. (2012) yhdistävät nämä eriävät näkökulmat yhdeksi verkko-oppimisen määrittelyksi, joka on:

”Verkko-oppiminen on lähestymistapa opettamiseen ja oppimiseen, joka edustaa kaikkea tai osaa käytettävästä opintomallista, ja joka perustuu sähköisen median ja laitteiden käyttöön työvälineinä parantamaan pääsyä harjoituksiin, kommunikaatioon ja vuorovaikutukseen joka mahdollistaa omaksua uusia tapoja ymmärtää ja kehittää oppimista”. (Sangra ym., 2012).

Määritelmänä se on hyvin kattava, ja ottaa kyseiset eri näkökulmat hyvin huomioon. Keskeistä verkko-oppimiselle ei siis ole vain itse väline, jolla oppiminen tapahtuu, tai edes itse oppiminen, vaan näiden optimaalinen yhdistäminen sellaiseksi kokonaisuudeksi, joka paitsi parantaa oppimista käytettävien välinein, myös mahdollistaa opetuksen kehittämisen edelleen.

2.2 Erilaiset verkko-oppimisympäristöt

Verkko-oppimisympäristöistä puhuttaessa on tärkeä huomata, miten verkko-oppimisympäristöt eroavat toisistaan: koska verkko-oppimisympäristöt voivat sisältää hyvin erityyppistä sisältöä, ja koska myös niiden tavoitteet voivat erota toisistaan, myös niiden tarkempi luokittelu on tarpeen. Kuitenkin on paljon epävarmuutta siitä, miten määritellä eri verkko-oppimisympäristöjä, ja esim. sellaisia termejä kuin oppimisen hallintajärjestelmä, kurssinhallintajärjestelmä sekä virtuaalinen oppimisympäristö käsitellään usein synonyymeinä, vaikka jotkut käsittävät nämä termit eri tavalla (Moore, Dickson-Deane & Galyen, 2011). Johnsonin, Hornikin ja Salasin (2008) mukaan verkko-oppimisympäristöt ovat ympäristöjä, jotka tarjoavat oppimateriaalia verkossa, jossa kurssin interaktiot, kommunikaatio ja suoritus onnistuu teknologian välityksellä. Kuitenkin määritelmässään Johnson ym. (2008) kuvaavat ympäristössä suoritettavia opintoja kursseina, kun taas muissa määritelmässä opintoja määritellään mm. ohjelmina ja oppimistavoitteina. Tämä aiheuttaa ongelmia mm. käsiteltävien opintojen laajuudessa, sillä nämä käsitteet määritellään eri tavoin. (Moore ym., 2011.)

Tutkielmassa keskitytään näiden ympäristöjen teknisiin ominaisuuksiin, sillä päätavoitteena on määrittää eri verkko-oppimisympäristöjen erot verkko-oppimisympäristöön, joka käyttää big dataa. Apuna tässä toimii Udupin,

Sharman ja Jhan (2016) tutkimuksessaan tekemä taulukko (taulukko 1), jossa määritellään keskeisimpiä ominaisuuksia, jotka määrittävät erityyppisiä verkko-oppimisympäristöjä. Tässä vertailussa nostettiin esille big dataa hyödyntävän verkko-oppimisympäristön lisäksi kaksi muuta verkko-oppimisympäristöä: e-portaali sekä oppimisen hallintajärjestelmä.

E-portaali on käsiteltävistä verkko-oppimisympäristöistä ominaisuuksiltaan suppein, mikä ilmenee taulukosta 1. Smith (2004) määrittelee sen infrastruktuuriksi, joka tarjoaa turvallisen, kustomoitavan, integroidun pääsyn sisältöön useasta eri lähteestä, useassa eri lähdemuodossa. Verkko-oppimisympäristönä sen pääominaisuus on siis luotettavan, muunnettavan oppimateriaaliin pääsyn turvaaminen.

Oppimisen hallintajärjestelmät taas ovat jo ominaisuuksiltaan monipuolisempia: ne tarjoavat sisällön lisäksi myös eri tapoja valvoa ja hallita sekä materiaalia että yksittäisiä kursseja, ja se lisäksi voi pitää kirjaa oppilaiden suorituksista ja arvioida niitä ennalta määrättyjen ohjeiden mukaan. Eräs ympäristön pääpiirre onkin sen systeeminen luonne: oppimisen hallintajärjestelmä on viitekehys, joka hallitsee oppimisen prosessin kaikkia osia – se paitsi esittää ja hallinnoi oppimateriaalia, myös osaa arvioida oppimisen tuloksia suoritusten arvioinnin kautta (Forouzes & Darvish, 2012). Moodle on yksi esimerkki oppimisen hallintajärjestelmästä, josta on tullut erittäin suosittu työkalu kurssien verkkosivujen luontiin eri oppilaitoksissa. (Cigdem & Topcu, 2015).

Ominaisuudet	E-portaali	Oppimisen hallintajärjestelmä	Datan louhimisen & big datan integraatio
E-sisällön näyttö	✓	✓	✓
Interaktiivinen oppiminen	✓	✓	✓
Suoritusten valvonta	✗	✓	✓
Kurssien ylläpito	✗	✓	✓
Kurssien hallinta	✗	✓	✓
Suoritusten arviointi	✗	✓	✓
Suoritusten seuranta	✗	✓	✓
Ryvästen luonti	✗	✗	✓
Mallien luonti	✗	✗	✓
Mallien analyysi	✗	✗	✓
Ennustava analyysi	✗	✗	✓
Tiedon löytäminen	✗	✗	✓

TAULUKKO 1: Eri verkko-oppimisympäristöjen ominaisuudet (Udupi, Sharma & Jha, 2016)

Viimeisenä esimerkkinä verkko-oppimisympäristöstä on ympäristö, johon on integroitu big datan keräämistä ja hyödyntämistä. Tämä ympäristö eroaa muista siten, että se oppimisen hallintajärjestelmien ominaisuuksien lisäksi osaa

kerätä tietoa siitä, miten oppilaat käyttävät ympäristöä, ja löytää näin uutta tietoa (Udupi ym., 2016). Käytännössä tämä tapahtuu siten, että ympäristön keräämästä datasta tehdään eri algoritmien ja mallien kautta tilastoanalyysiä, jolla puolestaan voidaan havaita uutta tietoa oppijoiden käyttäytymisestä, ja jota taas puolestaan voidaan hyödyntää ympäristön kehittämiseksi eteenpäin (Siemens, 2013). Näin voidaan saada tietoa mm. siitä, miten oppilaat suoriutuvat eri kursseista, tehdä suosituksia yksittäisten oppilaiden kurssi-ilmoittautumisten perusteella ja visualisoida helpommin tietoa siitä, miten ympäristöä käytetään ja millaiset trendit ja relaatiot eri datajoukoilla on (Sin & Muthu, 2015). Taulukossa 1 viitattu ryvästen ja mallien luonti viittaavatkin niihin tekniikoihin, joilla tätä dataa voidaan tilastollisesti havaita: ryvästen teko viittaa siihen, kuinka eri datajoukkoja järjestellään omiin ryppäisiin siten, että saman ryppään sisäiset datajoukot ovat keskenään mahdollisimman samankaltaisia, ja toisaalta mahdollisimman erilaisia niistä datajoukoista, jotka eivät kuulu ryppäeseen (Chatti ym., 2013). Mallien luonti taas viittaa yleisluontoisemmin samankaltaiseen datan järjestämiseen ja luokitteluun. Tutkielmassa ei keskitytä tarkasti prosessin tekniseen puoleen, mutta on kuitenkin tärkeä tietää, mitkä ovat prosessin perusteet.

Yksi tärkeimmistä ympäristön ominaisuuksista muihin verkko-oppimisympäristöihin verrattuna on uuden tiedon löytäminen: taulukon 1 ominaisuus ”tiedon löytäminen” viittaakin siihen, miten hyödyntämällä mm. tilastotiedettä, koneoppimista, tietokantoja sekä datan visualisointia voidaan löytää uusia, hyödyllisiä malleja suurista datajoukoista (Daniel, 2015). Se, millä tavoin löydettyä tietoa voidaan hyödyntää edelleen, tutkitaan tarkemmin luvussa 4.

2.3 Verkko-oppimisen edut ja haitat

Jotta voitaisiin ymmärtää, mitä hyötyjä big datan integroimisesta perinteisiin verkko-oppimisympäristöihin olisi, olisi myös hyvä tietää, mitä etuja ja rajoitteita perinteinen verkko-oppiminen on oppimiselle ja opetukselle suonut. Siksi seuraavissa kappaleissa käsitellään sitä, miten verkko-oppiminen tehostaa oppimista ja millaisia haittoja sillä on. Verkko-oppimista verrataan perinteiseen luokkahuoneopetukseen, jonka apuna toimii Zhangin, Zhaon, Zhoun ja Nunamakerin (2004) tutkimuksessaan kokoama taulukko perinteisen luokkahuoneoppimisen ja verkko-oppimisen eduista ja haitoista (taulukko 2). Verkko-oppimisella tarkoitetaan tässä luvussa niitä metodeja, jotka ovat yleisimpiä nykyisessä opetuksessa, eivätkä ne siis kata niitä tekniikoita, jotka edellä mainittu big datan integroiminen sallisi. Myöhemmin luvussa 4 käsitellään sitä, miten nämä edut ja haitat näkyvät big data -integroidun oppimisen kanssa.

2.3.1 Edut

Yksi verkko-oppimisen keskeisimmistä ja helpoiten havaittavista eduista perinteiseen, kasvotusten tapahtuvaan opetukseen verrattuna on sen riippumattomuus tietyistä ajasta ja paikasta: oppilaiden ei tarvitse kokoontua yhteen paikkaan samanaikaisesti (vrt. luokkahuone), vaan oppiminen tapahtuu maantieteellisesti riippumattomassa verkko-oppimisympäristössä Tämä taas mahdollistaa paremman joustavuuden opetuksessa: oppimateriaaliin voi tutustua silloin, kun haluaa, lähes missä tahansa ympäristössä (Zhang ym., 2004). Tätä ominaisuutta voikin soveltaa myös työelämään: jos esimerkiksi työntekijät suorittavat harjoituksia verkko-opetuksena, se mahdollistaa opetuksen aikatauluttamisen sen mukaan, miten he itse ehtivät niitä käydä (Navimipour & Zareie, 2015), eikä siis vie aikaa muiden työtehtävien suorittamiselta. Samalla tämä ominaisuus tekee opetuksesta paljon kustannustehokkaampaa (Zhang ym., 2004), sillä opetusta varten ei tarvitse varata tiloja tai palkata henkilöstöä.

Toisaalta etuna voidaan pitää myös verkko-oppimisen avoimuutta: oppimateriaalin ja muun oppimisresurssien digitointi on mahdollistanut näiden resurssien uudelleenkäytön eri maiden ja kouluympäristöjen välillä (Richter & McPherson, 2012, Stoffregen, Pawlowski & Pirkkalainen, 2015, mukaan), ja näin pääsy oppimateriaaliin on tullut paljon helpommaksi. Myös jo internetiin pääsy voi taata edullisen, korkeatasoisen opetuksen (Aparicio, Bacao & Oliveira, 2017), sillä vastaavanlaista oppimateriaalia löytyy jo paljon. Täten verkko-opiskelu on luonteeltaan paljon perinteistä opetusta avoimempaa, eikä välttämättä tuota yhtä suuria kustannuksia kuin perinteinen opetus.

Eräs jopa yllättävä verkko-oppimisessä tutkittu etu liittyy tiimityöskentelyyn ja luovuuteen: Lee, Lee, Seo ja Choi (2015) osoittavat tutkimuksessaan, kuinka verkko-oppimisympäristöissä toteutettu, ryhmän kesken tapahtuva tiedon jakaminen ja jaettu johtajuus voivat olla hyödyllisiä paitsi ryhmätyöskentelylle, myös ryhmän luovuudelle. Tästä voidaan päätellä, ettei verkko-oppimisympäristö itsessään rajoita ryhmätyöskentelyä tai toimi esteenä käyttäjiensä sosiaalisten taitojen kehitykseen, vaan voi jopa edistää niitä. Toisaalta Leen ym. (2015) tutkimuksessa keskeisessä roolissa olivat ympäristön sallimat ominaisuudet tiedon jakaminen ja jaettu johtajuus, joten voidaan olettaa, että jotta näitä taitoja voitaisiin edistää, täytyy opetuksen ja oppimisympäristön tarjota riittävät mahdollisuudet ryhmätyöskentelyyn.

Verkko-opiskelu tuo mukanaan myös mahdollisuuden oppimisen personointiin, vaikkei tätä ominaisuutta ole juuri hyödynnetty nykyaikaisissa oppimisympäristöissä. Suurin osa verkossa toteutetusta opetuksesta käyttää vielä perinteistä, homogeenistä oppimismallia, missä kaikille opiskelijoille tarjotaan tietyt oppimismateriaalit ja -välineet, vaikka heidän taustansa, oppimistyytinsä ja kognitiiviset taitonsa voivat erota paljonkin toisistaan. (Xu, Huang, Wang & Heales, 2014.) Tämä taas johtaa siihen, että opiskelijat eivät pääse etenemään samaan tahtiin, sillä kaikille tietty opetusmuoto ei syystä tai toisesta tehoa. Perinteinen oppiminen onkin usein tällaista: laajalle joukolla

erilaisia opiskelijoita tarjotaan samat oppimateriaalit, joita opetetaan samalla tyylillä. Kuitenkin oppimisympäristön sekä oppimisen sisällön ja tahdin muokkaus yksittäisten oppijoiden erityispiirteiden ja oppimistyylien mukaiseksi voi toimia valtavana etuna korostamaan verkko-oppimisen tehokkuutta (Aparicio ym., 2017), ja verkko-oppiminen onkin kokonaisuudessaan siirtymässä yhä enemmän opettajakeskeisestä oppimisesta oppijakeskeiseen oppimiseen, jossa ympäristön joustavuus ja personointi korostuvat entisestään. (Xu ym., 2014). Näistä mahdollisuuksista kerrotaan enemmän luvussa 4.

	Perinteinen luokkahuoneoppiminen	Verkko-oppiminen
Edut	<ul style="list-style-type: none"> - Välitön palaute - Tuttuus sekä opettajille että oppilaille - Motivoi oppilaita - Yhteisöllisyyden kehittäminen 	<ul style="list-style-type: none"> - Oppijakeskeinen ja oman tahdin mukainen - Joustava ajan ja paikan suhteen - Kustannustehokas oppijoille - Potentiaalisesti saatavilla maailmanlaajuisesti - Rajaton pääsy tietoon - Arkistomainen kyky tiedon jakamiseen ja uudelleenkäyttöön
Haitat	<ul style="list-style-type: none"> - Opettajakeskeinen - Rajoitukset ajan ja paikan suhteen - Kalliimpaa verkko-oppimiseen verrattuna 	<ul style="list-style-type: none"> - Välittömän palautteen puute - Vaatii enemmän valmistelua opettajalta - Ei sovi kaikille ihmisille - Herättää mahdollisesti enemmän turhautumista, ahdistumista ja hämmennystä

TAULUKKO 2: Perinteisen oppimisen ja verkko-oppimisen edut ja haitat (Zhang ym., 2004)

2.3.2 Haitat

Verkko-oppiminen tuo mukanaan myös haittoja. Jotkut verkko-oppimisympäristöt tarjoavat vain tekstipohjaista oppimateriaalia, mikä voi aiheuttaa opiskelijoissa tylsistymistä ja siten vaikeuttaa oppimateriaalin omaksumista (Zhang ym., 2004). Tämä toimii esimerkkinä siitä, kuinka verkko-oppimisympäristön ominaisuuksia ei aina hyödynnetä riittävästi, minkä osasyynä voidaan pitää epäsuhtaa ominaisuuksien tarjonnan ja käytön suhteen: vaikka erilaisia teknologioita ominaisuuksia tarjotaan runsain määrin osaksi verkko-oppimisympäristöjä, on vain vähän opetuksellisia periaatteita niiden käyttöön. Ei ole ohjeistuksia siitä, miten pedagogisesti analysoida, suunnitella, tuottaa ja hallita verkko-oppimisympäristöjä (Alonso ym., 2005.), mikä voi toisaalta luoda haasteita ympäristöjen käytölle.

Toinen ympäristön tuoma haaste on sen jäykkyys: verkossa opiskellessa oppilaalle voi herätä kysymyksiä, joihin vastaaminen ei suju yhtä nopeasti, kuin luokkahuoneen dynaamisemmassa ympäristössä. Täten verkko-opiskelussa palaute ei ole niin välitöntä kuin perinteisessä opiskelussa, mikä ilmenee myös Zhangin ym. (2004) tekemässä tutkimuksessa. Kun ympäristö ei tarjoa riittävästi interaktiota oppijan ja oppimateriaalin välillä, ja kun se ei ole riittävän joustava, myös oppimiseen on huonoja vaikutuksia: on esimerkiksi paljon tehottomampaa ja enemmän aikaa vievää etsiä vastausta kysymykseen monen tekstikappaleen tai videon seasta, kun sitä voisi vain kysyä opettajalta. (Zhang ym., 2004).

Kolmas haaste verkko-oppimisessa liittyy sen käyttöönoton vaikeuteen ja tästä seuraaviin ongelmiin. Vaikka yleensä ottaen teknologia pyrkii avustamaan käyttäjiään, voi sen käyttöönotossa ilmetä vaikeuksia – etenkin sen käyttöön tottumattomille. Haastetta korostaa entisestään opetukseen tarkoitettujen sovellusten suuri määrä, mikä ilmenee myös Sjödénin (2017) tekemässä tutkimuksessa: opetukseen tarkoitettuja sovelluksia on satoja tuhansia jo pelkästään AppStoressa, eikä juuri minkään näiden sovellusten tehokkuutta ole tieteellisesti arvioitu – eikä niitä jo määränsä takia realistisesti voikaan arvioida. Toisaalta perinteistä opetusta on järjestetty jo pitkään, ja sillä on myös omat positiiviset vaikutuksensa sekä oppimiseen että myös yhteisöllisyyden kehittämiseen: se on tuttua sekä opettajille että oppilaille, kehittää oppilaiden ja opettajien muodostamaa kouluyhteisöä ja motivoi oppilaita muodostamansa yhteisön avulla (Zhang, 2004). Tästä syystä uuden teknologian integroiminen vanhaan, totuttuun järjestelmään voi olla erittäin haastavaa ja riskialtista – etenkin, jos uusi teknologia muuttaisi opetusta radikaalisti.

Kaiken kaikkiaan verkko-oppimisella on paljon haasteita, jotka keskittyvät etenkin verkko-oppimisympäristöjen rajallisiin ominaisuuksiin, jäykkyyteen ja käytön vaikeuteen. Ratkaisuna joihinkin näistä haasteista esitellään oppimisen analytiikka, jonka tuomia mahdollisuuksia esitellään luvussa 4.

3 BIG DATA

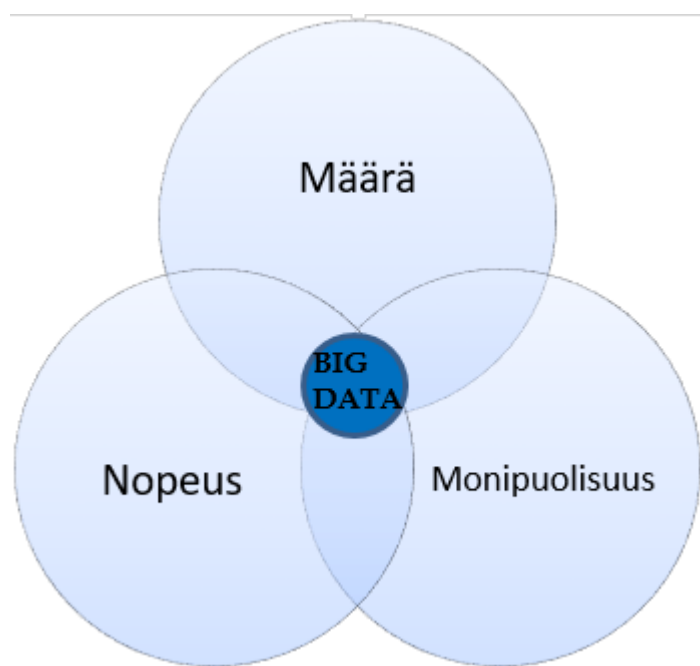
Big datasta puhuttaessa tarkoitetaan yleensä massiivisia datamääriä, joita hyödynnetään erilaisilla kehittyneillä teknologioilla. Kuitenkin jopa aiheen tutkijoiden parissa käsitteestä liikkuu erilaisia määritelmiä, jotka voivat olla ristiriidassa toistensa kanssa. Tässä kappaleessa määritellään big datan käsite, selvitetään sen prosessimalli sekä havainnollistetaan, miten sitä hyödynnetään nyky-yhteiskunnassa.

3.1 Big datan määritelmä

Big data on abstrakti käsite, jolle ei ole virallista määritelmää. Yleisesti ottaen se tarkoittaa mitä tahansa dataa, jota ei pystytä havaitsemaan, keräämään, prosessoimaan ja hallitsemaan perinteisellä IT:llä, ohjelmistoilla tai laitteilla – ainakaan siedettävällä aikavälillä (Chen, Mao & Liu, 2014; Sin & Muthu, 2015.) Kuitenkin edistysaskeleet tietokantateknologioissa ovat mahdollistaneet tällaisten suurien ja monimutkaisten datamäärien keräämisen ja säilyttämisen. Lisäksi on olemassa analyttisiä työkaluja, joilla tämä monimutkainen data voidaan muuntaa mielekkäiksi malleiksi ja arvoiksi ilmiöllä, joka yleensä tunnetaan nimellä big data. (Daniel, 2015.) Täten big datalla viitataan sekä konkreettisiin datayksiköihin että eri työkaluihin ja teknologioihin, joilla sitä pystytään käsittelemään (Sin & Muthu, 2015). Tässä luvussa big dataa käsitellään datajoukkona, joka tiivistetään sen olennaisimpiin käsitteisiin, ja seuraavassa luvussa keskitytään siihen, miten big dataa voidaan kerätä, analysoida ja hyödyntää.

Useissa määritelmissä big datan käsite tiivistetään sen keskeisimpiin ominaisuuksiin. Kuitenkin useilla tutkijoilla on eriäviä mielipiteitä siitä, mitkä nämä keskeiset ominaisuudet ovat: useat tutkijat määrittävät big datan kolmen ominaisuuden mukaan: datan määrän, monipuolisuuden ja nopeuden (Chen, Chiang & Storey, 2012; Kwon, Lee & Shin, 2014), kun taas osa lisää joukkoon datan arvon (Chen, 2014) ja osa vielä sen todennuksen ja totuudenmukaisuuden

(Daniel, 2015). Näillä lisätyillä ominaisuuksilla korostuu se, miten arvokasta kerätty data on, kuinka luotettavaa ja varmaa se on sekä kuinka totuudenmukaista se on (Daniel, 2015). Kuitenkin määritelmiä yhdistää se, kuinka ne kaikki sisältävät kolme yleisintä datan ominaisuutta – datan määrän, monipuolisuuden ja nopeuden – ja siksi voidaan päätellä, että nämä ominaisuudet ovat erittäin olennaisia big datalle. Tästä syystä tutkielmassa sitä määritellään näiden kolmen ominaisuuden pohjalta. Tutkielmaa varten tehdyssä kuviossa 1 esitetään big datan rakenne näiden kolmen ominaisuuden risteymänä.



KUVIO 1: Big datan kolme ominaisuutta

3.1.1 Määrä

Määrä viittaa big datan massiiviseen kokoon: tyypillisesti big datasta puhuttaessa sen kokoluokka on monia tera- tai petatavuja. Vertailun vuoksi, yhteen teratavuun mahtuisi 1500 CD:tä tai 220 DVD:tä (Gandomi & Haider, 2015.) Big datan koosta kertoo myös se, kuinka Google käsittelee satoja petatavuja dataa osana operaatioitaan ja Facebook luo kuukausittain yli 10 petatavu dataa (Chen ym., 2014). Tällainen datamäärä on nykyisin valtava, mutta varastointikapasiteetin kasvaessa se ei kenties tulevaisuudessa riitä kuvaamaan "suuren datamäärän" määritelmää, joten se yksin ei voi määritellä big dataa. Big datan suosion äkillisen kasvun myötä ei myöskään ole yleistä konsensusta siitä, miten suurta datan tulee olla tullakseen määritellyksi big dataksi, ja tätä rajausta vaikeuttaa entisestään erilaiset datan muodot: keskenään samankokoiset mutta erityyppiset datajoukot voivat vaatia erilaisia tapoja hallita dataa, kuten esimerkiksi verrattaessa laskennallista dataa videomuodossa olevaan dataan. (Gandomi & Haider, 2015.) Näihin

erityyppisiin datajoukkoihin viitataan myös datan monipuolisuudella, jota käsitellään myöhemmin. Siksi big datan nimestä huolimatta sen pelkkä koko ei voi määrittellä sitä tarpeeksi tarkasti, joten tarvitaan muita attribuutteja määrittämään se tarkemmin.

3.1.2 Nopeus

Nopeus viittaa toisaalta siihen, miten nopeaan tahtiin uutta dataa luodaan, ja toisaalta siihen, miten nopeasti tätä dataa tulisi vuorostaan pystyä analysoimaan ja hyödyntämään. Esimerkiksi älypuhelinien suosio on johtanut ennennäkemättömään datan tuotannon määrään, joka puolestaan on johtanut tarpeeseen tämän datan reaaliaikaiseen analysointiin ja analyysiin perustuvaan suunnitteluun. (Gandomi & Haider, 2015.) Myös yleisellä informaatioteknologian kehityksellä on ollut valtava vaikutus datan tuotannon kasvuun, sillä se on tehnyt datan tuotannon paljon helpommaksi: esimerkiksi YouTubeen ladataan joka minuutti keskimäärin 72 tuntia videoita (Chen ym., 2014). Datan tuotannon nopeudesta kertoo paljon myös se, kuinka esimerkiksi Facebook tuottaa kuukausittain yli 10 petatavua lokidataa, ja Google prosessoi toiminnassaan satoja petatavuja dataa (Chen ym., 2014). Tällaisen datamäärän käsittelyn ja sen eri lähteiden välisen integroinnin vaatima aika suhteessa siihen, miten nopeasti uutta dataa syntyy, onkin suuri rajoite ja haaste big datan parissa. Samalla haasteita muodostuu siitä, kuinka tätä jatkuvasti syntyvää dataa tulkitaan: jos suuressa määrässä dataa ilmenee virheitä, on niitä vaikea paikallistaa datamassan keskeltä, ja ne voivat vuorostaan johtaa vääriin datan tulkintoihin (Abdelouarit, Sbihi & Aknin, 2015.)

3.1.3 Monipuolisuus

Big datan monipuolisuus viittaa siihen, kuinka keskenään erilaista kerätty data on. Kerätty data voi olla jäsentynyttä, puolijäsentynyttä tai jäsentymätöntä. Jäsentynyt data on taulukoitavaa dataa, joka koostaa eri relaatiotietokantoja ja taulukkolaskentaohjelmia. (Gandomi & Haider, 2015.) Toisin sanoen se on erittäin tarkasti jäsennettyä dataa, jota tietokoneiden on helppo käsitellä. Kuitenkin tällainen data koostaa vain noin 5% kaikesta datasta (Cukier, 2010), ja enemmistö datasta on siis joko puolijäsentynyttä tai jäsentämätöntä.

Teksti, kuvat, äänet ja video ovat esimerkkejä jäsentymättömästä datasta. Tällaisella datalla ei ole säännöllistä, järjestäytyntä rakennetta, jota koneen olisi helppo analysoida (Gandomi & Haider, 2015.) Tämä tuokin omat haasteensa datan käsittelyyn: miten analysoida nopeasti ja tarkasti dataa, jota luodaan jatkuvasti, ja joka on luonteeltaan jäsentymätöntä ja määrältään valtavaa?

Jäsentyneen ja jäsentymättömän datan välimaastossa on puolijäsentynyt data. Tällainen data ei mukaudu tiukkiin standardeihin, vaan sisältää elementtejä, jotka esimerkiksi käyttäjä määrittää itse. Esimerkiksi verkossa datan siirtoon käytetty metakieli XML (Extensible markup language) on

puolijäsentynyttä dataa, sillä se sisältää käyttäjien luomia määrittelyjä, mitkä tekevät datasta koneelle lukukelpoista (Gandomi & Haider, 2015.)

Konkreettisten tiedostojen lisäksi big data mahdollistaa myös uudenlaisten muuttujien keräämisen ja seurannan: esimerkiksi sellaisia muuttujia kuin klikkausten lukumäärää ja sivulla vietettyä aikaa voidaan valvoa ja kerätä (Daniel, 2015), ja tämä tuokin vastaavasti paljon mahdollisuuksia tämän datan hyödyntämiselle. Näitä mahdollisuuksia tutkitaan enemmän luvussa 4.

3.2 Big datan hyödyntäminen osana oppimista

Edellisessä luvussa big dataa käsiteltiin vain sen ominaisuuksiensa puolesta, mutta olennaista on myös se, miten sitä käytetään. Seuraavaksi määritellään tutkielmalle olennainen käsite oppimisen analytiikka, ja samalla mallinnetaan, kuinka big datan keräys ja hyödyntäminen onnistuu.

Läheinen käsite big datan ja oppimisen yhdistämisessä on oppimisen analytiikka. Yleisesti ottaen se luo eri metodeja, joilla oppimisaiheiset datajoukot voidaan paremmin hyödyntää osana oppimisprosessia. (Chatti ym., 2013). Alana se on kuitenkin hyvin laaja: Chatti ym. (2013) mukaan oppimisen analytiikka on monitieteinen ala, johon kuuluu koneoppiminen, tekoäly, tiedonkeruu, tilastot ja visualisaatio. Se ei myöskään ole aidosti uusi tutkimusalue, sillä se lainaa paljon muista aloista ja yhdistää useita olemassaolevia tekniikoita (Chatti ym., 2013).

Käytännössä oppimisen analytiikkaa käytetään hyväksi mallien tunnistamiseksi oppilaiden digitaalisten palvelujen käytöstä, jotta ymmärrettäisiin paremmin sekä oppimista että sen hallinnollisia käytänteitä. (Godwin-Jones, 2017). Chatti ym. (2013) esittelevät tutkimuksessaan mallin prosessista, jolla tämä onnistuu (kuvio 2). Malli koostuu kolmesta vaiheesta: datan keräyksestä ja esikäsittelystä, sen analysoinnista ja käsittelystä sekä datan jälkikäsittelystä. Eräs olennainen piirre mallissa on, että vaikka kyseinen kuva keskittyy oppimisen analytiikkaan ja täten oppimistarkoitukseen kerättyyn dataan, myös muunlaiset, esim. liiketoiminnalliset hyödyntämistavat big datalle noudattavat toimitaan samaa prosessia.



KUVIO 2: Oppimisen analytiikka prosessina (Chatti ym., 2013)

Prosessin ensimmäisessä vaiheessa kerätään opetuksellista dataa, jota prosessin myöhemmissä vaiheissa käytetään hyväksi. Tätä dataa kerätään useasta verkko-oppimisympäristöstä ja -järjestelmästä, ja toimii koko prosessin perustana. Tässä vaiheessa dataa myös esikäsitellään prosessiin sopivaksi: siitä karsitaan prosessille epäolennaisia piirteitä, ja sitä muunnetaan sopivampaan muotoon prosessin onnistumiseksi. (Chatti ym., 2013.) Tätä vaihetta voikin pitää yhtenä prosessin tärkeimmistä vaiheista, sillä tulevien analyysien tarkkuus perustuu siihen, miten paljon ja kuinka sopivaa kerätty data on. Kuten jo aikaisemmin mainittiin, kerätty data voi big datana olla todella monimuotoista ja määrältään laajaa, eikä rajoitu vain esimerkiksi oppilaan sivulle lataamille erimuotoisille tiedostoille, sillä verkko-oppimisympäristö voi erittäin monipuolisesti valvoa ja kerätä tietoa oppilaan käyttäytymisestä verkko-oppimisympäristöstä. Esimerkiksi sellaisia muuttujia kuin klikkausten lukumäärää ja sivulla vietettyä aikaa voidaan kerätä, ja tästä voidaan päätellä yhä monipuolisempaa tietoa, kuten oppilaan sinnikkyyttä sekä opitun tiedon hallintaa (Daniel, 2015). Tällaiset toimet ovatkin keskeisiä prosessin seuraavalle vaiheelle: kerätyn datan analysoinnille ja dataan kohdistuvaan toimiiin.

Datan käsittelyvaiheessa kerättyä dataa analysoidaan ja sille voidaan suorittaa erilaisia toimenpiteitä sen mukaan, mikä päämäärä sen keräykselle oli. Erilaiset analyysitekniikat voivat paljastaa datasta piilotettuja malleja, joita voidaan edelleen hyödyntää tehokkaamman oppimisen aikaansaamiseksi (Chatti ym., 2013). Tällä tekniikalla voidaan esimerkiksi analysoida oppilaiden suoriutumista ja tehdä tämän pohjalta ennusteita oppilaiden tulevaisuudesta, jolloin esimerkiksi sellaiset oppilaat, jotka ovat vaarassa keskeyttää kurssin tai saada hylätyn arvosanan voidaan tunnistaa oppilaiden joukosta (Sin & Muthu, 2015.), ja tilanteeseen voidaan vaikuttaa ajoissa. Yksi keskeinen oppimisen analytiikan tekniikka, mikä mahdollistaa tällaisen mallien tunnistamisen

suuren datamassan keskeltä, on tiedon visualisointi: oppimisen analytiikka auttaa tekemään näistä piilotetuista malleista näkyviä, jolloin ne havaitaan helpommin ja ne voidaan käsitellä oikealla tavalla (Sin & Muthu, 2015). Dataan kohdistuvat toimet puolestaan voivat olla todella monipuolisia: mahdollisia ovat mm. edellä mainittu kurssin "hylkäämisvaarassa" olevien oppilaiden tunnistaminen (Sin & Muthu, 2015) ja suositusjärjestelmän soveltaminen kerätyn datan perustella (Dwivedi & Roshni, 2017) sekä oppimateriaalin ja verkko-oppimisympäristön personointi (Siemens, 2013). Näihin toimiin keskitytään enemmän luvussa 4, mutta ne kaikki perustuvat siihen, että dataa kerätään, havaitaan ja niiden muodostumiseen johtaviin piirteisiin voidaan puuttua.

Jälkikäsittelevaiheessa korostuu koko prosessin jatkuva parantaminen. Tässä vaiheessa muista lähteistä kootaan uutta dataa, kerättyä ja käsiteltyä dataa jalostetaan edelleen, uusia attribuutteja määritellään uusia iteraatioita varten, uusia indikaattoreita tunnistetaan, muuttujia muunnetaan analyysia varten tai valitaan aivan uusi tapa tehdä analyysiä (Chatti ym., 2013). Keskeistä vaiheelle on, että huomio kiinnittyy itse prosessiin ja sen toimintaan, ja että näin turvataan sekä prosessin onnistuminen että jatkuvuus tekemällä siihen tarvittavia muutoksia sen kehittämiseksi eteenpäin (Chatti ym., 2013).

3.3 Big datan nykyisiä käyttökohteita

Havainnollistuksena big datan käytöstä nyky-yhteiskunnassa ja osoituksena siitä, että sen käyttö myös osana verkko-oppimisympäristöjä on paitsi mahdollista myös potentiaalisesti hyödyksi oppimiselle, on eri lähteistä kerätty tietoa siitä, miten big dataa nykyisin käytetään. Nämä käyttökohteet painottuvat tällä hetkellä yritystoimintaan, mutta myös oppilaitoksilla on nykyteknologian avulla mahdollisuus käyttää näitä tekniikoita hyväkseen.

Monet organisaatiot käyttävät hyväkseen keräämäänsä dataa perustellakseen operatiivisia ja strategisia toimiaan. Tällainen datan keräys ja käyttö päätöksenteossa ei ole uutta: yritykset ovat keränneet ja analysoineet suuria määriä dataa aina tietovarastojen synnystä saakka 1990-luvulla. (Daniel, 2015.) Kuitenkin tiedon määrän ja monimutkaisuuden kasvaessa myös tästä prosessista on tullut entistä hankalampaa, ja siksi organisaatiot ovat alkaneet osoittaa kasvavaa kiinnostusta big datan osoittamaan suureen potentiaaliin (Chen ym., 2014). Tätä väitettä tukee se, miten viime vuosina lähes kaikki suuren kokoluokan yritykset kuten Google, IBM, Oracle, Microsoft, Amazon ja Facebook ovat aloittaneet omat big data-projektinsa. Näistä yksin IBM on sijoittanut vuodesta 2005 alkaen 16 miljardia dollaria big dataan liittyviin hankintoihin. (Chen ym., 2014.)

Eräs big datan suurimmista ja eniten esille tuoduista sovelluskohteista liiketoiminnassa on asiakaskunnan ymmärtäminen ja kohdentaminen. Tällä keinolla yritykset käyttävät asiakkaistaan keräämäänsä dataa hyväkseen

ymmärtääkseen paremmin asiakaskunnan käytöstä, preferenssejä ja asiakaskuntaa ylipäättään. (Marr, 2017). Tähän perustuvat myös useat suositusjärjestelmät: esimerkiksi Netflix analysoi suuria määriä dataa, jotta se pystyisi suosittamaan käyttäjilleen elokuvia, joista he voisivat pitää. Käytännössä palvelun sisältöä luokitellaan sen eri elementtien mukaan ja verrataan sitä niihin elokuviin, joita käyttäjä on katsonut. Tästä käyttäjädatasta ja sisällön luokittelusta taas saadaan monimutkaisten algoritmien seurauksena joukko elokuvia, joista käyttäjä voisi pitää. (Plummer, 2017.) Keskeistä siis on, että käyttäjistä kerätään paljon dataa, ja tätä dataa jalostetaan sitten pidemmälle, jotta siitä voisi tehdä malleja, jotka puolestaan pystyvät ennustamaan käyttäjän käytöstä ja mieltymyksiä. Samankaltaiset suositusjärjestelmät ovat hyvin suosittuja myös verkkokaupoissa, sosiaalisessa mediassa sekä muissa eri palveluissa, jotka tukeutuvat muiden käyttäjien mieltymyksiin ja suosituksiin. Big datan analyysia käyttämällä jopa pk-yritykset voivat louhia suuria määriä puolijäsentynyttä dataa parantaakseen verkkosivujansa ja käyttääkseen tehokkaita, kohdennettuja suositusjärjestelmiä (Gandomi & Haider, 2015), mikä herättääkin kysymyksen: mikseivät oppilaitokset voisivat käyttää samaa tekniikkaa? Vaikka nykyisin suositusjärjestelmiä käytetäänkin yksinomaan osana yritystoimintaa, hiljattain suositusjärjestelmä on kerännyt kiinnostusta myös koulutuksen parissa, missä suosituksia kerättäisiin vastaavasti oppilaille, opettajille ja kouluille. (Dwivedi & Roshni, 2017). Tästä mahdollisuudesta kerrotaan enemmän seuraavassa luvussa.

4 BIG DATA OSANA VERKKO-OPPIMISYMPÄRISTÖJÄ

Big datan hyödyntämiselle osana verkko-oppimisympäristöjä on useita eri malleja. Tässä luvussa tutustutaan niihin ja käydään läpi, miten ne voivat hyödyttää oppimista ja opettamista.

4.1 Big datan hyödyntämistavat verkko-oppimisympäristöissä

Vaikka luvussa 3 esiteltä, big dataa hyödyntävän verkko-oppimisympäristön toimintaprosessi on sama, on sillä paljon eri käyttömahdollisuuksia. Tässä luvussa käydään läpi näitä eri tapoja hyödyntää big dataa osana verkko-oppimisympäristöjä. On kuitenkin tärkeää huomata, että tällaista data-analyysia käytetään opetuksen apuna, eikä sen tarkoituksena ole täysin korvata kyvykkäiden opettajien kokemusta, ammattitaitoa, intuitiota ja arvostelua (Picciano, 2015).

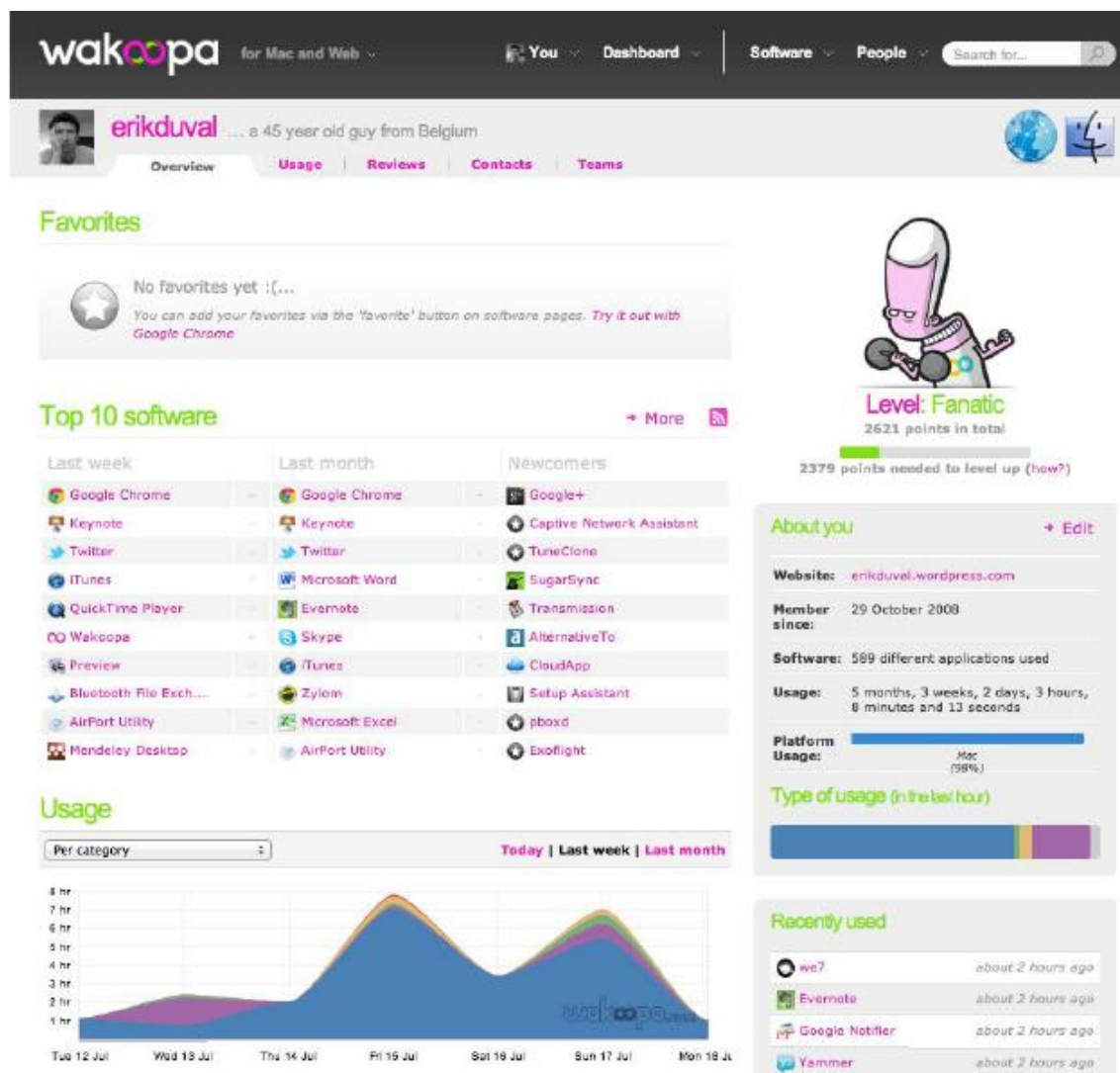
4.1.1 Datan visualisaatio ja ennusteet

Eräs big datan sovelluskohde verkko-oppimisympäristöissä on datan visualisaatio ja ennusteet kerätyn datan perusteella. Kun opiskelijat käyttävät oppimisteknologiaa hyväkseen, he jättävät toimiensa tuloksena datajälkiä, joista voidaan selvittää heidän tunteitaan, sosiaalisia yhteyksiä, aikomuksia ja tavoitteita. Tätä dataa voidaan edelleen hyödyntää oppilaiden suorituksen valvomiseen ajan myötä -lukuvuodesta tai vuodesta toiseen (Daniel, 2015.) Tämä data koostuu sekä eksplisiittisestä toiminnasta, kuten tehtävien tekemisestä ja kokeiden suorittamisesta, kuin myös esimerkiksi verkossa tapahtuvasta sosiaalisesta vuorovaikutuksesta, forumeille postauksesta ja muista aktiviteeteista, joita ei suoraan arvioida osana oppilaan koulutuksellista edistystä (Johnson, Smith, Willis, Levine & Haywood, 2011). Big datan rooli ja haasteet korostuvatkin siinä, miten näitä erilaisia ja suuria datajoukkoja

lähdetään tutkimaan, vaikka sen tekniikat antavatkin työkaluja datan tilastolliseen esittämiseen.

Kuitenkin datan tulkinta erimuotoisista taulukoista ja raporteista ei aina ole helppoa järjestelmän käyttäjille, jotka usein eivät ole tottuneet tilastollisen datan käsittelyyn. Siksi datan esitys käyttäjäystävällisessä, visuaalisessa muodossa voisi parantaa datan tulkittavuutta ja analyysiä. (Chatti ym., 2013.) Esimerkkinä tästä on kuvio 3, jossa näytetään tietoja Duvalin (2011) käyttämistä sovelluksista osana wakoopa-ympäristöä: näkymä kertoo niistä sovelluksista, joita käyttäjä on viimeisen viikon ja kuukauden ajan käyttänyt, ja pystyy tekemään tästä datasta tilastollisia, helposti ymmärrettäviä malleja. Ympäristö pystyy myös vertaamaan tätä dataa muihin käyttäjiin ja suosittelemaan sen perusteella ohjelmistoja tai yhteystietoja. Tämän ansiosta käyttäjä tulee paljon enemmän tietoiseksi omista aktiviteeteistaan, ja oppimiseen sovellettuna vastaavilla menetelmillä voitaisiinkin kartoittaa oppimisaktiviteetteja sekä käytettyjä tai suositeltuja työkaluja ja yhdistää oppilaita tai opettajia eri sosiaalisissa verkostoissa. (Duval, 2011.) Jo tällä tavalla kerätyt tiedot auttaisivat toiminnan kehittämiseen: pelkästään oppilaiden oppimisympäristöjen käyttämisen ja saavutusten analysointi voivat auttaa opettajia löytämään datajoukoista malleja, mikä auttaa päätöksenteossa siitä, miten oppimista tulisi jatkossa suunnitella (Chatti ym., 2013). Toisin sanoen pelkkä ympäristön käytön valvonta voi paljastaa malleja ja trendejä oppimisesta, joita hyödyntämällä oppimista voi kehittää eteenpäin.

Valvonnan lisäksi big datan käyttö mahdollistaa toiminnan analyysin ja ennustamisen kerätyn datan perusteella. Toiminnan tavoitteena on luoda malli, joka pyrkii ennustamaan sekä yksittäisten opiskelijoiden nykyistä tietämystä että tulevia suorituksia sen perusteella, millaisia aktiviteetteja ja saavutuksia opiskelija on nykyhetkellä saanut aikaan (Chatti ym., 2013). Tätä mallia voidaan hyödyntää esimerkiksi tunnistamaan opiskelijoita, jotka ovat vaarassa pudota kurseilta, ja edelleen varmistaa, että he saavat tarvitsemaansa apua ajoissa (Siemens, 2013; Daniel, 2015; Johnson ym., 2011; Picciano, 2015). Toisin sanoen dataa voidaan tuoda paremmin näkyville nykytilanteen lisäksi myös siitä, miten tilanne luultavasti tulee kehittymään, ja ennaltaehkäisemään ongelmatilanteita. Tällainen riskiryhmään kuuluvien opiskelijoiden tunnistus onkin ollut oppimisen analytiikan keskeisimpiä tutkimusaiheita korkeakoulutuksen piirissä (Johnson ym., 2011), mikä näkyy myös sitä koskevien tutkimusten määrästä. Se ei kuitenkaan ole ainoa tapa, jolla big dataa voidaan hyödyntää verkko-oppimisympäristöissä.



KUVIO 3: Wakoopa-ympäristön käyttöliittymä (Duval, 2011)

4.1.2 Suositusjärjestelmät

Myös luvussa 3 esiteltyä, perinteisesti yritystoiminnassa käytettyä suositusjärjestelmää kohtaan on levinnyt kiinnostusta oppimisen piirissä (Dwivedi & Roshni, 2017). Tätä järjestelmää voisi soveltaa esim. sopivien kurssien, aineiden ja oppimateriaalien suosittelemiseen yksittäisille opiskelijoille sen mukaan, mistä opiskelijat ovat kiinnostuneet (Lu, Wu, Mao, Wang & Zhang, 2015). Tämän tiedon saanti puolestaan onnistuisi joko eksplisiittisesti – keräämällä esim. käyttäjän tekemiä arvosteluja ympäristön sisällöstä – tai implisiittisesti, jolloin valvottaisiin käyttäjän toimintaa, ja tehtäisiin suositukset niiden pohjalta (Bobadilla, Ortega, Hernando & Gutierrez, 2013). Esim. Zaiane (2002) keskittyi tutkimuksessaan datan implisiittiseen keruuseen tutkimalla, kuinka datan louhimista voisi hyödyntää tunnistamaan tiettyjä oppilastyyppejä, ja tehdä suosituksia mm. suoritettavista tehtävistä sen perusteella, mitä oppimistavoiltaan samankaltaiset, hyvin menestyvät oppilaat

ovat valinneet. Tällä tavalla huomattavasti menestyvät oppilaat voisivat käyttää samoja tapoja oppimisensa edistämiseen kuin paremmin menestyvät, oppimistyylyiltään samankaltaiset oppilaat. Esimerkkinä Zaiane (2002) mainitsee, kuinka ennen kokeen suorittamista verkko-oppimisympäristössä ympäristö voisi suositella tiettyjä oppimateriaalin osioita, joita opiskelijat, jotka ovat pärjänneet kokeessa hyvin, ovat käyttäneet. Tämä auttaisi opiskelijoita löytämään oppiaineelle relevantteja tietoja nopeammin käyttämällä eräänlaisia oikoteitä hyödyllisiin materiaaleihin, jotka on löydetty analysoiduilla aiempien käyttäjien toimilla. Tällä tavalla suositusjärjestelmä ”oppisi” yhden tai useamman käyttäjän menneistä aktiviteeteista ja osaisi sen pohjalta ennustaa sivuja tai toimintoja, jotka käyttäjää voisi kiinnostaa. (Zaiane, 2002.) Sovelluskeinona tällä onkin paljon samankaltaisuuksia pelkän datan ennustamisen kautta, mutta sitä vastoin, että ympäristö vain osaisi visualisoida dataa ja luoda ennusteita sen perusteella, se pystyisi itse suoraan suositteluun tarvittavia oppimateriaaleja (Zaiane, 2002; Benhamdi, Babouri & Chiky, 2016; Lu ym., 2015), kursseja (Sin & Muthu, 2015; Lu ym., 2015) tai muuta oppimista tukevaa toimintaa, kuten oppimismuotoja (Lu ym., 2015).

Toinen vaihtoehtoinen tapa toteuttaa implisiittinen suositusjärjestelmä olisi personoitu oppimateriaalin suositusjärjestelmä. Tällainen järjestelmä käyttäisi laskennallista analyysiä yhdistämällä oppimateriaalin muodostavaa tietokantaa sekä yksittäisen oppijan oppimistavoitteita, ja loisi tämän perusteella otoksen oppimateriaalista, jota se suosittelee opiskelijalle (Lu ym., 2015). Tässä mallissa suositusjärjestelmä siis arvioisi tarvittavia toimia tapauskohtaisesti, analysoiden ensin itse opiskelijan tietoja ja sitten oppimateriaalia, ja loisi sitten näistä yhdistelmän, joka soveltuisi parhaiten yksittäiselle opiskelijalle. Edelliseen malliin verrattuna muihin käyttäjiin ei tarvitsisi kiinnittää huomiota, vaan suositukset olisivat henkilökohtaisia, ja toiminta keskittyisi yksittäisiin oppijoihin koko opiskelijajoukon sijaan. Keskeistä molemmille malleille kuitenkin on, että ne ottavat huomioon käyttäjien erityispiirteitä, ja pyrkivät tilastollisia malleja hyödyntämällä tekemään suosituksia oppimisen ohjaukseen siten, että se sopisi mahdollisimman hyvin näihin erityispiirteisiin.

4.1.3 Personointi

Datan visualisoinnin ja suositusten tekemisen lisäksi oppimisen analytiikan soveltaminen oppimisympäristöön sallii personoitujen, mukautuvien oppimisympäristöjen tarjonnan. Vaikka toisaalta myös edellä mainittu suositusjärjestelmä pystyy tuottamaan personoitua sisältöä suositteluun sitä käyttäjän mieltymysten mukaan (Lu ym., 2015.), on tähän muitakin keinoja kuin vain suositukset. Tutkielmassa eräänä tällaisena keinona esitellään oppimisympäristön personointi, jossa ympäristö räätälöi eri tavoin sisältöään oppijoille sopivaksi.

Personoiduissa oppimisympäristöissä keskeistä on niiden mukautuvuus. Suositusjärjestelmistä puhuttaessa painottui näkökulma siitä, että suositusjärjestelmät tekisivät suosituksia tehtävistä toimista: esimerkiksi Zaianan (2002) esimerkissä mainittiin, että järjestelmä voisi suositella hyödyllisiä oppimateriaaleja tenttiin valmistautuvalle opiskelijalle. Personoiduissa oppimisympäristöissä taas ympäristöt muuttuvat itse, usein käyttäjän toimien epäsuorana seurauksena: esimerkiksi ennustettaessa, että oppija luultavasti putoaa kurssilta, kuten luvussa 4.1.1., voi tämä tieto toimia pohjana oppimateriaalin personointina sopivammaksi käyttäjälle (Siemens, 2013). Tällaiset oppimisympäristöt perustuvat siis kerättyyn dataan oppilaista ja heidän yksittäisistä oppimisprosesseistaan (Schumacher & Ifenthaler, 2018), ja mahdollistavat henkilökohtaisten oppimiskokemusten luonnin, jotka vastaavat oppilaiden yksittäisiin oppimistyyliin (Daniel, 2015). Tällainen ympäristö voisi esimerkiksi seurata oppilaan interaktioita verkossa, analysoida oppilaan taitoja ja sitten verrata oppilaan tietämystä oppiaineen vaatimuksiin. Tämän arvioinnin seurauksena oppimisympäristö voisi tuottaa personoitua sisältöä. (Siemens, 2013.) Toisin sanoen tähänkin prosessiin liittyy keskeisesti datan aktiivinen keräys järjestelmästä ja sen hyödyntäminen ympäristön kehittämiseen, mutta huomio on keskittynyt yksittäisiin oppilaisiin koko ympäristön kehittämisen sijaan.

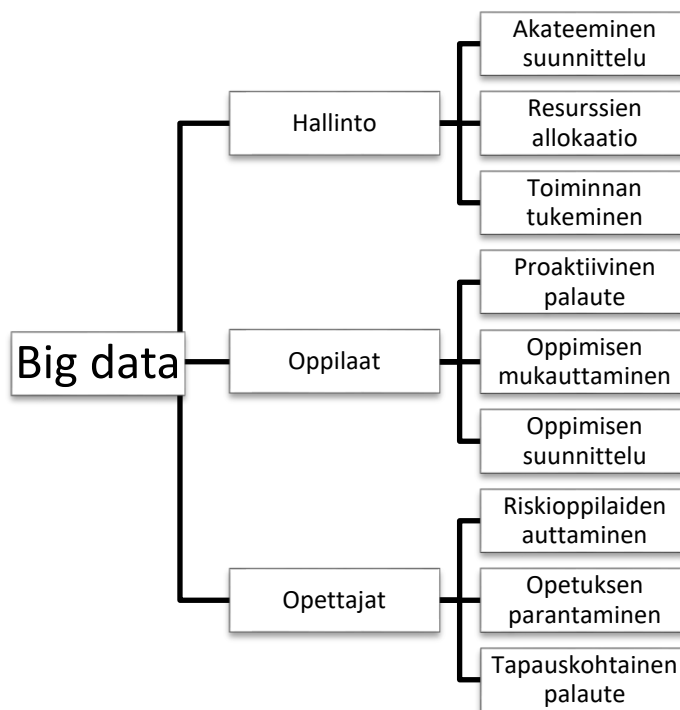
Keskeisintä tällaisille ympäristöille on kuitenkin se, miten ne käyttävät toiminnassaan oppimisen analytiikkaa – ja siten myös big dataa – hyväkseen. Suurin osa tällaisista big data -pohjaisista, personoiduista oppimisympäristöistä käyttää toiminnassaan oppimisen analytiikkaa joko mukauttaakseen toimintaansa – tarjoamalla automaattisesti oppimateriaalia oppijan yksittäisten tarpeiden mukaan – tai mukautuakseen oppijan mieltymyksiin suosittelemalla erilaisia oppimismuotoja (Chatti ym., 2013). Toisin sanoen nekin voivat suositella materiaalia ja tehdä analyysiä oppilaiden ympäristön käytön pohjalta. Big datan käyttö onkin olennaista, jotta personoidut oppimisympäristöt toimisivat.

4.2 Hyödyt

Tutkittaessa big datan integroimista oppimisympäristöihin tärkeitä ovat paitsi hyödyntämistavat, myös hyödyt itse. Tässä luvussa käydään läpi, kuinka big datan käyttö osana verkko-oppimisympäristöjä voi hyödyttää koulutuksen päätoimijoita.

Keskeisintä big datan käytössä osana oppimisympäristöä on datan kerääminen ja analysointi: sen olennaisin ero muista verkko-oppimisympäristöistä, kuten luvussa 2 kerrottiin, on se, kuinka vain se pystyy keräämään ja analysoimaan big dataa. Tämä big data on kuitenkin hyödytöntä, mikäli sitä ei keräämisen jälkeen hyödynnetä jollain tavalla: sen potentiaalinen arvo saavutetaan vasta, kun sitä käytetään päätöksenteon pohjana (Gandomi & Haider, 2015), ja oppilaitosten ylimmän portaan muodostavan hallinnon

suhteen tämä tieto onkin valtava voimavara juuri päätöksentekoon. Vaikka päätöksenteon voi yleensä määritellä koostuvan kahden tai useamman vaihtoehdon välillä valitsemisesta, moderneissa oppilaitoksissa se on keskeinen osa monimutkaisia, hallinnollisia prosesseja, kuten akateemista suunnittelua, käytäntöjen muodostamista ja budjetointia (Picciano, 2015), mikä ilmenee myös Danielin (2015) tutkimuksessa ja kuviossa 6. Yleensä ottaen tiedot siitä, miten oppilaat käyttävät ympäristöä, voidaan ajatella olevan vaikeaselkoista ja monimutkaista, mutta tämän monimutkaisen datan interaktiivisen visualisaation avulla myös hallinto pystyy tarkemmin keskittymään sen eri alueisiin (Siemens ym., 2011) ja tekemään päätöksiä sen perusteella, sillä data voidaan esittää käyttäjäystävällisessä, visuaalisessa muodossa (Chatti ym., 2013). Siispä hallinnon vastuulle jää toiminnan tukeminen ja suunnittelu, ja he saavat yksityiskohtaisempaa tietoa päätöksenteon ja resurssien allokoinnin pohjalle. Tämä näkyy myös kuviossa 4.



KUVIO 4: Big datan päämahdollisuudet koulutuksen kolmelle toimijalle (Daniel, 2015) (muokattu)

Oppilaille big data antaa paljon mahdollisuuksia: ympäristön kehittyessä parempaan suuntaan on oppilailla mahdollisuus entistä tehokkaampaan oppimiseen, ja vaikeudet ympäristön käytön suhteen näkyvät puolestaan dataa kerätessä. Näin ollen oppilaat kontribuoivat antamalla jatkuvasti proaktiivista, todenmukaista palautetta ympäristön käytöstään (kuviota 4), ja pystyvät tämän datan ja valitun big datan sovelluskeinon turvin mukauttamaan ja suunnittelemaan oppimistaan entistä tarkemmin. Näin oppijoiden saavutuksia ja motivaatiota voidaan parantaa antamalla datan visualisaatiolla tietoa heidän

suorituksistaan ja suosittelemalla sisältöä sen mukaan (Siemens ym., 2011). Suositusten pohjalta oppimista voidaankin tehostaa luomalla ”oikoteitä” hyödyllisiin materiaaleihin aiempien oppilaiden toimien pohjalta (Zaiane, 2002), mikä puolestaan tehostaa oppimista vähentämällä tämän materiaalin etsintään kuluvaan aikaan. Käytettäessä oppimisen personointia voidaan puolestaan varmistaa, että jokainen oppija saa resursseja ja sisältöä, joka sopii oppijan yksittäiseen tietämykseen (Siemens ym., 2011). Gell-Mannin mukaan (1996) oppimisprosessi on optimissaan, kun se on avustettua ja personoitua, ja tietokoneet ovatkin oppimisjärjestelmän potentiaalisia pelastajia sallimalla tällaisen personoidun oppimisen (Alonso ym., 2005).

Tämä ympäristön mukautuvuus voidaankin nähdä eräänlaisena ratkaisuna verkko-oppimisen perinteisiin ongelmiin: luvussa 2 avattu, Zhangin ym. (2004) esittelemä ongelma oppimateriaalin yksinkertaisuudesta voitaisiin välttää käyttämällä tällaista ympäristöä, joka käyttää erityyppisiä materiaaleja erityyppisille oppijoille. Samalla tämä mukautuvuus myös mahdollistaa ympäristön kehittymisen, mikä taas voisi olla apuna toiselle verkko-oppimisympäristöjen ongelmalle: niiden jäykkyydelle ja muuttumattomuudelle. Toisaalta ympäristön vastaavuudet muihin Zhangin ym. (2004) esittämiin haasteisiin, kuten opettajien työmäärän lisäämiseen ja käyttöönoton hankaluuteen, ei vielä voida tarkasti vastata ympäristön alkeellisuuden vuoksi: koska monet tutkimuksista käsittelevät aihetta puhtaasti teoreettiselta pohjalta, eikä kattavaa tutkimusta tällaisen ympäristön käytöstä ole, ei vielä pystytä sanomaan, miten se käytännössä voisi vaikuttaa näihin haasteisiin.

Opettajien suhteen toimintaa voidaan ohjata paremmin käyttämällä oppimisen analytiikkaa: tunnistamalla kaikkien oppilaiden joukosta juuri ne oppilaat, jotka tarvitsevat apua, ja vastaavasti tunnistamalla tehokkaimmat opetusmenetelmät on opetustoiminta paljon tehokkaampaa ja vähemmän aikaa vievää (Siemens ym., 2011). Opettajat voivat siis oppilailta kerätyn tiedon pohjalta muokata opetustaan paremmin ymmärrettävään muotoon, ja parantaa näin oppimateriaalin omaksumista ja siten myös oppimiskokemusta (Daniel, 2015). Kun prosessia vertaa perinteisiin opetuksen kehitysmenetelmiin, on prosessi paljon tehokkaampi: perinteisin metodein toteutettu oppijoiden arviointi ja analyysi on yleensä koostunut esimerkiksi kurssipalautteesta, mikä taas on kärsinyt annetun palautteen rajallisesta määrästä ja laadusta. Palautteen pohjalta tehdyt toimenpiteet taas ovat usein kärsineet suuresta viiveestä palautteen saannin ja tehtyjen toimenpiteiden välillä. (Elias, 2011.) Vahvuutena perinteiseen palautteen keruuseen oppimisen analytiikalla kerätty data onkin määrältään paljon suurempaa, sillä se kerääntyy jatkuvasti oppilaiden käyttäessä oppimisympäristöä.

Yleensä ottaen oppimisen analytiikka tukeutuu erityyppisten oppijoiden profilointiin luodakseen mukautuvia ja personoituja oppimisympäristöjä, jotka pystyvät ottamaan huomioon eriävät tarpeet, ja luomaan räätälöityjä oppimiskokemuksia, jotka sopivat näihin eri tarpeisiin (Chatti ym., 2013). Samalla kun ympäristön käyttöä pystytään paitsi valvomaan, myös visualisoimaan, korostuu entisestään ympäristön jatkuva kehitys: ympäristöä

valvottaessa voidaan paljastaa malleja oppilaiden käyttäytymisestä, mikä antaa tukea päätöksentekoon ympäristöä koskien (Chatti ym., 2013), ja näin ollen oppiminen on jatkuvasti kehittyvää ja mukautuvaa, ja osaa ottaa huomioon eri oppimisen tyylit. Tehokkaasti käytettynä big data voikin auttaa instituutioita tehostamaan näitä oppimiskokemuksia, edistämään oppilaiden suorituksia, vähentämään koulun keskeytyksiä sekä lisäämään valmistuneiden määrää (Daniel, 2015). Vaikka big data ja oppimisen analytiikka eivät kuitenkaan ole ihmelääke kaikkiin opiskelua koskeviin ongelmiin ja päätöksiin, on toivoa siitä, että niistä voitaisiin muodostaa ratkaisuja oppimiseen ja hiljattain integroida hallinnollisiin ja opetuksellisiin toimiin (Picciano, 2014). Tässä tutkielmassa esiteltiin kolme tapaa integroida big dataa opetukseen ja oppimiseen, ja jokaisesta niistä koituu samankaltaisia hyötyjä.

5 YHTEENVETO

Tutkielmassa pyrittiin ottamaan selvää, kuinka big dataa voidaan hyödyntää osana verkko-oppimisympäristöjä ja millaisia hyötyjä niistä koituisi. Vastauksena tähän löydettiin kolme päätapaa, joilla big dataa voidaan hyödyntää osana verkko-oppimisympäristöjä: datan visualisaatio ja ennusteet, suositusjärjestelmät ja oppimisympäristön personointi. Vaikka nämä teknologiat ovat keskenään erilaisia, on niillä paljon yhteistä: kaikki ne käyttävät hyväkseen ympäristön käytöstä kerättyä dataa parantaakseen toimintaansa ja kehittyäkseen. Tähän tarjottiin monia eri tapoja: visualisointi ja ennusteet voivat mm. paljastaa oppilaat, joilla on vaikeuksia, suositusjärjestelmät osaavat tarjota parhaiten yksittäisille opiskelijoille sopivia materiaaleja ja personoidut oppimisympäristöt pystyvät mukautumaan oppilaiden taitojen mukaisiksi, ja pystyvät samalla tehdä suosituksia opetuksen suhteen ja valvoa oppilaan ympäristön käyttöä.

Hyötyjen suhteen mallin suurin rikkaus on itse datassa, jota voidaan hyödyntää mm. oppilaitosten hallinnossa tukemaan päätöksentekoa, opettajilla parantamaan opetusta ja oppilailla parantamaan oppimista. Verkko-oppimisen perinteisten haasteiden suhteen suurimmat edistysaskeleet ovat puolestaan ympäristöjen monipuolisessa käytössä: siinä missä ennen ympäristöt olivat jäykkiä, kehittymättömiä ja muuttumattomia, nyt big data antaa työkalut ympäristöjen kehittämiseen eteenpäin keräämällä jatkuvasti tietoa siitä, miten ympäristöä käytetään. Kuitenkin muihin haasteisiin vastaaminen vielä tässä vaiheessa on hankalaa: koska ympäristöjen käyttö on niin kokeellisella asteella, ei vielä pystytä ottamaan kantaa siihen, miten ne voivat vaikuttaa opettajan työmäärään tai käyttöönoton vaikeuteen. Tässä osin tarvitaankin lisätutkimusta. Tutkielma ei myöskään ottanut kantaa itse big datan integroimisesta syntyviin ongelmiin, vaan keskittyi sen mahdollisuuksiin. Tässäkin osin lisätutkimukset olisivat tarpeen jotta ymmärrettäisiin, mitä käytännössä vaatisi käyttää big dataa osana verkko-oppimisympäristöjä.

LÄHTEET

- Abdelouarit, K. A., Sbihi B. & Aknin, N. (2015). Big-Learn: Towards a Tool Based on Big Data to Improve Research in an E-Learning Environment. *International Journal of Advanced Computer Science and Applications*, 6(10), pp. 59-63.
- Alonso, F., López, G., Manrique, D. & Viñes, J. (2005). An instructional model for web-based e-learning education with a blended learning process approach. *British Journal of Educational Technology*, 36(2), pp. 217-235.
- Aparicio, M., Bacao, F. & Oliveira, T. (2017). Grit in the path to e-learning success. *Computers in Human Behavior*, 66, pp. 388-399.
- Benhamdi, S., Babouri, A. & Chiky, R. (2017). Personalized Recommender System for e-Learning Environment. *Education and Information Technologies*, 22(4), pp. 1455-1477.
- Bobadilla, J., Ortega, F., Hernando, A. & Gutierrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems*, 46, p. 109.
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U. & Thüs, H. (2013). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5 6), p. 318.
- Chen, H., Chiang, R. H. L. & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), p. 1165.
- Chen, M., Mao, S. & Liu, Y. (2014). Big Data: A Survey. *Mobile Networks and Applications*, 19(2), pp. 171-209.
- Cigdem, H. & Topcu, A. (2015). Predictors of instructors' behavioral intention to use learning management system: A Turkish vocational college example. *Computers in Human Behavior*, 52, pp. 22-28.
- Cukier K. (2010, Helmikuu 25), Data, data everywhere: A special report on managing information. *The Economist*. Haettu osoitteesta <http://www.economist.com/node/15557443>
- Daniel, B. (2015). Big Data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46(5), pp. 904-920.
- Duval, E. (2011). *Attention please! Learning analytics for visualization and recommendation*. In Proceedings of the 1st International Conference on Learning Analytics and Knowledge, LAK '11 (pp. 9-17). New York, NY: ACM. doi:10.1145/2090116.2090118
- Dwivedi, S. & Roshni K. (2017). Recommender system for big data in education. *E-Learning & E-Learning Technologies (ELELTECH), 2017 5th National Conference on*, pp. 1-4.
- Elias, T. (2011). Learning Analytics: Definitions, Processes and Potential. Haettu 1.2.2018 osoitteesta <https://pdfs.semanticscholar.org/732e/452659685fe3950b0e515a28ce89d9c5592a.pdf>

- Forouzesh, M. (2012). Characteristics of learning management system (LMS) and its role in education of electronics. *The 8th International Scientific Conference eLearning and software for Education*, Bucharest, April 26-27, 2012
- Fournier, H., Kop, R. & Sitlia, H. (2011). *The value of learning analytics to networked learning on a personal learning environment*. First International Conference on Learning Analytics and Knowledge 2011, Banff, Alberta, Canada, February 27–March 1, 2011.
- Friesen, N. (2009). *Re-thinking e-learning research: Foundations, methods and practices*. New York: Peter Lang.
- Gandomi, A. & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), pp. 137-144.
- Gell-Mann, M. (1996). A commentary to R Schank. In J. Brockman (Ed.), *The third culture: beyond the scientific revolution*. New York: Touchtone Books, 167-180.
- Godwin-Jones, R. (2017). Scaling up and zooming in: Big data and personalization in language learning. *Language Learning & Technology*, 21(1), 4-15.
- Hijon R. and Carlos, R. (2006). E-learning platforms analysis and development of students tracking functionality. *Proceedings of the 18th World Conference on Educational Multimedia, Hypermedia & Telecommunications*, pp. 2823-2828
- Johnson, L., Smith, R., Willis, H., Levine, A. & Haywood, K. (2011). *The 2011 Horizon Report*. Austin, Texas: The New Media Consortium.
- Johnson, R. D., Hornik, S., & Salas, E. (2008). An empirical examination of factors contributing to the creation of successful e-learning environments. *International Journal of Human - Computer Studies*, 66(5), pp. 356-369.
- Kahiigi, E. K., Ekenberg, L., Hansson, H., Tusubira, F.F., & Danielson, M. (2008). Exploring the e-learning state of the art. *The Electronic Journal of e-Learning*, 6 (2), 77-88.
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, 34(3), 387-394.
- Lee, D. S., Lee, K. C. & Choi, D. Y. (2015). An analysis of shared leadership, diversity, and team creativity in an e-learning environment. *Computers in Human Behavior*, 42, pp. 47-56.
- Lu, J., Wu, D., Mao, M., Wang, W. & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems*, 74, p. 12.
- Marr, B. (2017), How is Big Data Used in Practice? 10 Use Cases Everyone Must Read. Haettu 10.12.2017 osoitteesta <https://www.bernardmarr.com/default.asp?contentID=1076>
- Moore, J. L., Dickson-Deane, C. & Galyen, K. (2011). e-Learning, Online Learning, and Distance Learning Environments: Are They the Same? *Internet and Higher Education*, 14(2), pp. 129-135.

- Navimipour, N. J. & Zareie B. (2015). A model for assessing the impact of e-learning systems on employees' satisfaction. *Computers in Human Behavior*, 53, p. 475.
- Picciano, A. G. (2014). Big Data and Learning Analytics in Blended Learning Environments: Benefits and Concerns. *International Journal of Interactive Multimedia and Artificial Intelligence*, 2(7), pp. 35-43.
- Plummer, L., (2017). This is how Netflix's top-secret recommendation system works. Haettu 30.12. osoitteesta <http://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like>
- Sangra, A., Vlachopoulos, D. & Cabrera, N., (2012). Building an Inclusive Definition of E-Learning: An Approach to the Conceptual Framework. *International Review of Research in Open and Distance Learning*, 13(2), pp. 145-159.
- Schumacher, C. & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior*, 78, pp. 397-407.
- Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), pp. 1380-1400.
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., ym. (2011). Open learning analytics: An integrated & modularized platform. Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques. Society for Learning Analytics Research
- Sin, K., Muthu, L. (2015). Application of big data in education data mining and learning - a literature review. *ICTACT Journal on Soft Computing*, 5(4), pp. 1035-1049
- Sjödén, B. (2017). What teachers should ask of educational software: identifying the integral data values. *10th annual International Conference of Education, Research and Innovation (ICERI2017), Seville, Spain, 16-18 November, 2017*
- Smith, M. (2004). Portals: Toward an application framework for interoperability. *Communications of the ACM*, 47(10), pp. 93-97.
- Stoffregen, J., Pawlowski, J. M. & Pirkkalainen, H., (2015). A Barrier Framework for open E-Learning in public administrations. *Computers in Human Behavior*, 51, p. 674.
- Udupi, P. K., Sharma, N. & Jha, S. K. (2016). Educational data mining and big data framework for e-learning environment. *Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2016 5th International Conference on*, pp. 258-261.
- Xu, D., Huang, W., Wang, H. & Heales, J., (2014). Enhancing e-learning effectiveness using an intelligent agent-supported personalized virtual learning environment: An empirical investigation. *Information & Management*, 51(4), pp. 430-440.
- Zaiane, O. (2002). Building a recommender agent for e-learning systems. *Computers in Education, 2002. Proceedings. International Conference on Computers in Education* , pp. 55-59.

- Zhang, D., Zhao, J. L., Zhou, L. & Nunamaker, J. F. (2004). Can e-learning replace classroom learning? *Communications of the ACM*, 47(5), pp. 75-79.
- Zhang, H., Almeroth, K., Knight, A., Bulger, M., and Mayer, R. (2007). Moodog: Tracking Students' Online Learning Activities in C. Montgomerie and J. Seale (Eds.). *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Chesapeake, VA: AACE, 2007, pp. 4415-4422.