

Avoim data Lahden ammattikorkeakoulussa

LAHDEN
AMMATTIKORKEAKOULU
Liiketalouden laitos
Tietojenkäsittelyn koulutusohjelma
Opinnäytetyö
Kevät 2016
Anni Manninen

Lahden ammattikorkeakoulu
Tietojenkäsittelyn koulutusohjelma

MANNINEN, ANNI:

Avoin data Lahden
ammattikorkeakoulussa

Tietojenkäsittelyn opinnäytetyö, 35 sivua

Kevät 2016

TIIVISTELMÄ

Tässä opinnäytetyössä suunnitellaan prosessi avoimen tiedon julkaisulle Lahden ammattikorkeakoulussa. Lahden ammattikorkeakoulussa ei ole suunnitelmaa rajapintojen avaamiseksi, joten tutkimukselle aiheesta on tarvetta alusta saakka. Opinnäytetyössä keskitytään taustatekijöihin avoimen tiedon julkaisun taustalla. Opinnäytetyössä ei suunnitella verkkopalvelua lomakkeineen rajapinta-avaimien luovuttamiseksi ja rajapintojen tai muun datan avulla tehtyjen sovelluksien tarkasteluun.

Opinnäytetyö on kvalitatiivinen tutkimus jossa hyödynnetään deduktiivista päättelyä ja suunnittelutiedettä. Opinnäytetyön aluksi käydään läpi tutkimustehtävä ja tutkimuksen metodologia jonka jälkeen teoriaosassa esitellään avoin data-termi ja siihen liittyvät tekijät. Luvussa viisi käsitellään itse tutkimuksen suorittaminen. Opinnäytetyön lopuksi pohditaan tutkimuksen onnistumista sekä validiteettia ja reliabiliteettia. Lopussa annetaan myös ehdotuksia jatkotutkimukselle. Opinnäytetyön teoriaosaan sekä suunnittelutieteelliseen viitekehykseen perustuen muodostetaan prosessi avoimen tiedon julkaisulle.

Aiempien tutkimuksien ja julkaisujen perusteella luotiin prosessi datan julkaisun taustalle suunnittelutieteellisen viitekehyksen ohjatessa prosessia. Suunnitellun prosessin avulla avoimen datan julkaisu Lahden ammattikorkeakoulussa on helppoa toteuttaa tulevaisuudessa kun mahdollisia julkaistavia aineistoja tulee lisää.

Asiasanat: avoin tieto, julkinen tieto, avoin rajapinta

Lahti University of Applied Sciences
Degree Programme in Business Information Technology

MANNINEN, ANNI:

Open data at Lahti University of
Applied Sciences

Bachelor's Thesis in Information Technology, 35 pages

ABSTRACT

The aim of this thesis was to create a process to publish open data at Lahti University of Applied Sciences. Firstly, the data will be published via application programming interfaces (APIs) of a system called Peppi, but later some data will be published through other systems too. The thesis process did not include the implementation of a web service for new users to send a request for API keys and browse applications or web services that have already been done with the APIs or other open data material. There was no plan at Lahti University of Applied Sciences on how to publish these data, and therefore there was a need to understand the underlying process.

This thesis is a qualitative study based on deductive reasoning and design science. The second and third chapter discuss and introduce the study and the related research methodology. The fourth chapter is about the theoretical background of thesis where the term open data is explained. Chapter five explains how the study was carried out. Finally, results and the conclusion of the study are discussed. The validity and reliability of the study and ideas for further study are also explained.

The process was created based on a theoretical framework drawn on design science. Based on this process, it is easier to publish open data at Lahti UAS in the future.

Keywords: open data, open government data, open API

SISÄLLYS

1	JOHDANTO	1
2	TUTKIMUSTEHTÄVÄ	2
2.1	Tutkimuskysymys ja tavoite	2
2.2	Tutkimusongelma ja rajaus	3
2.3	Teoreettinen viitekehys	4
2.4	Tutkimusmenetelmät	5
3	TUTKIMUKSEN METODOLOGIA	6
3.1	Suunnittelutiede	6
3.2	Artefaktit ja niiden luokittelu	7
3.3	Suunnittelutieteellinen prosessi	8
4	AVOIN TIETO	11
4.1	Mitä tieto on?	12
4.2	Avoimen datan hyödyt	14
4.3	Avoimen datan käyttökohteet	15
4.4	Avoimen datan periaatteet	16
4.5	Avoimen datan julkaisuprosessi	18
4.6	Tekninen toteutus	21
5	TUTKIMUKSEN SUORITTAMINEN	23
5.1	Avoim data Lahden ammattikorkeakoulussa	23
5.2	Peppi-järjestelmä	23
5.3	Tutkimusprosessi	24
6	TUTKIMUSTULOKSET	26
7	YHTEENVETO	29
7.1	Avoim LAMK-palvelu	30
7.2	Jatkotutkimus ja yleistettävyys	31
7.3	Validiteetti ja reliabiliteetti	31
	LÄHTEET	33

1 JOHDANTO

Tietoa ei enää pidetä vain organisaation yksinoikeutena vaan yhä enemmän ilmaiseksi jaettavissa olevana, eri sidosryhmille hyödyllisenä aineistona. Viime vuosina ajatusmaailma tiukasti varjellusta datasta on asteittain siirtynyt kohti kaikille avointa, helposti hyödynnettävää tietoa joka on vapaasti käytettävissä, muokattavissa sekä jaettavissa tulevaisuuden teknologiaa sekä palveluita suunniteltaessa.

Ensimmäisenä suurena valtiona Yhdysvallat avasi portaalin aineistojensa julkiseen käyttöön vuonna 2009, jonka jälkeen monet valtiot sekä organisaatiot ovat seuranneet perässä (Kitchin 2014, 49). Suomessa esimerkiksi Tampereen ammattikorkeakoulu ja Metropolia-ammattikorkeakoulu ovat julkaisseet avoimia rajapintoja tietojärjestelmistään (Metropolia 2013; Tampereen ammattikorkeakoulu 2016). Tämän yleisen trendin mukaan Lahden ammattikorkeakoulussa suunnitellaan siirtymistä tiedon avoimuuteen tutkimuksen, kehityksen ja uusien innovaatioiden turvaamiseksi. Julkaisemalla avointa dataa esimerkiksi uusien sovelluksien kehittäminen monipuolistuu tiedon ollessa vapaasti saatavilla. Alkuvaiheessa julkaisu koskee vain opetuksen suunnittelujärjestelmä Pepin avoimia ohjelmointirajapintoja, mutta myöhemmässä vaiheessa dataa on tarkoitus julkaista myös muista lähteistä. Koska kyseessä on korkeakoulu, voidaan myös olettaa yritys- ja muiden sidosryhmien olevan kiinnostuneita avoimista rajapinnoista ja niiden mahdollisuuksista opiskelijayhteistyön tiivistäjänä.

Tässä opinnäytetyössä perehdytään avoimen tiedon määritelmään sekä suunnitellaan prosessi tiedon julkistamiselle. Idea opinnäytetyöhön lähti Lahden ammattikorkeakoulun toimeksiannosta sekä opinnäytetyön kirjoittajan omasta mielenkiinnosta aihetta kohtaan. Opinnäytetyön aihe on varsin ajankohtainen tiedon määrän lisääntyessä jatkuvasti sekä ilmapiirin tiedon jakamisen suhteen muuttuessa koko ajan avoimemmaksi.

2 TUTKIMUSTEHTÄVÄ

Tässä opinnäytetyössä tutustutaan avoimen datan periaatteisiin, tutkimuksessa käytettyyn metodologiaan, käydään läpi tutkimuksen kannalta olennaisia termejä sekä tutkitun materiaalin perusteella suunnitellaan käytäntö avoimen datan julkaisemiseen Lahden ammattikorkeakoulussa. Opinnäytetyössä suunnitellaan prosessi tiedon avaamisen taustalle sekä sivutaan hieman avoimelle datalle suunniteltavaa verkkopalvelua sekä sitä, mihin avointa dataa Lahden ammattikorkeakoulussa voidaan hyödyntää. Opinnäytetyötä tehdessä keväällä 2016 rajapintojen julkaisu koskee vain Peppi-järjestelmää, mutta myöhemmässä vaiheessa myös muista lähteistä julkaistaan dataa.

Toisessa ja kolmannessa luvussa tutustutaan tutkimustehtävään menetelmiseen sekä esitellään teoreettinen viitekehys tutkimuksen taustalla. Neljännessä luvussa käydään läpi taustaa avoimeen tietoon liittyen, jonka jälkeen avataan lukijalle keinoja tutkimuksen suorittamiselle. Varsinaiset tutkimustulokset esitellään luvussa kuusi. Opinnäytetyön lopuksi pohditaan tutkimuksen onnistumista sekä jatkotutkimusmahdollisuuksia ja tutkitaan reliabiliteettia ja validiteettia tutkimustuloksissa.

Opinnäytetyössä dataan viitataan joko sanalla data (monikko sanasta *datum*) tai vaihtoehtoisesti tieto, sillä tieto on käypä suomennus englannin kielen sanalle data. Tarkempi määritelmä datalle ja siihen liittyville termeille sekä tiedon arvoketju käsitellään kappaleessa 4.1.

2.1 Tutkimuskysymys ja tavoite

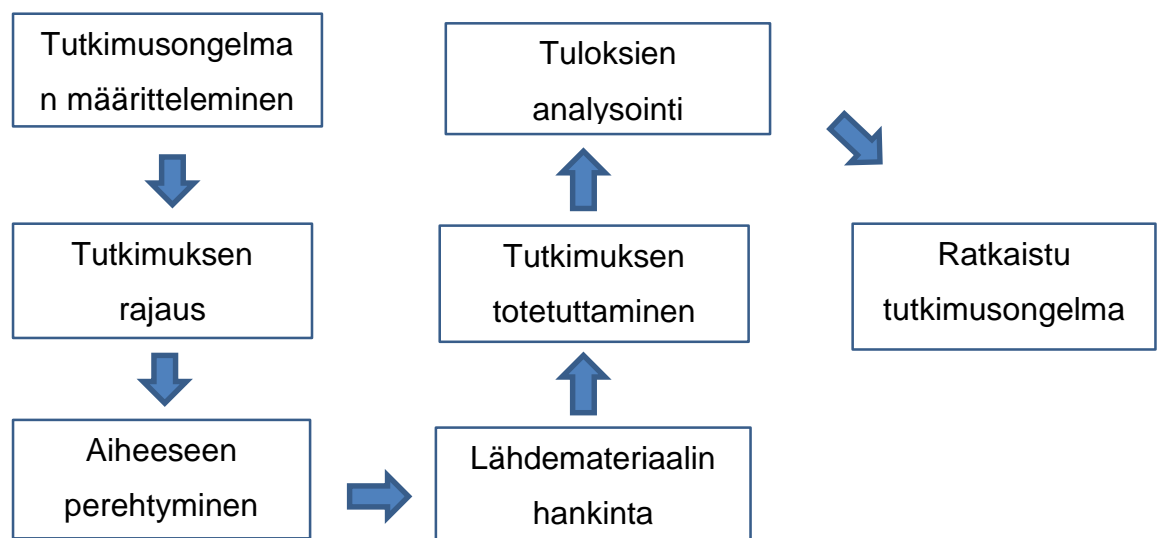
Opinnäytetyöprosessin aikana pyritään suunnittelemaan toimiva malli sille, kuinka tietoa julkistetaan ulkopuolisille. Tavoitteena on, että opinnäytetyöprosessin lopuksi Lahden ammattikorkeakoululla on selkeä visio siitä, millainen prosessi datan julkistaminen on. Tutkimuskysymys on seuraava:

- Kuinka avoimen datan julkaiseminen Lahden ammattikorkeakoulussa tulisi toteuttaa?

Lahden ammattikorkeakoulu on profiloitunut ympäristö- ja energiatietoisena korkeakouluna (Turunen 2016). Tällä hetkellä meneillään on sääasemaprojekti Helsingin yliopiston kanssa, ja tästä sääasemasta saatavaa dataa haluttaisiin saada julkaistua myöhemmässä vaiheessa. Vuonna 2017 suunnitelmissa on saada julkaistavaa tietoa Tekniikan laitoksen konesalista ja palvelintietokoneiden energiankäytöstä. Tulevaisuudessa näitä tietoja halutaan julkaista Peppi-järjestelmästä julkaistun aineiston rinnalla (Turunen 2016).

2.2 Tutkimusongelma ja rajaus

Opinnäytetyön kirjoittamishetkellä keväällä 2016 Lahden ammattikorkeakoulussa ei ole suunnitelmaa aineistojen julkaisemisen suhteen. Tiedossa on, että Peppi-järjestelmästä tullaan julkaisemaan avoimia rajapintoja, ja että tulevaisuudessa myös muuta avointa dataa halutaan tuoda julki. Opinnäytetyössä tutkimusongelma pyritään ratkaisemaan seuraavan kuvion mukaisesti:



KUVIO 1 Tutkimusongelman ratkaisu

Tässä opinnäytetyössä tutustutaan open data-käsitteeseen sekä suunnitellaan prosessi tiedon julkistamisen suhteen. Opinnäytetyössä ei suunnitella alustaa julkaistavalle tiedolle. Avoimeen tietoon läheisesti liittyvät termit määritelmien jätetään käsittelemättä opinnäytetyön aiheen rajaamisen helpottamiseksi. Näitä käsitteitä ovat muun muassa big data, Internet of Things (esineiden internet) sekä semanttinen verkko.

2.3 Teoreettinen viitekehys

Tutkimuksen viitekehystä voidaan pitää tutkimuksen teoreettisena osana, jossa tuodaan esille tutkimuksen keskeiset käsitteet sekä suhteet näiden käsitteiden välillä. Viitekehukseen kuuluvat myös tutkimuksen metodologia ja se, mitä tutkittavasta aiheesta jo tiedetään (Tuomi & Sarajärvi 2009, 18-19).

Tämän opinnäytetyön teoreettinen viitekehys koostuu avoimen datan määritelmästä sekä aiemmista tutkimuksista informaatiojärjestelmien viitekehysten mukaan (KUVIO 2.) Tutkimusprosessin alkuvaiheessa tutustutaan avoimen datan käsitteeseen sekä haetaan aiempia tutkimuksia ja julkaisuja kuvion 2 oikean puolen eli *tietämiskannan* (Knowledge Base) mukaisesti. Prosessin aikana huomioidaan *ympäristö* (Environment), joka muodostuu Lahden ammattikorkeakoulun henkilöstöstä sekä opiskelijoista, jotka tulevat hyötymään suunnitellusta artefaktista. *Ympäristöön* otetaan mukaan myös organisaation jo olemassa olevat järjestelmät sekä tekninen infrastruktuuri. Kuvion 2 keskellä suoritetaan itse tutkimus, joka tämän opinnäytetyön tapauksessa on avoimen datan julkaisun taustalle suunniteltava prosessi. Tämä prosessi on suunnittelutieteellinen artefakti, jonka rakentaminen esitellään tutkimusprosessin jälkeen

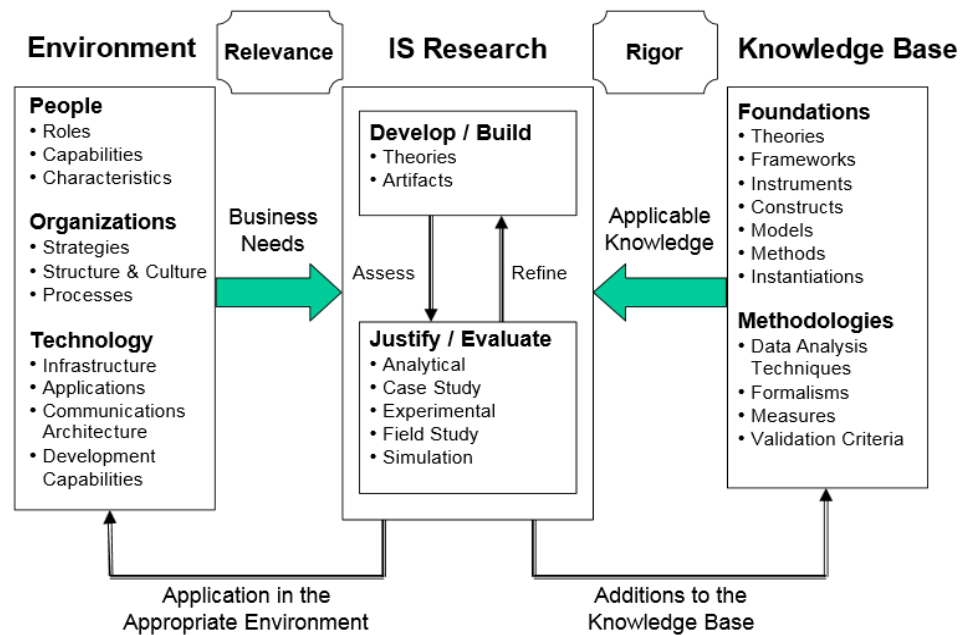


Figure 2: Information Systems Research Framework

KUVIO 2 Hevnerin ym. määrittelemä informaatiotieteiden viitekehys

2.4 Tutkimusmenetelmät

Tutkimus on kvalitatiivinen eli laadullinen tutkimus jossa hyödynnetään deduktiivista päättelyä ja suunnittelutiedettä.

Tuomi & Sarajärvi (2009, 95;97) määrittelevät deduktiivisen logiikan olevan teorialähtöinen, luonnontieteellisen tutkimuksen yleinen analyysimalli joka perustuu jo olemassa olevaan teoriaan tai malliin. Tällöin tutkimuksen kohteena oleva ilmiö määritellään jo olemassa olevan tiedon mukaan. Opinnäytetyön teoriaosuus koostuu avoimeen dataan liittyvistä tutkimuksista ja julkaisuista. Nämä aineistot ohjaavat tutkimuksen kulkua deduktiivisen lähestymistavan mukaisesti. Käytettyihin metodeihin päädyttiin, sillä opinnäytetyön aikana pyritään jo olemassa olevien keinojen ja materiaalin pohjalta suunnittelemaan uusi toimintamalli, joka tämän opinnäytetyön tapauksessa on prosessi avoimen tiedon julkaisulle.

3 TUTKIMUKSEN METODOLOGIA

Puhuttaessa tutkimuksen metodologiasta, tarkoitetaan yleistä lähestymistapaa tutkimusaiheeseen (Metsämuuronen 2006, 83.) Peffers ym. (2007) määrittelevät suunnittelutieteen metodologian muodostuvan suunnittelutieteellisen tutkimuksen määritelmästä, prosessista artefaktin toteutuksen taustalla sekä tutkimuksen varsinaisesta toteuttamisesta ja esittämisestä. Tuomi & Sarajarvi (2009, 13) pitävät metodologiaa sääntöinä ja metodeina siitä, kuinka tutkimuksen päämäärään päästään.

3.1 Suunnittelutiede

Suunnittelutiede (engl. design science) on erityisesti tietojärjestelmien kehityksessä sekä insinööritieteissä käytetty lähestymistapa, jonka tarkoituksena on luoda malleja, metodeja ja järjestelmiä eli artefakteja, joiden avulla voidaan helpottaa ihmisten elämää. Siinä missä luonnontieteellisessä tutkimuksessa pohditaan syitä asioiden taustalla, pyritään suunnittelutieteen avulla suunnittelemaan ennalta määrätty artefakti. (March & Smith 1995, 253.)

Järvisen ja Järvisen (2004, 103) mukaan suunnittelutieteellinen tutkimus vastaa esimerkiksi seuraaviin kysymyksiin: ”Voimmeko rakentaa tietyn innovaation ja kuinka hyödyllinen on joku innovaatio?” tai ”Millainen tietyn innovaation tulisi olla ja kuinka meidän tulisi se rakentaa?” Järviset (2004,12;107) viittaavat artefaktiin termillä *innovaatio*. Järviset esittävät innovaation mahdollistavan laajemman katsannon aiheesta, sillä siihen voidaan sisällyttää niin tekniset, sosiaaliset, tiedolliset ja teoreettiset innovaatiot yhdistelmiseen. Myös livari (2007, 51) muistuttaa, etteivät kaikki suunnittelutieteelliset artefaktit ole tietojärjestelmiä. March & Smith (1995) taas pitävät artefaktia enemmän teknisenä toteutuksena.

Suunnittelutieteellisen tutkimuksen keskiössä olevan artefaktin määritelmä jaotteluineen käsitellään seuraavaksi.

3.2 Artefaktit ja niiden luokittelu

Johannesson ja Perjons (2014, 1) määrittelevät artefaktin olevan ihmisen kehittämä luomus, joka on kehitetty ratkaisemaan jokin tietty käytännön ongelma. Artefakti voi olla esimerkiksi tietojärjestelmä, osa lähdekoodia tai polkupyörä. Kaikkia artefakteja yhdistää päämäärä ratkaista ongelma, joka voi olla hyvinkin arkipäiväinen. Tämän tutkimuksen artefaktina toimii datan ja rajapintojen julkaisuun suunniteltava prosessi. Koska kuten Johannesson ja Perjons (2014, 3) määrittelevät, artefakti voi olla paitsi konkreettinen asia kuten auto, se voi myös olla suunniteltava menetelmä tai ohjeistus. Täten tutkimuksen kohteena suunniteltava tiedon julkaisun prosessi sopii suunnittelutieteellisen artefaktin määritelmään.

March & Smith (1995, 253) lähestyvät suunnittelutieteellisiä tutkimuksia kahdesta näkökulmasta, rakentamisesta (engl. build) ja arvioinnista (engl. evaluate). Rakentaminen viittaa artefaktin suunnitteluun ja arviointi toteutuneen artefaktin toiminnallisuuteen. *Rakentaessa*, määritellään artefaktille tietty tehtävä, jolloin tilannetta voidaan lähestyä kysymyksellä: ”Toimiiko se?” Artefaktin ollessa valmis, osoitetaan suunnitteluongelman olevan ratkaistu. Arvioitaessa suunniteltua artefaktia voidaan kysyä: ”Kuinka hyvin se toimii?” Tällöin verrataan artefaktin suoriutumista sille asetettujen kriteerien perusteella. Hevner ym. (2004) pitävät suunnittelutieteellistä tutkimusta joko kokonaan uuden, innovatiivisen artefaktin suunnitteluna tai jo olemassa olevan artefaktin merkittävänä parantamisena.

Tässä opinnäytetyössä keskitytään artefaktin rakentamiseen, sillä tarkoituksena ei ole testata toteutettavan artefaktin toiminnallisuutta ympäristössään. Rakennettavana artefaktina toimii avoimen tiedon julkaisun taustalle suunniteltava prosessi.

Suunniteltavat artefaktit voidaan jakaa neljään eri tyyppiin, kuten Johannesson ja Perjons (2014, 29) esittävät. Seuraavaksi käydään läpi nämä neljä tyyppiä, joille Järviset (2004, 103) ovat antaneet seuraavat suomennokset: konstruktio, malli, metodi ja toteutus.

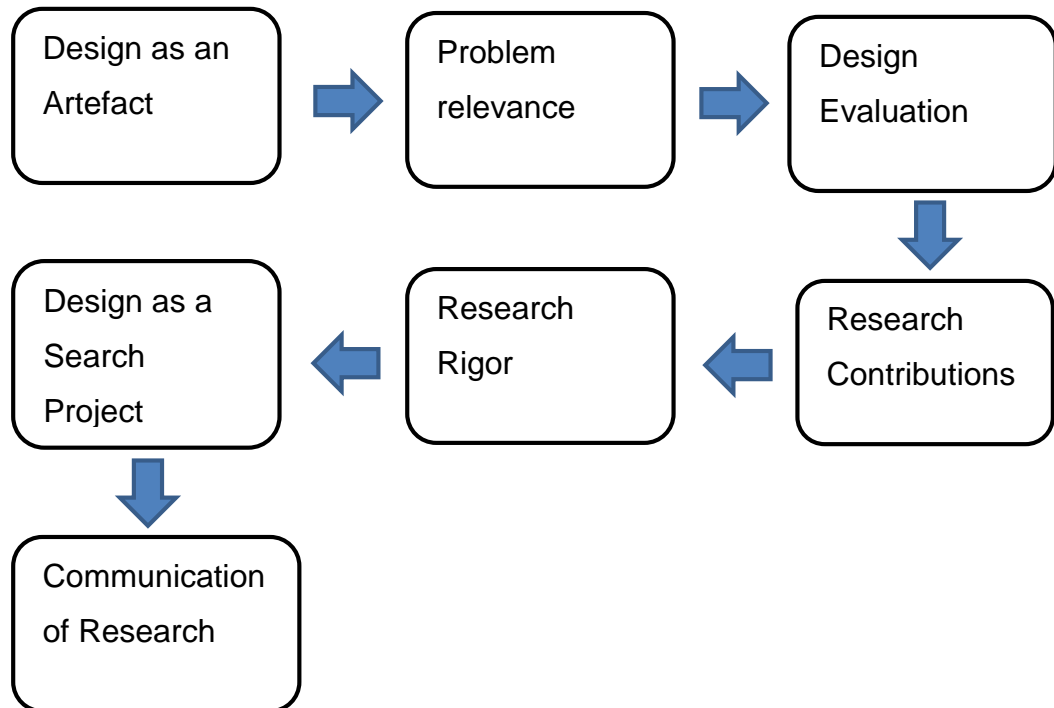
Konstruktiot (engl. constructor) määrittävät tutkimusaiheen sanaston (March ym. 1995, 256.) Ne jakavat tutkimuksen kohteena olevan ilmiön pieniin osiin, kuten ohjelmointikielen metodeiksi, mikä helpottaa tutkittavan ilmiön ymmärtämistä. Konstruktioiden tehtävänä ei ole ilmiön laaja analysointi tai määritelmien tekeminen (Johannesson ja Perjons 2014).

Malleja (engl. model) käytetään tukemaan muiden artefaktien rakennetta. Mallien avulla pyritään hakemaan ratkaisuja käytännön ongelmiin jakamalla ilmiö osiin. Metodit (engl. method) määräävät suositukset ja ongelmanratkaisun prosessin tavoitteeseen pääsyn taustalla. Järvisen (2004, 114) mukaan metodin voi tuottaa tutkimuksen ainoana tuloksena tai osana suurempaa prosessia jonka loppuun saattamiseksi tarvitaan useita askelia eli metodeja. Tällöin askeleita (metodeja) seuraten tulisi päästä tavoitetilään eli esimerkiksi uuden toimivan järjestelmän käyttöönottoon (Johannesson & Perjons 2014). Toteutukset (engl. instantiations) taas ovat esimerkiksi valmiita järjestelmiä tai toteutuksia, joihin konstruktiolla, malleilla ja metodeilla pyritään (Hevner ym. 2004, 79.)

Opinnäytetyössä suunniteltavaa artefaktia voidaan pitää metodina, sillä kuten Järviset (2004, 107) määrittelevät: "Metodi on joukko askelia (algoritmi tai ohjeisto), joita käytetään suorittamaan tehtävä." Tässä tutkimuksessa pyritään luomaan metodi sille, kuinka tiedon julkistaminen tulisi Lahden ammattikorkeakoulussa toteuttaa. Tämän metodin pohjalta suunnitellun prosessin avulla aineistojen julkaisu on helppoa toteuttaa.

3.3 Suunnittelutieteellinen prosessi

Suunnittelutieteellisen artefaktin suunnittelun, toteuttamisen ja arvioinnin taustalle Hevner ym. (2004, 83) ovat laatineet seitsemän ohjetta kuvion 3 mukaan. Tässä kappaleessa käydään läpi tämä prosessi sekä avataan hieman kuvion 3 sisältöä.



KUVIO 3 Artefaktin suunnitteluprosessi

Ensinnäkin, ”*Suunnittele artefaktia*” (Järviset 2004, 115.)

Suunnittelutieteellisen tutkimuksen lopputulemana tulisi olla jokin tarkoituksellinen artefakti, jonka tarkasta määrittelystä tutkimuksen tulee lähteä.

Toiseksi, tutkimuksen tavoitteena on teknologialähtöisen ratkaisun löytäminen ratkaisemattomaan ja tärkeään liiketoiminnan ongelmaan. ”*Painota suunnittelussa liiketoimintaongelman relevanttiutta.*” (Järviset 2004, 115.) Relevanttius tuodaan esille ratkaistaessa suunnittelutieteellinen ongelma (Hevner ym. 2004; Järvinen & Järvinen 2004, 115).

Kolmanneksi, ”*Osoita artefaktin relevanttius arvioimalla se*” (Järviset 2004, 115.) Hevner ym. määrittelevät artefaktin arvioinnin (arvioinnin) tapahtuvan esimerkiksi sen toiminnallisuutta, valmiusastetta, tarkkuutta, suorituskykyä, käytettävyyttä tai luotettavuutta arvioimalla. Arvioinnin tulee tapahtua liiketoimintaympäristön vaatimuksien mukaan.

”*Tuota tutkimuksella uutta tietoa, uusia menetelmiä tai merkittävä artefakti*” (Järviset 2004, 115.) Neljäntenä ohjeena Hevner y. (2004, 87) esittävät, että suunnittelutieteellisen tutkimuksen tulisi tuottaa kolmenlaisia eri kontribuutioita jotka liittyvät seuraaviin: a) suunniteltu artefakti b) konstruointitietämys eli perusteet ja c) metodologia. Näistä vähintään yhden tulisi löytyä suunnitellusta artefaktista.

Viidentenä ohjeena Hevner ym. (2004, 87-88) painottavat tutkimuksen tieteellistä tarkkuutta, joka voidaan päätellä siitä kuinka tutkimus on toteutettu. Suunnittelutieteellisessä tutkimuksessa sekä artefaktin rakentaminen että arviointi tulee suorittaa tarkkoja metodeja seuraten jotta voidaan saavuttaa tieteellinen tarkkuus.

Kuudentena ohjeena pidetään seuraavaa: Tutkimus ja artefaktin suunnitteluprosessi tulisi nähdä parhaan mahdollisen ratkaisun etsimisenä. Tätä ratkaisua etsitään saatavilla olevien toimenpiteiden ja resurssien eli *keinojen* (engl. means) avulla (Järviset 2004, 115.) Suunnittelutiede on iteratiivista, eli lopputulokseen päästäkseen tutkijan on jaettava ongelma pienempiin osiin ja tehtävä toistoja tarvittaessa (Hevner ym. 2004).

Viimeisenä ohjeena Hevner ym. (2004) painottavat tutkija- ja soveltajayhteisön informoimista uusista tuloksista. Tällöin tulee ottaa huomioon niin teknisesti orientoituneet tahot kuin liiketoiminnasta vastaavat henkilöt. Teknisestä toteutuksesta kiinnostuneille tulee tarjota tarvittavat tiedot artefaktin rakentamiseen organisaation tarpeet huomioon ottaen. Liiketoiminnasta kiinnostuneille taasen on mielekkäämpää tietää prosessi artefaktin toteutuksen taustalla.

4 AVOIN TIETO

Julkisen sektorin avoimet tietovarannot ovat nykypäivän muoti-ilmiö. Suosion voidaan katsoa lähteneen liikkeelle vuodesta 2009, jolloin Yhdysvallat julkaisi portaalin julkishallinnon avoimiin tietovarantoihin. Tätä seurasi pian Ison-Britannian vastaava palvelu jonka jälkeen mielenkiinto muissa valtioissa ja kansainvälisissä organisaatioissa on lisääntynyt huomasti (Janssen 2012).

Esimerkiksi Euroopan Unioni määrittelee avoimen datan yhdeksi avaintekijäksi EU 2020-strategiassa, jonka tarkoituksena on avittaa Unionin talouskasvua. Datan julkistamisella haetaan paitsi tuloja EU-alueelle sekä lisäkiinnostusta politiikkaan, myös uusia innovaatioita ja tiedonvälityksen sujuvuutta eri instituutioiden välille. (Euroopan komissio, 2011).

Myös Suomessa useat kaupungit ja organisaatiot ovat avanneet aineistojaan julkiseen käyttöön. Helsinki Region Infoshare (2016) kerää yhteen pääkaupunkiseudun tapahtumiin liittyviä rajapintoja ja niistä tehtyjä sovelluksia. Vastaavan palvelun ovat avanneet muun muassa Turku, Oulu ja Jyväskylä. Suomen valtiovarainministeriön (2016) mukaan Suomessa tietoa on avattu useista lähteistä muun muassa säähän, ilmastoon, liikenteeseen sekä talouteen liittyen. Merkittävä osa julkishallinnon hallitsemasta datasta on paikkatietoa, eli tietoa jostakin fyysisestä sijainnista alueesta tai osoitteesta (Poikola ym. 2010, 70.)

Avointa dataa ei tule sekoittaa avoimeen lähdekoodiin eikä julkiseen tietoon. Puhuttaessa avoimesta lähdekoodista tarkoitetaan järjestelmiä, joiden lähdekoodi on vapaasti hyödynnettävissä kun taas avoimella datalla nimensä mukaisesti avoimesti saatavilla ja hyödynnettävissä olevaa tietoa (Lindman & Nyman 2014, 2.) Julkinen tieto tarkoittaa tietoa, joka on julkisesti luettavissa esimerkiksi verkkosivulla (Helsinki Region Infoshare 2016.) Tällöin tietoa ei voi jalostaa käyttäjän omiin tarkoituksiin, esimerkiksi mobiilisovelluksien tekemiseen.

Tässä opinnäytetyössä sovelletaan julkishallinnon tietovarantojen julkaisuun painottuvia ohjeita sekä teoriaa, sillä kuten Poikola ym. (2004, 15) toteavat, julkishallinnollisen datan julkaisuun perustuvia periaatteita voidaan hyödyntää muiden organisaatioiden tapauksessa. Teoria avoimen datan taustalla käydään myös läpi julkishallinnollisesta näkökulmasta katsoen.

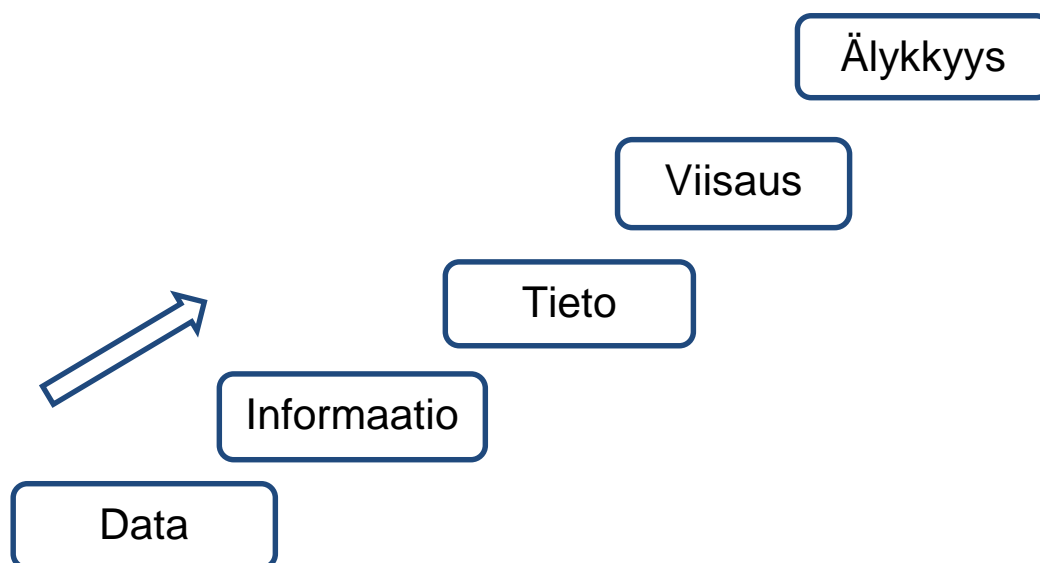
Jotta ymmärrämme paremmin avoimen tiedon käsitettä, sen hyötyjä sekä sitä, mihin avointa dataa hyödynnetään, tässä luvussa käydään läpi teoriaa sekä termejä avoimeen dataan liittyen.

4.1 Mitä tieto on?

Poikolan ym. (2010, 13-14) mukaan suomen kielen termit data, informaatio (engl. information) ja tieto (knowledge) eivät arkipuheessa oikeastaan eroa toisistaan, joten niille on vaikea antaa tiettyä määritelmää. Tällöin datan, informaation sekä tiedon mielletään usein tarkoittavan samaa.

Kitchin (2014, 1) pitää dataa raakamateriaalina, jonka avulla ympäröivä todellisuus voidaan kategorisoida, jakaa mitattaviin olioihin tai muihin selittäviin malleihin. Näitä malleja voivat olla esimerkiksi numerot, kuvat, äänet tai vaikkapa bitit, joista on mahdollista jalostaa tietoa (engl. information) ja ymmärrystä (engl. knowledge). Samoilla linjoilla ovat Poikola ym. (2010, 14) joiden mukaan dataan usein viitataan kun puhutaan digitaalisesti tallennetusta informaatiosta. Englanninkielinen termi information voidaan määritellä olevan dataa joka on sellaisessa muodossa, että sillä on jokin merkitys vastaanottajalle (March & Smith 1995.)

Sydänmaanlakka (2003, 176-177) esittää tiedolle seuraavan hierarkian:



KUVIO 5 Tiedon hierarkia

Kuviossa 5 Sydänmaanlakka (2003, 176) määrittelee datan informaation raaka-aineeksi, jolla ei ole suhteita eikä merkityksiä. Tämä raaka-aine, kuten Kitchin (2014, 1) myös määrittelee, on numeroita, kuvia, tekstiä tai niiden yhdistelmiä. Informaatioksi data muuttuu kun se sisältää jonkun viestin. Tällöin informaatiota voi olla vaikka sähköpostiviesti (Sydänmaanlakka 2003, 176-177.) Tiedon määritelmää Sydänmaanlakka (2003, 178) pitää vaikeampana. Tietoa muodostuu ajattelemalla ja kokemusten kautta ja sitä voidaan arvioida aiempien tapahtumien perusteella sekä pohtia mihin tapahtumat johtavat. Älykkyyteen johtaa oikeaan aikaan käytetty tieto, jolloin yksilö voi tehdä oikeita ratkaisuja, päätöksiä ja valintoja. Viisaudella taas tarkoitetaan jotakin huolellisesti sisäistettyä, sovellettua ja itse kehitettyä tietoa johon liittyvät arvot, moraalit ja yksilön omat kokemukset.

Poikola ym. (2010) käyttävät sanaa data tai siitä johdettuja termejä kuten datalähde tai katalogi, kuvaamaan digitaalisessa muodossa olevaa raakadataa. Tietoon viitataan myös termillä *tietovaranto*. Tässä opinnäytetyössä, kuten luvussa kaksi jo sivuttiin, puhutaan termeistä data, tieto tai tietovaranto kun puhutaan datan julkistamisesta.

4.2 Avoimen datan hyödyt

Poikola ym. (2010) alleviivaavat avoimen tiedon avulla avautuvia mahdollisuuksia esimerkiksi uusille yhteistyökumppaneille ja käyttötarkoituksille. Datan julkaisu tulisi nähdä vuorovaikutteisena prosessina, sillä organisaation ulkopuolisilla tahoilla voi olla datalle kansalaisia hyödyttäviä käyttökohteita, usein yhdessä muiden datalähteiden kanssa.

Gigler ym. (2011) määrittelevät julkishallinnon datan julkaisun taustalle neljä motivaatiotekijää. Ensimmäiseksi, tiedon julkistamisella voidaan parantaa päätöksenteon läpinäkyvyyttä, jonka myötä tehokkuus paranee ja korruption mahdollisuus pienenee. Tällöin julkiset päätökset ja vastuunkantajat päätöksien taustalla olisivat kansalaisten tiedossa. Gigler ym. (2011) eivät enää pidä kansalaisia vain passiivisina tiedon vastaanottajina vaan aktiivisina tiedon hyödyntäjinä. Avoimella pääsyllä julkishallinnon dataan eri käyttäjäryhmät voisivat helpommin arvioida päätöksentekoa sekä puoltaa palveluiden tuottamista.

Toiseksi, avoin data voidaan nähdä alustana uusille palveluille ja julkisille hyödykkeille. Tämä tarjoaisi yksittäisille henkilöille, yksityiselle sektorille ja eri organisaatioille mahdollisuuden yhteistyöhön, jolloin uusien innovaatioiden myötä voidaan helpottaa talouden kasvua (Gigler ym. 2011).

Kolmantena motivaatiotekijänä Gigler ym. (2011) pitävät kattavaa osallistumista ja kiinnostusta päätöksentekoon ja hallinnon tuottamaan dataan. Avoimen datan puolestapuhujat korostavat datan julkaisun demokratisoivaa vaikutusta, jolloin kansalaisten on helppo ottaa osaa keskusteluun hallinnon päätöksiin liittyen.

Viimeisenä motivaatiotekijänä julkishallinnon avoimien tietovarantojen takana Gigler ym. (2011) esittävät hallinnon tehokkuuden parantuvan tiedon infrastruktuurin, virastojen yhteistyön sekä työntekijöiden suorituskyvyn noustessa yhdessä digitalisoituvan hallinnon kanssa.

Tämän tekijän taustalla julkinen sektori nähdään sekä tiedon tuottajana että hyödyntäjänä.

Parannettavaa avoimen datan hyödynnettävyyteen liittyen Ubaldi (2013, 11) näkee datan julkaisumuodossa sekä avoimen datan yhteisön epäjärjestäytyneisyydessä. Eri käyttäjäryhmät haluavat datan eri muodossa, esimerkiksi ohjelmoija haluaa koneluettavaa raakadataa ja yksityishenkilö datasta tehtyjä hyödykkeitä kuten julkisen liikenteen reittioppaita. Yrittäjille taas tärkeää olisi kehittää datalle selvät uudelleenkäyttöehdot sekä luotettava lisensointi.

4.3 Avoimen datan käyttökohteet

Avointa dataa hyödynnetään monenlaisissa käyttökohteissa, mikä voi unohtua jos pohditaan vain jokapäiväisessä käytössä olevia verkkopalveluita sekä mobiilisovelluksia. Tässä kappaleessa avataan lukijalle mahdollisia kohteita julkishallinnon avoimelle datalle.

Yhdistelmäpalveluihin (engl. mashup) kuuluvat tietoa useista eri lähteistä hakevat sovellukset, joiden tehtävänä on helpottaa kansalaisten arkea jollakin alueella visualisoimalla, suodattamalla tai kokoamalla tietoa näistä lähteistä (Poikola ym. 2010, 22). Näihin sovelluksiin kuuluvat esimerkiksi Google Mapsin avointa rajapintaa hyödyntävät sovellukset, jotka yhdistävät tapahtumia maantieteellisiin sijainteihin.

Avoimella datalla ei ole tarkoitus ainoastaan suunnitella pienemmän mittakaavan sovelluksia, joilla on tarkoituksena jonkun arkipäiväisen ongelman ratkaiseminen. Myös koulutus, tutkimus ja tuotekehitys tarvitsevat avoimen tiedon tarjoamia mahdollisuuksia, jolloin tavoitteena on kokonaan uuden tiedon aikaansaaminen ja laajojen data-aineistojen hyödyntäminen toiminnan parantamiseksi. Esimerkiksi eri liikennetietojen avulla voidaan parantaa kaupungin liikennejärjestelmän toimivuutta. Tämän kaltaisia palveluita organisaatioiden omista datavarannoista tehdään jo, mutta avoimella tiedolla lisälähteiden hyödyntäminen muilta

organisaatioilta avaisi ovia monipuolisimpiin ratkaisuihin (Poikola ym. 2010, 22).

Erilaiset automaatioimet, kuten sää- ja sähköverkkodatasta tietonsa hakeva lämmitys- ja ilmastointijärjestelmä, voisivat hyötyä avoimesta datasta prosessin ohjauksessa tai helpottamisessa. Tämän kaltaiset järjestelmät voisivat laajempien avoimien aineistojen avulla säädellä halutun automaation toimintaa mahdollisimman edulliseksi (Poikola ym. 2010, 23).

Avoin data voidaan nähdä myös mahdollistajana suurien käyttäjäkuntien hyödyntämisessä erilaisten ongelmien ratkaisussa, jolloin perinteinen organisaation sisällä tapahtuva tiedonhaku jäisi vähemmälle. Tämä niin sanottu crowdsourcing-ajattelu nojaa ajatukseen uudenlaisesta työnjaosta internetin anonyymeja käyttäjiä hyödyntäen. Tällöin käyttäjille annetaan ongelma johon kuka vain tarvittavan tietotaidon omaava voi tuoda osaamisensa mukaan ongelman ratkaisemiseksi. Tämänkaltaisessa tapauksessa kansalaiset ja yritykset tulisi nähdä tiedon tuottajina uusien sovelluksien synnyttämiseksi (Poikola ym. 2010, 23).

4.4 Avoimen datan periaatteet

Open Government Data (2016) määrittelee avoimelle tiedolle periaatteet, jotka käydään läpi tässä kappaleessa. Useat eri organisaatiot ovat luoneet omat periaatteensa, jotka voivat hieman poiketa seuraavaksi esitettävistä.

Ensinnäkin, julkaistavan tiedon tulee olla kokonaan julkisesti saatavilla huomioon ottaen turvallisuustekijät sekä yksityisyydensuojan. Poikola ym. (2010, 35) tarkentavat että datasta ei saa puuttua osia tai pääsy aineistoon ei koske vain tiettyä osaa.

Toiseksi, data tulee julkaista alkuperäisessä muodossaan, ei muokattuna eikä yhdisteltynä. Tähän Poikola ym. (2010, 36) lisäävät tietosuojan huomioimisen. Heidän mukaansa data voi olla raakadatan lisäksi datan jalostettua tai yhdistettyä muotoa.

Kolmanneksi, datan julkaisun tulee tapahtua niin pian kuin mahdollista tiedon arvon säilyttämiseksi. Poikola ym. (2010, 36) peräänkuuluttavat nykyaikaisen teknologian mahdollisuuksia tiedon ajantasaisuuden suhteen. Heidän mukaansa viivästyneesti julkaistun datan arvo on rajallisempi, jolloin esimerkiksi julkisen liikenteen aikataulujen ajantasaisuudesta kertova sovellus menettää tarkoituksensa.

Neljäntenä esitetään, että datan tulee olla helposti saatavilla mahdollisimman monelle käyttäjälle sekä eri tarkoituksiin. Poikola ym (2010, 27) pitävät tärkeinä niin teknisiä ominaisuuksia kuin perinteisiä mainontakanavia jotta data olisi mahdollisimman helposti löydettävissä. Teknisestä löydettävyydestä puhuttaessa tarkoitetaan esimerkiksi parantamalla hakukoneoptimointia tai hyödyntämällä linked data-ajattelua.

Viidenneksi, avoimesti julkaistun datan tulee olla koneluettavaa, johon Poikola ym. (2010, 37) lisäävät, että datalla tulee olla pysyvä sijainti internetissä ja sen tulee olla tarpeeksi strukturoitua jotta siitä voidaan lyhyessä ajassa hyödyntää esimerkiksi mobiilisovelluksen tekemisessä. Koneluettavuus ja muut tekniset näkökohdat datan julkaisuprosessin taustalla käydään läpi kappaleessa 4.6.

Datan tulee myös olla saatavilla ilman rekisteröitymistä. Tässä Poikola ym. (2010, 35) kirjoittavat myös käyttöehtojen tasa-arvoisuudesta, jolloin kaikki datan halukkaat käyttäjät tulisi olla yhtäläisessä asemassa riippumatta taustasta. Datan käyttökohteita ei myöskään rajata mihinkään, esimerkiksi vain tutkimustyöhön ja lisensointiehtojen tulee noudattaa standardeja.

Open Government Data (2016) määrittelee avoimelle datalle vielä seuraavat periaatteet:

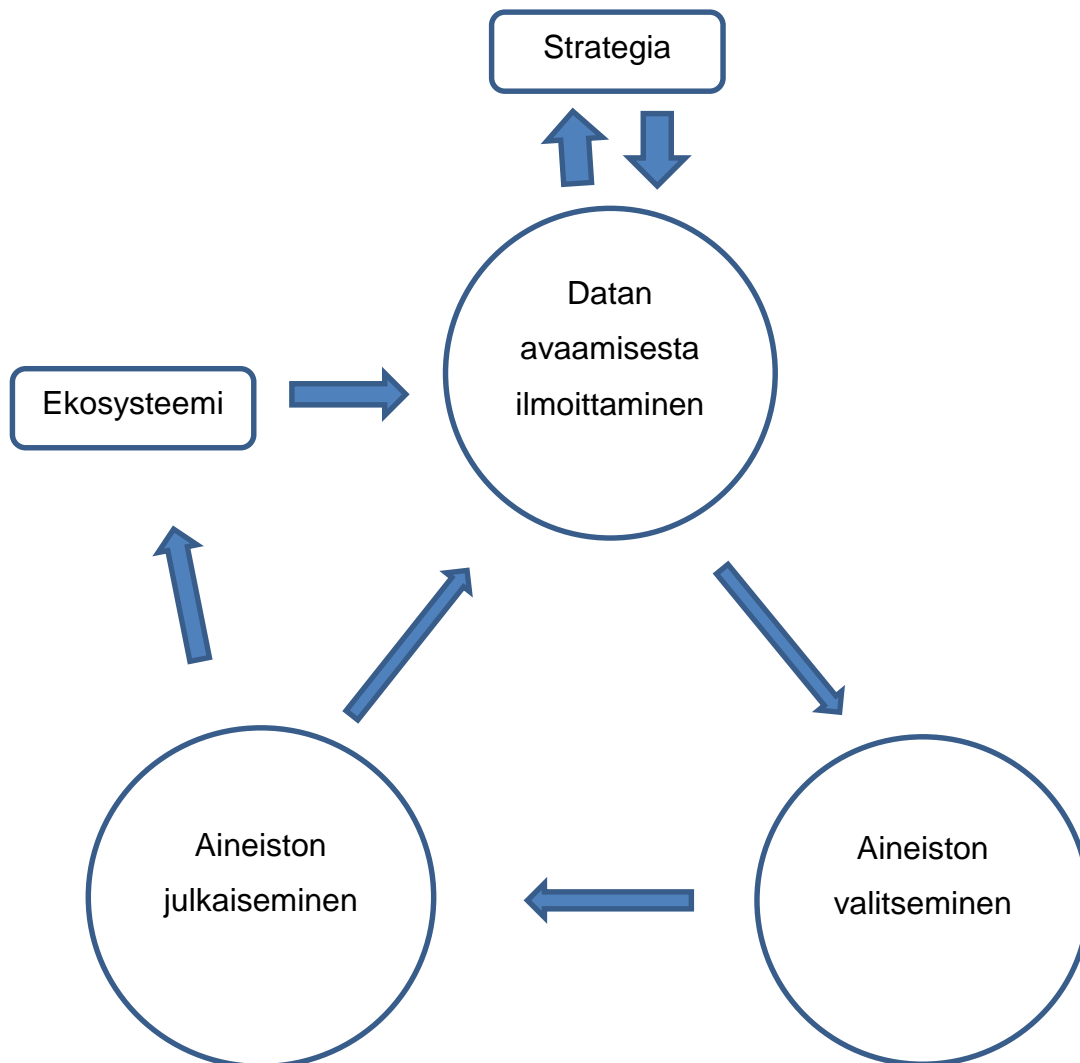
- Datalla ei tule olla omistajaa. Toisin sanoen, yhdelläkään organisaatiolla ei saa olla keskitettyä omistajuutta datalle.
- Datan käyttöön ei tarvita lisenssiä, sillä ei ole tekijänoikeuksia, patenttia eikä tavaramerkkiä.

- Datalla tulee olla kontaktihenkilö datan hyödyntämiseen ja väärinkäyttöön liittyviä kysymyksiä varten.

Poikola ym. (2010, 35) lisäävät avoimen datan periaatteeksi maksuttomuuden. Tähän voidaan lisätä, että pieni maksu ei tee datasta vielä suljettua, mutta voi huomattavasti lisätä haluttomuutta datan uudelleenkäytölle.

4.5 Avoimen datan julkaisuprosessi

Datan julkaiseminen tulee nähdä organisaatiossa prosessina, jossa tutustutaan olemassa oleviin tietovarantoihin sekä niiden hyödyntämismahdollisuuksiin. Dataa julkaistaessa tulee huomioida prosessin olevan pitkä, iteratiivinen prosessi joka ei pääty ensimmäisiin julkaistuihin rajapintoihin. Datan julkaisuprosessissa uusien iteraatiokierroksien avulla pyritään oppimaan datan julkaisusta organisaatiossa ja näin suunnittelemaan uusien avauksien tekemistä (Poikola ym. 2010, 38). Tässä kappaleessa pohditaan organisaation datavarantojen sisällön evaluointia sekä prosessia avoimen datan julkaisun taustalla kuvion 6 (Poikola ym. 2010, 40) mukaan. Kuviossa ekosysteemi tarkoittaa datan hyödyntäjiä sekä julkaisijoita yleisesti, niin kansalaisia kuin eri organisaatioiden työntekijöitä.



KUVIO 6 Esimerkkiprosessi datan julkaisulle

Datan julkaisu tulisi aloittaa olemassa olevan aineiston läpikäynnillä jolloin selvitetään millaista dataa organisaatiolla on. Tällöin kiinnitetään huomiota myös siihen, kuinka julkista julkaistava data jo on. Ensimmäisenä julkaistava aineisto on tämän datan läpikäynnin tuloksia, ei niinkään vielä avoimia rajapintoja. Tällöin organisaatio voi viedä alustavaa tietoa datasta mahdollisille hyödyntäjille, joita voivat olla esimerkiksi kehittäjäyhteisöt (Poikola ym. 2010, 39;41).

Kun aineistoa on ensin käyty läpi, julkaistaan ensimmäiset aineistot. Nämä ensin julkaistavat aineistot tulisi valita niin, että julkaistava data olisi teknisesti helposti avattavissa. Tällä toisella kierroksella pohditaan myös

käyttöehtoja sekä sitä, kuka tai mikä taho on vastuussa datasta. Ensimmäisiä aineistoja julkaistessa tarkoituksena on tehdä dataa tunnetuksi ja mainostaa sitä mahdollisille kehittäjille, jolloin yhteistyössä eri yhteisöjen kanssa tulevaisuuden aineistojen julkaisu on helpompaa. Tässä vaiheessa on myös helppoa luokitella aineisto siten, että salainen materiaali erotetaan julkaistavaksi kelpaavasta (Poikola ym. 2010, 39).

Kolmannella iteraatiokierroksella tulisi pohtia sitä, kuinka hyvin data on dokumentoitu jotta mahdollisimman moni mahdollinen käyttäjä löytää aineiston. Tällä kierroksella tulisi ottaa esille tiedon saatavuuden ja käytettävyyden helppous sekä jo tiedossa olevat käyttäjäryhmät. Tällä kierroksella julkaistujen aineistojen jakelua parannetaan esimerkiksi julkaisemalla avoin rajapinta, jolloin voidaan lisätä jo julkaistun aineiston kiinnostavuutta (Poikola ym. 2010, 39;41).

Neljännellä iteraatiokierroksella organisaatiolla on jo teknisesti hyvin dokumentoituja, laadukkaita aineistoja julkaistuna ja ehkä kokemuksia avatun datan hyödyntämisestä. Hyvässä tapauksessa organisaatiolla on yhteiskäyttöistä dataa muiden organisaatioiden kanssa. Tässä vaiheessa aineiston kehittäminen tulee ajankohtaiseksi, jolloin käyttäjien toiveet ja kokemukset otetaan huomioon sekä pohditaan, onko tieto varmasti hyödynnettävässä muodossa. Aineiston kuvaamiseen tulee kiinnittää huomiota, jolloin vältetään yhteensopivuusongelmat ja saadaan mahdollisimman laaja katsaus hallussa olevista datavarannoista (Poikola ym. 2010, 43).

Jokaisella iteraatiokierroksella aineistoja voidaan luonnollisesti julkaista lisää. Datan käyttökohteita myös seurataan ja datan hyödyntäjiä tavataan esimerkiksi erilaisissa avoimen datan työpajoissa, jolloin organisaation on helppoa pohtia tulevia datan avauksia. Datan julkaisuprosessin tuloksena julkaistava data tulisi tehdä mahdollisimman helposti ja laajasti eri käyttäjille löydettäväksi. Tärkeänä seikkana tulee ottaa huomioon muut datan avaamisesta hyötyvät organisaatiot ja yhteistyön mahdollisuus.

4.6 Tekninen toteutus

Avointa dataa julkaistessa tulee miettiä aineiston julkaisumuotoa siten, että se on mahdollisimman helposti jatkokäytettävissä ja mahdollisimman suurelle käyttäjäryhmälle saatavissa. Dataa voidaan julkaista niin koneluettavassa tiedostomuodossa kuin avaamalla rajapinta jostakin tietojärjestelmästä. Tässä kappaleessa käydään läpi teknisiä ratkaisuja rajapintojen tai muun datan julkaisun taustalla.

Aiemmin opinnäytetyössä on sivuttu termiä koneluettavuus puhuttaessa datan sijoittamisesta avoimesti saataville.

Teksti tai raportti voi olla digitaalisesti saatavilla esimerkiksi PDF-muodossa, jolloin ihmisen on helppoa lukea tiedoston viesti mutta tietokone ei tätä viestiä voi lukea samalla tavalla. PDF sekä HTML ovat toimivia tiedostomuotoja puhuttaessa ihmiselle helposti luettavasta materiaalista jota tietokoneet eivät sellaisenaan pysty käsittelemään. Hyvä esimerkki koneluettavasta datasta on yksinkertainen tekstitiedosto (Open data handbook 2016).

Taulukkomuotoista dataa on selkeää julkaista pilkku- ja sarkainerotellussa muodossa jolloin sen käsittely taulukkolaskentaohjelmalla on helppoa. Monimutkaisemmissa tapauksissa data voidaan tarjota esimerkiksi seuraavissa muodoissa:

XML, extensible Markup Language, joka on yksinkertainen merkintäkieli datan kuvailuun ja siirtämiseen.

JSON, JavaScript Object Notation on kevyt tiedostomuoto tiedonvälitykseen ja joka on sekä ihmiselle että tietokoneelle helposti luettavissa.

CSV, comma-separated values, tiedostomuoto taulukkomuotoisen datan esittämiseen tekstitiedostossa.

Avoimesta datasta puhuttaessa puhutaan myös ohjelmointirajapinnoista (engl. application programming interface, API). Näiden rajapintojen avulla

mahdollistetaan tietojärjestelmien osien keskenäinen kommunikointi ja mahdollistetaan integraatiot toisiin järjestelmiin. Mikäli tietoa halutaan jakaa internetin välityksellä esimerkiksi toisen sovelluksen tarpeisiin, voidaan toteuttaa verkko-ohjelmointirajapinta eli niin sanottu web-API. Avoimeen dataan liittyen nämä nämä web-API:t ovat merkittävä tekijä, sillä haettaessa tietoa web-rajapinnasta ohjelmointityö vähenee datan ollessa valmiiksi saatavilla ajantasaisessa muodossa. Yksinkertainen esimerkki rajapinnan hyödyntämisestä on tietynlaisen kutsun lähettäminen rajapintaan, jolloin vastauksena saadaan pyydetty tiedosto. Erilaiset rajapintaratkaisut ovat saaneet suurta suosiota kehittäjien keskuudessa, näistä hyvänä esimerkkinä Googlen tarjoamat rajapinnat (Poikola ym. 2010, 71).

5 TUTKIMUKSEN SUORITTAMINEN

5.1 Avoin data Lahden ammattikorkeakoulussa

Lahden ammattikorkeakoulussa on suunnitelmissa julkaista avointa dataa aluksi opetuksen suunnittelujärjestelmä Pepistä. Pepin tapauksessa julkaisu tapahtuu järjestelmän avoimien rajapintojen kautta.

Tulevaisuudessa tarkoituksena on julkaista tietoa myös muista lähteistä kuin Pepistä. Opinnäytetyön kirjoittamishetkellä keväällä 2016 avoimen datan suhteen ei ole tehty suunnitelmia Lahden ammattikorkeakoulussa.

Muita data-avauksia on suunniteltu Peppi-järjestelmän avoimien rajapintojen lisäksi. Lahden ammattikorkeakoululla ja Helsingin yliopistolla on meneillään yhteinen projekti Niemen kampuksella sijaitsevan sääaseman kanssa. Tulevaisuudessa dataa tullaan julkaisemaan myös tästä sääasemasta sekä Tekniikan laitoksen palvelintietokoneiden energiankäytöstä (Turunen 2016).

5.2 Peppi-järjestelmä

Peppi on SOA-menetelmin rakennettu tietojärjestelmäkokonaisuus koulutusorganisaatioiden opetuksen suunnittelu- ja toteutustehtäviin. Peppi tarjoaa palvelut niin opetuksen suunnittelijalle, opettajalle kuin molemmille yhteisesti. SOA (Service Oriented Architecture, palvelukeskeinen arkkitehtuuri) on arkkitehtuurityyli jonka perusajatuksena on luoda useita teknologioita hyödyntävä, palvelukeskeinen suunnittelumalli tietojärjestelmien suunnitteluun. SOA-ajattelussa hyödynnetään itsenäisiä, avoimia ja joustavia palveluita joiden välillä integraatioiden tekeminen on helppoa (Poikola ym. 2010, 90).

Opettajan ja suunnittelijan yhteisien palveluiden avulla voidaan hallita muun muassa tilojen-, henkilöiden-, ryhmien- tai välineiden varauksia sekä työaikoihin liittyviä tehtäviä ja henkilöiden resursointia. Suunnittelijan palvelut mahdollistavat esimerkiksi vuosi- ja opetussuunnitelmien

tekemisen. Opettajan näkymässä opettaja voi muun muassa hallita omaa työjärjestystään sekä oppimateriaalejaan (Peppi-konsortio 2015).

Peppi on toteutettu konsortiomallilla, jossa konsortion vastuuna on järjestelmän kehittäminen ja uusien jäsenien hankkiminen. Lahden ammattikorkeakoulu liittyi konsortioon keväällä 2014. Suurin osa konsortion jäsenistä on koulutusorganisaatioita, mutta mukana on myös useita yrityksiä. Järjestelmä on käytössä muun muassa Tampereen ammattikorkeakoulussa, Hämeen ammatti-instituutissa sekä Jyväskylän yliopistossa. Kuten aiemmin todettiin, konsortio on avoin myös opetuksen suunnittelun ja opetushallinnon järjestelmien kehityksestä kiinnostuneille yrityksille. Tutkimuksen kirjoittamishetkellä keväällä 2016 mukana on kahdeksan IT-alan yritystä (Peppi-konsortio 2015).

Pepin palvelumoduuli tarjoaa SOAP-protokollaan ja Rest-standardeihin pohjautuvat rajapinnat, joista Lahden ammattikorkeakoulussa julkaistaan opetussuunnitelmat, opintojaksototeutukset ja varaukset.

5.3 Tutkimusprosessi

Opinnäytetyössä luotavan artefaktin suunnitteluprosessi perustuu kappaleessa 2.3 esitettyyn informaatiostysteemien viitekehukseen jossa otetaan huomioon ympäristö, itse tutkimus sekä tietämuskanta. Tässä luvussa esitellään artefaktin suunnitteluprosessi toimeksiantajalle suunnittelutieteellisen artefaktin toteuttamisen ohjeiden mukaan. Tämä prosessi on käyty läpi kappaleessa 3.3.

Opinnäytetyössä suunniteltavan artefaktin suunnittelu aloitettiin pohtimalla suunniteltavaa artefaktia ja sen tarvetta yhdessä toimeksiantajan kanssa. Tässä vaiheessa tultiin siihen tulokseen, että organisaatiolla on tarvetta kyseiselle artefaktille, jollaista ei ole aiemmin toteutettu. Tämän jälkeen prosessi jatkui lähdemateriaaliin ja vastaaviin palveluihin tutustumisella.

Suunnittelutieteellisellä prosessilla tulisi ratkaista toimeksiantajan ratkaisematon ongelma. Tämän opinnäytetyön tapauksessa toimeksiantajalla ei ole tietoa siitä, kuinka avointa dataa voitaisiin julkaista.

Tämä on opinnäytetyössä ohjaava ongelma johon pyritään saamaan vastaus.

Suunnittelutieteellisellä tutkimuksella tulisi tuottaa uutta tietoa, uusia menetelmiä tai merkittävä artefakti. Prosessin lopussa Lahden ammattikorkeakoululla on selkeä menetelmä avoimen tiedon julkaisemiseksi sekä uutta tietoa aiheesta. Tämä tieto, jota on prosessin aikana kerätty, auttaa toimeksiantajaorganisaatiota myös tulevaisuudessa ja menetelmää voidaan soveltaa uusien data-avauksien tapauksessa.

Tieteellinen tarkkuus opinnäytetyössä tulee esille sillä, että suunniteltu prosessi toteutettiin tarkasti artefaktin tuleva ympäristö huomioon ottaen uusia iteraatiokierroksia tehden. Prosessin aikana pohdittiin erilaisia mahdollisuuksia artefaktin toteuttamiseksi aina edellisistä parantaen.

Artefaktia ei testata sen luonnollisessa ympäristössä rajallisen aikataulun vuoksi. Opinnäytetyön tulokset julkaistaan Lahden ammattikorkeakoulun työntekijöille sekä opiskelijoille, jotka tarvitsevat tietoa avoimeen dataan liittyen. Prosessissa otettiin huomioon suunnittelutieteellisen viitekehyksen mukaisesti artefaktin tuleva ympäristö ja sen hyödyntäjät.

6 TUTKIMUSTULOKSET

Prosessi avoimen datan julkaisulle Peppi-järjestelmästä muodostettiin kappaleessa 4.5 esitellyn prosessin mukaan huomioiden Lahden ammattikorkeakoulun olevan oppilaitos. Prosessissa huomioitiin myös suunnittelutieteellisen artefaktin suunnittelussa käytetyt vaiheet, kuten kappaleessa 3.3 on esitetty. Tutkimuksessa pohdittiin suunnittelutieteellisen tutkimuksen lähestymistavan mukaan artefaktin *rakentamista*. Kuten Järvinen & Järvinen (2004, 12) toteavat, käsitteistöt, mallit ja metodit auttavat suunnittelemaan lopullisen toteutuksen. Nämä artefaktien jaottelut on käyty läpi kappaleessa 3.2. Tässä opinnäytetyössä suunniteltiin metodi, jota voidaan pitää suunnittelutieteellisenä artefaktina.

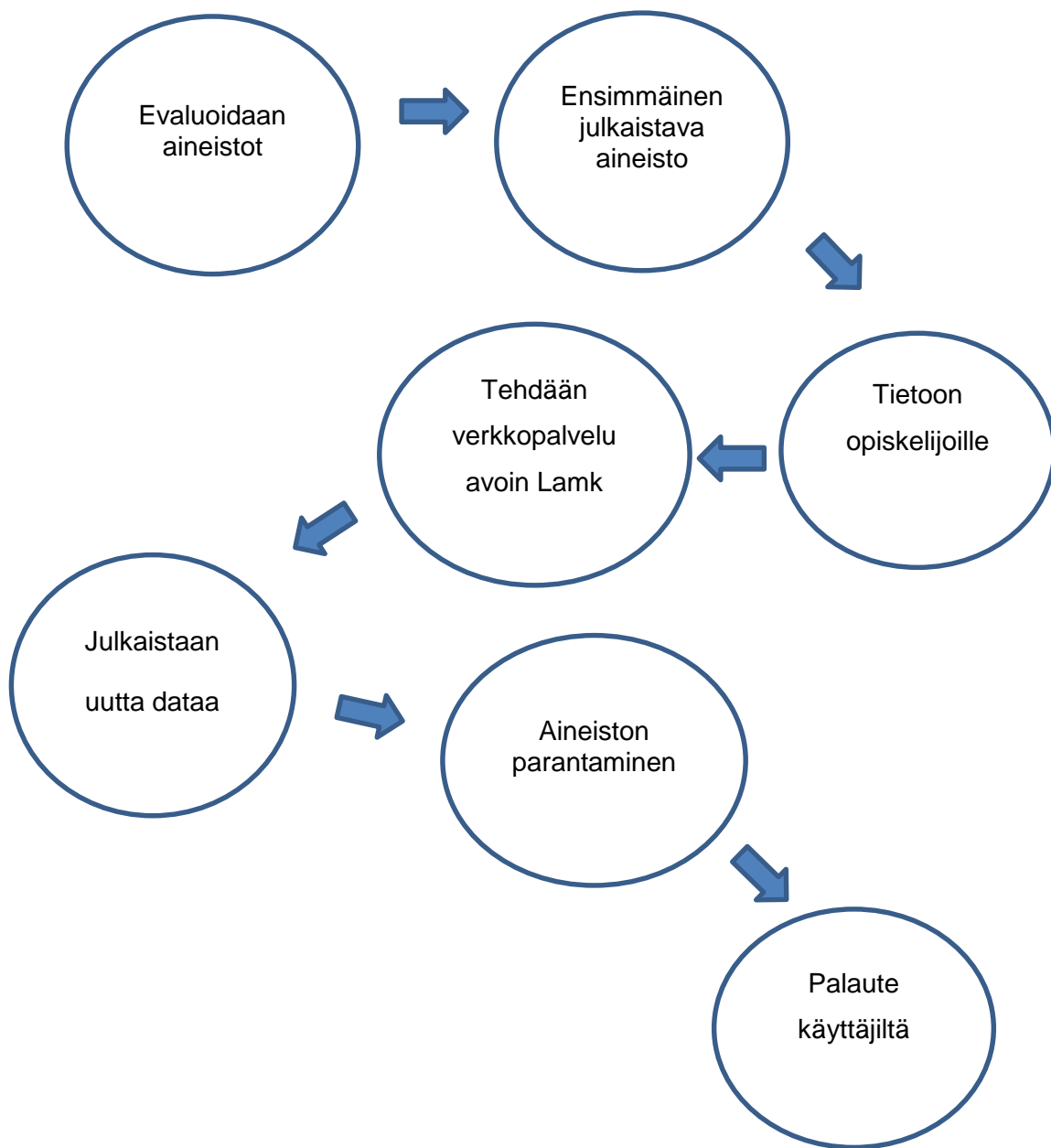
Alkuvaiheessa avointa dataa Peppi-järjestelmästä julkaistaan opintojaksojen toteutuksesta, opetussuunnitelmista sekä varauksista. Kuten Poikola ym. (2010, 21) toteavat, datan julkaisuprosessissa tulee ottaa huomioon datan käyttäjät, toimintatavat sekä oppimismahdollisuudet. Lahden ammattikorkeakoulun tapauksessa suurin käyttäjäryhmä on opiskelijat, etenkin tekniikan alalta. Avoimen datan suunnitteluprosessi aloitettiin pohtimalla kappaleessa 4.5 läpikäytyä viitteellistä datan julkaisuprosessia sekä suunnittelutieteellistä viitekehystä seuraten. Peppi-järjestelmän tapauksessa prosessi on helppoa aloittaa, sillä tiedossa on jo mitä aiotaan ensimmäiseksi julkaista eikä järjestelmään itsessään tarvitse tehdä muutoksia. Prosessi on suunniteltu niin, että tulevaisuudessa uusia data-avauksia tehdessä sitä on hieman soveltaen helppo seurata.

Julkaistessa avointa dataa Lahden ammattikorkeakoulussa, tulisi prosessissa lähteä liikkeelle pienin askelin, jolloin on helppoa lähteä evaluimaan avoimien rajapintojen tarvetta tulevaisuudessa. Aluksi pohditaan jo olemassa olevia aineistoja ja sitä, mitä halutaan julkaista. Ensimmäinen avattu rajapinta voitaisiin tuoda tietoon esimerkiksi tietotekniikan opiskelijoille, jotka voisivat hyödyntää rajapintaa opiskeluprojektissaan. Näitä projekteja tulisi tuoda esille avoin LAMK-verkkopalveluun, jolloin rajapintojen muut mahdolliset käyttäjät saisivat käsityksen siitä, mihin rajapintoja voidaan hyödyntää. Tämä verkkopalvelu

tulisi toteuttaa mahdollisimman pian ensimmäisten data-avauksien jälkeen jotta tietoisuus avoimen datan mahdollisuuksista leviää.

Oleellista prosessissa on, että datan julkaisu jatkuu mahdollisuuksien mukaan. Uusia rajapintoja voidaan avata tai dataa julkaista muussa muodossa. Jotta avoin data Lahden ammattikorkeakoulussa pysyy houkuttelevana kehittäjien keskuudessa, tulee huolehtia aineiston laadusta. Tähän kuuluvat muun muassa aineiston helppo löydettävyys sekä tiedon ajantasaisuudesta huolehtiminen.

Ideaalitulanteessa käyttäjiltä saadaan palautetta dataan liittyen jolloin muutoksia on tarvittaessa helppoa tehdä ja tulevaisuuden toimintaa suunnitella. Seuraavaksi esitellään prosessi avoimen tiedon julkaisulle toimintaohjeina.



KUVIO 7 Prosessi avoimen datan julkaisulle

7 YHTEENVETO

Opinnäytetyössä oli tarkoituksena luoda prosessi avoimen tiedon julkaisulle Lahden ammattikorkeakoulussa. Opinnäytetyön aikana suunniteltiin prosessi avoimen datan julkaisemiseksi Lahden ammattikorkeakoulussa. Tutkimustulokset johdettiin aiempien samankaltaisten tutkimuksien ja aiheeseen liittyvien artikkeleiden pohjalta suunnittelutieteellisen viitekehyksen ohjatessa tutkimusta. Artefaktin tuleva ympäristö otettiin huomioon suunnitteluprosessissa ja lopullinen artefakti rakennettiin aiempaan tietämykseen perustuen.

Opinnäytetyössä suunniteltu tiedon julkistamisen prosessi vastaa suunnittelutieteellisen artefaktin suunnittelussa käytettyä prosessia sekä suunnittelutieteellisen artefaktin määritelmää. Suunnittelun artefaktin avulla ratkaistiin tutkimusongelma, joka oli alussa esitelty seuraavasti:

- Kuinka avoimen datan julkaiseminen Lahden ammattikorkeakoulussa tulisi toteuttaa?

Opinnäytetyön tuloksena syntynyt artefakti tuotti uutta tietämystä ja uuden metodin. Toimeksiantajaorganisaatiolla ei ennen opinnäytetyöprosessia ollut tietoa eikä suunnitelmaa opinnäytetyöhön liittyvän artefaktin suhteen, joten suunnittelutieteellisen ongelman voidaan katsoa olevan ratkaistu. Tässä vaiheessa voidaan myös todeta tuloksien täyttäneen opinnäytetyön alussa määritellyt tavoitteet ja suunnittelun artefaktin täyttävän suunnittelutieteellisen artefaktin määritelmän.

Opinnäytetyö oli kvalitatiivinen tutkimus jossa hyödynnettiin deduktiivista lähestymistapaa sekä suunnittelutiedettä. Prosessi tiedon julkaisulle toimeksiantajalle suoritettiin aiempien tutkimuksien ja aineiston perusteella informaatiojärjestelmien viitekehykseen pohjautuen. Opinnäytetyöprosessin alussa määriteltyyn tutkimuskysymykseen saatiin vastaus.

Avoin data kokonaisuudessaan on varsin laaja käsite, ja tämä opinnäytetyö on vain pieni katsaus siitä mitä kaikkea tiedon avoimuudella voidaan nykypäivänä saavuttaa. Avoimen tiedon tulevaisuus vaikuttaa

lupaavalta, joten tästä opinnäytetyöstä on hyötyä niin tekijälleen, toimeksiantajalleen kuin muille aiheesta kiinnostuneille.

Vaikka opinnäytetyössä käsitelty aihe on varsin nuori ja suuressa nosteessa, on datan julkaisua käsitelty lähinnä julkishallinnon näkökulmasta. Avoimen datan julkaisuprosessista on vain vähän tieteellisiä artikkeleita etenkin muuhun kuin julkishallinnon avoimeen tietoon liittyen. Tämän takia opinnäytetyössä hyödynnettiin pitkälti avoimen julkishallinnollisen datan hyötyjä, määritelmiä sekä julkaisuprosessia. Edellä mainittuja sovellettiin tutkimuksen toimeksiantajan tapaukseen.

7.1 Avoin LAMK-palvelu

Toimeksiantajaorganisaation pyynnöstä tässä kappaleessa sivutaan hieman tulevaa verkkopalvelua, jonka on tarkoitus kerätä yhteen avoimia rajapintoja ja muuta avointa dataa Lahden ammattikorkeakoulussa. Tämä Avoin LAMK-verkkopalvelu tulisi olla Lahden ammattikorkeakoulun avointa dataa yhteentuova palvelu, jonne kerätään tietoa avatusta datasta ja rajapinnoista sekä niistä tehdyistä sovelluksista.

Tämän luvun asiat ovat opinnäytetyön kirjoittajan omaa pohdintaa avoimen tiedon mahdollisuuksista, vaikkakin muutama aiempi esimerkki tuodaan esille.

Avoimen datan hyödyntämiseksi Lahden ammattikorkeakoulussa voitaisiin pohtia erilaisia kannustimia avoimen datan hyödyntämiseen. Esimerkiksi erilaisilla kilpailuilla tai monialaisilla projekteilla saataisiin opiskelijat tietoisiksi rajapintojen mahdollisuuksista. Washington D. C:n hallinto järjesti ensimmäisen Apps for Democracy-kilpailun vuonna 2008, jonka tarkoituksena oli koota yhteen sovelluskehittäjiä ja muita aiheesta kiinnostuneita suunnittelemaan käyttötarkoituksia kaupungin hallitsemille tietovarannoille. Idea sai huippusuosion ja on sittemmin levinnyt ympäri maailmaa (Pioneerinstitute 2016.) Vastaavankaltaisia tapahtumia voisi järjestää myös Lahden ammattikorkeakoulussa, jolloin avoimet rajapinnat

otettaisiin mukaan esimerkiksi opiskelijoiden Hackathon-tapahtumia suunniteltaessa.

Eri organisaatioiden ja opiskelijoiden yhteistyössä toteuttamat projektit avoimeen dataan liittyen tulisi nostaa esille tulevaisuuden opintotarjontaa suunniteltaessa. Hyvä esimerkki tällaisesta toiminnasta on Aalto-yliopiston ohjelmistoprojektikurssin opiskelijoiden sekä Helsingin kaupungin yhteistyössä tekemä avoimeen rajapintaan perustuva verkkopalvelu, joka tarjoaa kaupunkilaisille mahdollisuuden antaa palautetta kaupungin asioista (Helsinki Region Infoshare 2016.) Tämän kaltaisia projekteja Lahden ammattikorkeakoulu voisi hyödyntää avoimiin rajapintoihin liittyen, jolloin opiskelijat pääsisivät tutustumaan työelämälähtöisiin projekteihin opiskeluaikana sekä voisivat luoda kontakteja eri sidosryhmiin.

7.2 Jatkotutkimus ja yleistettävyyys

Jatkossa voidaan pohtia rajapintojen käyttöä sekä niiden laajempaa hyödyntämistä. Tähän hyödyntämiseen voitaisiin ottaa huomioon alueen yritykset ja muut sidosryhmät. Erilaisilla kyselytutkimuksilla avoin LAMK-palvelun kiinnostavuutta olisi helppoa arvioida ja samalla suunnitella tulevaisuuden data-avauksia vastauksien mukaisesti.

Tässä opinnäytetyössä toteutettu prosessi tiedon avaamiselle voidaan yleistää myös muihin datan avaamisprojekteihin erilaisissa organisaatioissa. Mikäli avoin tieto jatkaa samaa kehityskulkua kuin opinnäytetyön kirjoittamishetkeen kevääseen 2016 asti, voidaan olettaa tulevaisuudessa yhä useamman yhteisön ja organisaation olevan kiinnostunut opinnäytetyössä läpi käydyistä asioista.

7.3 Validiteetti ja reliabiliteetti

Kvalitatiivisen tutkimuksen luetettavuutta voidaan mitata reliabiliteetin ja validiteetin avulla. Puhuttaessa reliabiliteetista tarkoitetaan tutkimustuloksien toistettavuutta. Reliabiliteettia voidaan lähestyä kysymällä: ”Kuinka samanlaisia tai toisistaan poikkeavia vastauksia

saataisiin?” mitattaessa samaa ilmiötä useaan kertaan samalla mittarilla (Metsämuuronen 2006, 56). Kvalitatiivisen tutkimuksen tapauksessa voidaan pohtia myös tutkimuksen johdonmukaisuutta, sillä perinteiset luotettavuuden mittarit, jotka toimivat hyvin kvantitatiivisessa tutkimuksessa, voidaan kyseenalaistaa kvalitatiivisen tutkimuksen luotettavuutta ja arviointia tehdessä (KvaliMOTV.)

Vaikka opinnäytetyössä suunniteltu prosessi on suunniteltu tiettyä viitekehystä noudattaen, jo olemassa olevaan aineistoon perustuen, voi joku toinen tutkija päätyä varsin erilaisiin ratkaisuihin vastaavaa suunniteltaessa. Täten voidaan tulla tulokseen, että tutkimus sinänsä tuskin on toistettavissa vaikka prosessi suunniteltiin johdonmukaisesti kerättyyn aineistoon perustuen.

Tutkimuksen validiteetti mittaa tutkimuksen pätevyyttä, eli sitä, kuinka perusteellisesti tutkimus on tehty sekä tuloksien ja tehtyjen päätelmien ”oikeutta”. Kvalitatiivisen tutkimuksen tapauksessa tätä voidaan pohtia uskottavuuden ja vakuuttavuuden näkökulmasta ja siitä, kuinka kehitetyt konstruktiot vastaavat tutkittuja (KvaliMOTV). Tällöin tulee esille, onko tutkimuksessa tutkittu sitä mitä oli tarkoituskin (Metsämuuronen 2006, 56.) Opinnäytetyö perustuu vahvasti olemassa oleviin aineistoihin, joiden pohjalta työssä suunniteltu prosessi muodostettiin tiettyyn viitekehukseen pohjautuen. Tutkimuksessa saatiin vastaus opinnäytetyön alussa esitettyyn tutkimusongelmaan, joten voidaan sanoa tutkimuksen olevan validi.

LÄHTEET

Kirjalliset lähteet:

Järvinen, A. & Järvinen, P. 2004. Tutkimustyön metodeista. Tampere: Opinpajan kirja.

Kitchin, R. 2014. The Data Revolution. London: SAGE Publications.

Sydänmaanlakka, P. 2003. Älykäs organisaatio – tiedon, osaamisen ja suorituksen johtaminen. 6. painos. Jyväskylä: Gummerus Kirjapaino Oy

Tuomi, J. & Sarajärvi, A. 2009. Laadullinen tutkimus ja sisällönanalyysi. Helsinki: Tammi.

Elektroniset lähteet:

Euroopan komissio. 2011. Open data – An engine for innovation, growth and transparent governance [Viitattu 15.2.2016]. Saatavissa: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>

Gigler, B.S., Custer, S. & Rahemtulla, H. 2011. Realizing the Vision of Open Government Data – Opportunities, Challenges and Pitfalls [Viitattu 21.3.2016]. Saatavissa: <http://ifap-is-observatory.ittk.hu/node/853>

Hevner, A., March, Salvatore T., Park, J. & Ram S. Design Science In Information Systems Research 2004 [Viitattu 11.4.2016]. Saatavissa: http://wise.vub.ac.be/thesis_info/design_science.pdf

Helsinki Region Infoshare-verkkosivu 2016 [Viitattu 28.3.2016]. Saatavissa: <http://www.hri.fi/fi/>

Iivari, J. 2007. A Paradigmatic Analysis of Information Systems as a Design Science [Viitattu 22.2.2016]. Saatavissa: <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1018&context=sjis>

Janssen, K. 2012. Open Government Data: right to information 2.0 or its rollback version? [Viitattu 27.3.2016]. Saatavissa:

http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2152566

Johannesson, P. & Perjons, E. 2014. An Introduction to Design Science

[Viitattu 24.2.2016]. Saatavissa: <http://link.springer.com/book/10.1007/978-3-319-10632-8>

KvaliMOTV- Menetelmäopetuksen tietovaranto. Tutkimuksen

luotettavuuden arviointi. Tampere: Yhteiskuntatieteellinen tietoarkisto

[Viitattu 12.4.2016]. Saatavissa:

http://www.fsd.uta.fi/menetelmaopetus/kvali/L3_3.html

March, S.T. & Smith, G.F. 1995. Design and Natural Science Research on Information Technology [Viitattu 9.4.2016]. Saatavissa:

<https://pdfs.semanticscholar.org/5bd7/3700d5b3bafbbd41d274e8f9c4be81950f39.pdf>

Metropolia-ammattikorkeakoulu 2013 - Avoin data Metropoliaassa [Viitattu

11.4.2016]. Saatavissa: <https://opendata.metropolia.fi/>

Open data handbook-verkkosivu 2016 [Viitattu 12.4.2016]. Saatavissa:

<http://opendatahandbook.org/>

Open Government Data-verkkosivu [Viitattu 5.3.2016]. Saatavissa:

<https://opengovdata.org/>

Peffers, K., Tuunanen, T., Rothenberger, M.A. & Chatterjee, S. 2007. A Design Science Research Methodology for Information Systems Research

[Viitattu 14.3.2016]. Saatavissa:

http://wise.vub.ac.be/thesis_info/Design_Science_Research_Methodology_2008.pdf

Peppi-Konsortio. 2015 [Viitattu 6.3.2016]. Saatavissa: [http://www.peppi-](http://www.peppi-konsortio.fi/)

[konsortio.fi/](http://www.peppi-konsortio.fi/)

Pioneerinstitute - Apps for Democracy 2016-verkkosivu [Viitattu 9.4.2016].
Saatavissa: <http://bgc.pioneerinstitute.org/apps-for-democracy/>

Poikola A., Kola P. & Hintikka K.A. Julkinen data, johdatus tietovarantojen avaamiseen. Liikenne- ja viestintäministeriö 2010 [Viitattu 25.3.2016].
Saatavissa:
<http://www.lvm.fi/documents/20181/815557/Julkinen+data/467e7da9-3038-46a1-b47e-994d7cd102d2?version=1.0>

Suomen valtiovarainministeriö. 2015. Avoimesta datasta innovatiiviseen tiedon hyödyntämiseen [Viitattu 9.4.2016]. Saatavissa:
<http://vm.fi/julkaisu?pubid=6902>

Tampereen ammattikorkeakoulu 2016 – Open data [Viitattu 11.4.2016].
Saatavissa: <http://avoindata.tamk.fi/>

Ubaldi, B. 2013. Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives [Viitattu 31.3.2016]. Saatavissa:
http://www.oecd-ilibrary.org/governance/open-government-data_5k46bj4f03s7-en

Suulliset lähteet:

Turunen, T. 2016. Lehtori, Lahden ammattikorkeakoulu. Keskustelu 26.2.2016