

Opinnäytetyö (AMK)

Kirjasto- ja tietopalvelu

NKITIS13

2017

Tuomas Varila

KÄYTTÄJÄTIETOJEN HYÖDYNTÄMINEN KIRJASTOTYÖSSÄ

Tuomas Varila

KÄYTTÄJÄTIETOJEN HYÖDYNTÄMINEN KIRJASTOTYÖSSÄ

Tämän opinnäytetyön tavoitteena on osoittaa käyttäjätietojen hyödyllisyys kirjastoille. Työssä tarkastellaan, mitä tietoja kirjastot tilastoivat ja miten näitä tietoja on hyödynnetty. Perinteisten tilastotietojen lisäksi tutkitaan, miten kirjastot ovat hyödyntäneet ja voisivat hyödyntää kirjastojärjestelmistä saatavaa anonyymia lainadataa. Lainadataa analysoimalla on mahdollista saada huomattavasti tarkempaa tietoa kirjaston toiminnasta ja käyttäjistä.

Opinnäytetyössä esitellään, miten HelMet- ja Vaski-kirjastokimpoissa on hyödynnetty anonyymia lainadataa palvelujen kehityksen ja vaikuttavuuden mittaamisen ja todistamisen näkökulmista. Lisäksi tutkitaan käytännön tasolla, mitä lähitulevaisuudessa kaikkien yleisten kirjastojen saatavilla oleva lainadata on ja miten sitä voidaan hyödyntää.

Opinnäytteen tärkeimpinä tietolähteinä toimivat haastatellut asiantuntijat sekä itse analysoitu lainadata. Lainadatan hyödyntäminen on aineiston määrällistä eli kvantitatiivista analysointia, jonka avulla pyritään löytämään valtavasta aineistosta ilmiöitä, poikkeamia sekä syy-seuraussuhteita. Näitä tutkimalla puolestaan saadaan käyttäjien tarpeista ja kirjaston toiminnasta ja kokoelmista selkeämpi ja yksityiskohtaisempi kuva.

Lainadataa hyödyntämällä kirjastojen palveluita voidaan kehittää vastaamaan käyttäjien tarpeisiin paremmin ja monipuolisemmin. Kunkin kirjaston on mahdollista tutustua käyttäjäkuntaansa aivan uudentlaisella tasolla. Yhdistelemällä lainadataa valtion muiden instituuttien keräämiin tietoihin kansalaisista, voidaan myös mitata minkälaisia vaikutuksia kirjastoilla on käyttäjiensä elämässä.

ASIASANAT:

Käyttäjätiedot, palvelukehitys, tiedonlouhinta, vaikuttavuus

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Library and information studies

2017 | 17

Tuomas Varila

UTILIZING USER DATA IN LIBRARIES

The purpose of this thesis is to show the usefulness of user data in libraries. In the thesis, it is examined what information is compiled into statistics and how this information is used. Alongside traditional statistics, the thesis studies how libraries have and could utilize the anonymous loaning data that can be extracted from the library system. By examining the loaning data, it is possible to get much more specific information about any library's services and users.

The thesis presents how the Helmet and Vaski library networks have utilized the anonymous loaning data in their service development and in the measuring and proving of efficacy. It is also presented what the loaning data is and how it is utilized in practice. This data will soon be available to all public libraries in Finland.

The most important information sources for this thesis were the specialists interviewed and the data being analyzed. Essentially the utilizing of the data means its quantitative analysis which attempts to find phenomena, anomalies and causations in the massive data pack. By examining the above, the library gets a clearer and more elaborate image of the needs of the users and the services and collections provided by the library.

By utilizing the loaning data, the library services can be developed to match the need of the users better and more diversely. Each library can get to know their users on a whole new level. It is also possible to unite the loaning data with data collected by other government institutions to examine what effects the library services have in the everyday life of its users.

KEYWORDS:

Data mining, development, efficacy, user data

SISÄLTÖ

1 JOHDANTO	1
2 KÄYTTÄJÄDATAN HYÖDYNTÄMINEN KIRJASTOSSA	2
2.1 Tutkimusmenetelmät	3
2.2 Palveluiden kehitys	3
2.2.1 Anonyymi lainadata Vaskissa	4
2.2.2 Tulevia anonyymien lainadatan sovellutuksia	5
2.3 Vaikuttavuuden mittaaminen	6
2.3.1 Kirjastojen vaikutusalueet vaikuttavuuden mittareina	6
2.3.2 Kirjaston ympäristövaikutusten tarkastelu lainadatan avulla	6
3 KIRJASTOAUTOJEN VAIKUTUSALUEEN MITTAAMINEN	8
3.1 Käytävissä oleva data	8
3.1.1 Kirjastojärjestelmän lainadata	9
3.1.2 Yhteiskuntarakenteen ja kartat	9
3.2 Datan hyödyntäminen	10
3.2.1 Käytetyt ohjelmistot	10
3.2.2 Työvaiheet	12
3.3 Tulokset	13
4 POHDINTAA	16
LÄHTEET	17

KUVAT

Kuva 1. R:llä erotetut Turun pääkirjaston kokoelmat ja lainat.	11
Kuva 2. Vaski-alueen kirjastoautopysäkit kartalla.	12
Kuva 3. Turun kirjastoauton vaikutusalue.	13
Kuva 4. Liedon kirjastoauton vaikutusalue.	14
Kuva 5. Salon kirjastoauton vaikutusalue.	15

TAULUKOT

Taulukko 1. Turun pääkirjaston kokoelmat ja lainat.

11

1 JOHDANTO

Ihmisen toiminnasta yhteiskunnassa jää tietoa talteen yritysten ja instituuttien tietokantoihin valtavia määriä. Varsinkin yritykset käyttävät jo tehokkaasti käyttäjätietoja markkinoinnin ja palveluiden kehittämiseen yksittäisille käyttäjille sopiviksi: Erityisesti internetissä toimivat jättyyritykset käyvät nykyään kauppaa nimenomaan käyttäjätiedoillaan ja pyrkivät ennakoimaan käyttäjiensä käyttäytymistä.

Myös kirjaston käyttämisestä jää jälki kirjaston eri järjestelmiin. Tätä käyttäjätietoa eli käyttödataa ei kuitenkaan vielä juurikaan hyödynnetä, vaikka se toki olisi jo mahdollista. Kirjastojärjestelmästä saatavia tietoja voi hyödyntää kirjaston palvelunkehityksessä sekä vaikuttavuuden osoittamisessa rahoittajan eli kunnan ja valtion suuntaan. Käyttödataa hyödyntämällä palveluita olisi mahdollista räätälöidä juuri kunkin kirjastoyksikön käyttäjien tarpeiden mukaisiksi. Toisaalta voitaisiin osoittaa valtiolle kirjaston käytön runsaus ja sitä kautta merkitys kansalaisten elämässä.

Tämän opinnäytetyön tavoitteena on esitellä erilaisia keinoja, joilla dataa tähän mennessä on hyödynnetty sekä miten sitä olisi mahdollista hyödyntää. Lisäksi kuvailen tarkemmin itse toteuttamani kirjastoautojen vaikutusalueiden mittaamisen. Käytettävissäni oli Turun pääkirjaston kirjastojärjestelmän Axiell Auroran anonyymi lainadata, Tilastokeskuksen tarjoama yhteiskuntarakenneddata sekä Qgis-ohjelmistoon liitettävät kartat. Lisäksi tein asiantuntijahaastatteluja datan ja tilastojen parissa työskenteleville kirjastolaisille.

2 KÄYTTÄJÄDATAN HYÖDYNTÄMINEN KIRJASTOSSA

Mille tahansa yritykselle tai instituutille on hyödyllistä tuntee asiakaskuntansa mahdollisimman hyvin. Tämä mahdollistaa palveluiden asiakaslähtöisen kehittämisen ja sitä kautta tuo lisää asiakkaita tai käyttäjiä.

Yksi keino tutustua asiakkaiden tai käyttäjien käyttäytymiseen on toiminnan tilastointi. Tämä on siitä hyvä keino, että tilastointia on helppo tehdä rikkomatta asiakkaiden yksityisyyden suojaa: Tilastoihin tallennetaan vain niin ympäröivä tietoja heistä, ettei niiden avulla ole mahdollista seurata yhdenkään yksittäisen ihmisen toimintaa. Esimerkiksi Suomen julkisten kirjastojen järjestelmistä saatavissa tiedoissa yksittäisestä asiakkaasta jää tietoon vain ikä, sukupuoli ja epätarkka osoitetieto (Varila, K. sähköpostihaastattelu, 8.11.2017). Sitten kun näitä tietoja kerätään pitkältä ajalta, on mahdollista nähdä kuvioita: Esimerkiksi mihin aikaan yrityksen tai instituutin palveluita käytetään eniten, missä niitä käytetään ja miten, sekä mikä asiakasryhmä niitä käyttää. Näiden kuvioiden avulla voidaan ennakoida vaikkapa ketkä voisivat olla potentiaalisia uusia asiakkaita, mihin aikaan heitä kannattaa lähestyä ja missä.

Myös kirjasto kerää tietoa käyttäjistään. Jo vuosia kirjastot ovat tilastoineet toimintaansa ja näiden tilastojen avulla on kehitetty kirjaston palveluita ja kokoelmia käyttäjien tarpeiden mukaisiksi. Kirjastot.fi -sivustolta löytyy yleisten kirjastojen tilastoimat tiedot vuodesta 1999 eteenpäin. Nämä tilastot on jaoteltu näin: Toimipisteet, kokoelmat, hankinnat, lainaus, lehdet, e-aineistot, poistot, kirjaston käyttäjät, kirjaston tapahtumat ja käyttäjäkoulutus, kirjaston henkilökunta ja talous. Lisäksi löytyy ryhmä ”Muut tilastot”, mikä käytännössä sisältää kunnan asukasluvun, sekä ryhmä ”TUNNUSLUVUT”. Tunnusluvuilla tarkoitetaan eri tilastolukujen suhteita toisiinsa, esimerkiksi kuinka monta lainaa tapahtuu yhtä kunnan asukasta kohden vuodessa. (Kirjastot.fi 2017.)

Kuitenkin kirjastojärjestelmistä saatavalla datalla on mahdollista kerätä vielä paljon tarkempaa tietoa kirjastonkäyttäjien käyttäytymisestä: Jokaisen kirjastoyksikön tai osaston lainatietoja voidaan tutkia erikseen, eri ikäluokkien lainauskäyttäytymistä voidaan analysoida ja ennakoida, kirjastonkäytön vaikutuksia elämän muihin osa-alueisiin voidaan tutkia.

Tässä opinnäytteessä keskityn tutkimaan lainadatan hyödyntämistä kirjaston palveluiden kehittämisessä sekä toisaalta kirjaston toiminnan vaikuttavuuden mittaamisessa.

2.1 Tutkimusmenetelmät

Tämän opinnäytetyön tärkeimmät tietolähteet olivat itse hyödynnettävä data sekä asiantuntijahaastatteluista saatu tieto. Haastateltavia oli kaksi: Turun pääkirjaston palvelupäällikkö Kalle Varila sekä Turun amk:n kirjastossa suunnittelijana toimiva Liisa Tiittanen.

Koska anonyymien lainadatan hyödyntäminen on melko harvinaista kirjastotyössä, ei aiheesta ole vielä paljonkaan kirjallisuutta tarjolla, joten asiantuntijahaastattelut olivat paras keino saada tietoa aiheesta. Haastattelut toteutettiin pääosin sähköpostitse, osittain myös kasvokkain. Molemmat haastateltavat ovat työskennelleet kirjastoissa tilastojen tai lainadatan parissa.

Toinen oleellinen tietolähde oli itse anonyymi lainadata: Sitä analysoimalla on mahdollista päätellä melko hyvin, mitä kaikkea sillä voi tehdä. Lainadatan hyödyntäminen ja sen analysointi on aineiston määrällistä eli kvantitatiivista analysointia, jonka pyrkimyksenä on saada tilastollista, yleistettävää tietoa eroteltua massiivisesta aineistomäärästä. Jaotteleamalla ja teemoittamalla aineistoa osiin ja vertailemalla osia toisiinsa pyritään löytämään poikkeamia, ilmiöitä ja syy-seuraussuhteita. (Virtuaali ammattikorkeakoulu 2017). Menetelmää voidaan kutsua myös tiedon louhimiseksi: Kyse on pohjimmiltaan haluttujen tietojen kaivamista esiin valtaisasta aineistomäärästä erilaisten ohjelmistojen avulla.

2.2 Palveluiden kehitys

Kirjastot.fi -sivustolta löytyvien tilastojen avulla on mahdollista tutkia esimerkiksi mitkä aineistolajit ovat suosituimpia ja miten kokoelma on kehittynyt vuosien varrella tai vaikkapa mitkä tapahtumat ovat olleet suosituimpia tai onko tapahtumien lisääntyminen vaikuttanut kirjaston käytön määrään.

Monet kirjastot käyttävät tilastoja lähinnä kokoelmatyön apuna: Suosittuja aineistoja lisää, käyttämättömiä ja vanhentuneita pois. Lisäksi tilastoja on hyödynnetty

päätöksenteossa sekä toiminnan arvioinnissa, suunnittelussa ja vertailussa. (Varila, K. sähköpostihaastattelu, 8.11.2017, Tiittanen, L sähköpostihaastattelu, 4.12.2017.) Tällä hetkellä Kirjastot.fi -sivuston tilastot ovat kuitenkin melko rajalliset: Esimerkiksi lainamäärät on ilmoitettu vain koko kunnan tasolla eri aineistolajeihin jaoteltuina. Yksittäisten kirjastoyksiköiden lainamääriä ei löydy, löytyy vain yksiköiden määrä.

Kuitenkin itse kirjastojärjestelmään jää huomattavasti tarkempaa tietoa lainoista, lainaajista ja kokoelmista ja näiden tietojen hyödyntäminen mahdollistaisi paljon tarkempien tilastojen luomisen. Varsinais-Suomen Vaski-kirjastokimppa sekä pääkaupunkiseudun HelMet-kirjastokimppa ovat olleet pilotteina mukana luomassa Kirjastot.fi -palvelimelle Raportointipalvelua, jonka on tarkoitus tulla kaikkien Suomen yleisten kirjastojen käyttöön. Raportointipalvelu perustuu kirjastojärjestelmästä saatavaan anonyymiin lainadataan, joka siis kattaa kirjaston fyysisen kokoelman. Kun Raportointipalvelu tulee käyttöön, on minkä tahansa yleisen kirjaston mahdollista pyytää sieltä käyttöönsä neljästi vuodessa päivittyvä datapaketti.

2.2.1 Anonyymi lainadata Vaskissa

Vaski-kirjastokimpassa anonyymia lainadataan on hyödynnetty tähän mennessä kokoelmatyön apuna sekä erilaisten palvelussa tapahtuvien muutosten vaikutusten tutkimisessa.

Kokoelmatyön tueksi lainadatan avulla on tutkittu kunkin kirjastoyksikön vaikutusalue ja sen käyttäjien ikärakenne. Kunkin kirjaston vaikutusalueesta on lisäksi tutkittu sen asukkaiden sosioekonominen rakenne: Tässä Tilastokeskuksen anonyymit sosioekonomiset väestökarttatiedot on yhdistetty lainadataan. (Varila, K. sähköpostihaastattelu, 8.11.2017.) Tämän tyyppinen kuntien eri asuinalueiden asukkaisiin tutustuminen on kirjastoille itseisarvoista. Jokaisella asuinalueella on omanlaisensa tarpeet ja näihin alueisiin tutustumalla kirjastot voivat tunnistaa tarpeet ja oppia täyttämään ne (Chant 2014).

Lisäksi Vaskissa on tutkittu kaunokirjallisuuden eri genrien suosiota eri kirjastoissa ja sitä, miten aineiston ikä tai tuoreus vaikuttaa sen käyttöön. On myös tarkasteltu mitä kokoelman aihealueita asiakkaat ikäryhmittäin käyttävät. Näitä kaikkia tietoja peilataan kirjaston kokoelmiin, jotta tiedettäisiin kuinka hyvin kunkin kirjaston kokoelmat vastaavat niiden käyttäjien tarpeita (Varila, K. sähköpostihaastattelu, 8.11.2017). Lisäksi

Tilastokeskuksen tietoja hyödyntämällä voidaan ennakoida, missä jokin aihealue tai aineistolaji voisi mennä paremmin lainaan.

Perinteisen kokoelmatyön lisäksi lainadatan avulla on seurattu kuljetuksia eri kirjastojen ja kuntien välillä Vaski-alueen sisällä. Kuljetuksia seuraamalla on voitu esimerkiksi ennakoida käyttöön tulleen kellutuksen vaikutuksia eri kirjastojen kokoelmiin. Lisäksi on tutkittu, miten varausmaksujen poistuminen vuoden 2017 alussa on vaikuttanut kokoelmien käyttöön ja kuljetuksiin. (Varila, K. sähköpostihaastattelu, 8.11.2017.)

Turun pääkirjaston palvelupäällikkö Kalle Varilan mukaan ”kaiken kaikkiaan datan ja sen käsittelytaidon merkitys kirjastotyössä kasvaa: voimme kehittää kirjastopalveluja aiempaa täsmällisemmän ja monipuolisemman tiedon perusteella”.

2.2.2 Tulevia anonyymien lainadatan sovellutuksia

Anonyymilla lainadatalla on kasvava rooli kirjastojen erilaisissa kehityshankkeissa. Monipuolisemmat palvelut vaativat monipuolisempaa tietoa käyttäjistä ja kokoelmista.

Tällä hetkellä Suomessa on vireillä useampikin hanke, jossa pyritään kehittämään kirjastopalveluihin erilaisia virtuaalisia suosittelijasovelluksia. Näissä kaikissa anonyymilla lainadatalla on osansa: Sen avulla voidaan esimerkiksi tarkastella, minkä tyyppistä aineistoa jostain aihepiiristä kiinnostunut ihminen todennäköisesti kaipaa. Ylipäätään asiakaskunnan lainauskäyttäytymisestä tarvitaan tarkempaa tietoa, jotta suosittelijasovellus saadaan toimimaan. (Varila, K. sähköpostihaastattelu, 8.11.2017.)

Lisäksi on kehitteillä aineiston sisätilapaikannukseen pyrkiviä sovelluksia, joissa lainadatalla myös on roolinsa (Varila, K. sähköpostihaastattelu, 8.11.2017). Lainadata tarjoaa tähän valmiiksi digitaalisessa muodossa olevaa, pelkkää luokittelua tarkempaa tietoa aineiston sijainnista. Nämä tiedot ovat helposti vietävissä kartalle.

Myös kokoelman sijainnin vaikutusta sen käyttöön voidaan tarkastella lainadatan avulla (Varila, K. sähköpostihaastattelu, 8.11.2017). Näin olisi mahdollista saada kirjaston tarjoamat aineistot monipuolisemmin käyttöön.

2.3 Vaikuttavuuden mittaaminen

Yleiset kirjastot ovat julkisrahoitteisia intistuuksia eli niitä pyöritetään kansalaisten rahoilla. Tämän takia on tärkeää, että kirjastot voivat osoittaa yleisölle ja päättäjille olevansa hyödyllisiä. Niinpä konkreettiset ja kiistattomat todisteet kirjastojen hyödyllisyydestä ja vaikutusmahdollisuuksista ovat tarpeen. Näitä todisteita on mahdollista saada tilastoista ja lainadatan avulla. Mitattavissa olevat toiminnan tulokset ovat aina vakuuttavia todisteita tarpeesta ja sen täyttymisestä (Dando 2014, 8-9).

2.3.1 Kirjastojen vaikutusalueet vaikuttavuuden mittareina

Mahdollisuus mitata kunkin kirjastoyksikön vaikutusalueita on mahdollistanut myös kirjastojen vaikuttavuuden arvioinnin uudella tavalla. Kunkin kirjaston vaikutusalueelle on laskettu niin sanottu markkinaosuus eli kuinka iso osa alueen väestöstä käyttää kirjastoa. Markkinaosuus on jaoteltu ikä- ja sukupuoliryhmittäin, joten se antaa kirjastoyksiköille mahdollisuuden tarkastella, mitkä keinot voisivat auttaa alueen asukkaiden osallistamisessa. (Kuntaliitto 2016, 80-82.)

Toinen tärkeä vaikuttavuuden mittari yksittäisille kirjastoille on sen paikallinen vahvuus. Kirjaston paikallista vahvuutta mitataan tutkimalla kuinka iso osa alueen lainoista on lainattu nimenomaan paikallisesta kirjastosta. Tämä puolestaan kertoo, kuinka hyvin paikallinen kirjasto tarjoaa alueen käyttäjien tarpeiden mukaisia palveluita ja kokoelmia. (Kuntaliitto 2016, 82.)

Näiden kahden mittarin avulla jokainen kirjastoyksikkö voi helposti mitata, kuinka hyvin se onnistuu aktivoimaan paikallisia käyttäjiään ja vastaamaan heidän tarpeisiinsa (Kuntaliitto 2016, 82).

2.3.2 Kirjaston ympäristövaikutusten tarkastelu lainadatan avulla

Helsingin HelMet-kirjastokimpassa on tutkittu lainadatan avulla kirjaston mahdollisuuksia vaikuttaa käyttäjiensä hiilijalanjälkeen eli siihen, kuinka paljon hiilikuormaa kirjastonkäyttö aiheuttaa. Lainadatan avulla tallentuu kunkin lainassa olevan niteen kohdalle sen omistava yksikkö (kotiyksikkö) sekä sen fyysisesti lainannut yksikkö (lainausyksikkö). Näiden tietojen avulla voidaan tarkastella kuinka suuri osa jostain

yksiköstä lainatuista niteistä on kuljetettu kyseiseen yksikköön asiakkaan tekemän varauksen seurauksena. Koko kirjastoverkoston tasolla voidaan laskea, kuinka paljon kuljetuksia ylipäätään tehdään. (Kuntaliitto 2016, 82-84.)

Näiden tietojen avulla voidaan puolestaan laskea, miten pitkiä matkoja aineistoa kuljetetaan kirjastoverkoston sisällä. Toisaalta voidaan myös laskea, kuinka pitkiä matkoja lainaajat ovat valmiita kulkemaan saadakseen haluamansa aineiston. Sekä käyttäjille että kirjastoille voidaan siis laskea tieverkkoa käyttäen syntyvä matkasumma, joka ilmoitetaan kilometreinä. Ympäristön kannalta on parempi, mitä pienempi käyttäjien matkasumma on: Aineiston kuljettaminen on ekologisempaa, kuin yksittäisen ihmisen liikkuttaminen vastaavan matkan. Tämä johtuu lähinnä siitä, että kuljetuksia voidaan tehdä erissä kun taas asiakas kulkee matkan yksin valitsemallaan tavalla. (Kuntaliitto 2016, 82-84.)

Tästä seuraa, että mitä paremmin kirjastoverkosto onnistuu vastaamaan käyttäjiensä tarpeisiin, sitä pienempi on heidän hiilijalanjälkensä. Myös tiheämpi kirjastoverkosto rohkaisee käyttäjiä kulkemaan jalan tai pyörällä, kun taas harvempi verkosto ohjaa käyttäjät julkisen liikenteen tai oman auton käyttöön. Näin ollen säästöt toimintaverkon tiheydessä siirtyvät ympäristön ja käyttäjien maksettaviksi. (Kuntaliitto 2016, 82-84.)

3 KIRJASTOAUTOJEN VAIKUTUSALUEEN MITTAAMINEN

Suomen kirjastolaissa on määritelty, että ”yleisen kirjaston tulee olla kaikkien käytettävissä ja saavutettavissa” (Laki yleisistä kirjastoista 1492/2016). Tämän lainkohdan toteutumista voidaan tarkkailla tutkimalla jääkö jonkin kunnan alueelle katvealueita, joilla kirjastoa ei käytetä.

Niinpä päätin datan hyödyntämisen havainnollistamiseksi mitata Varsinais-Suomen alueen kirjastoautojen vaikutusalueet. Tähän mennessä oli mitattu vain kiinteiden kirjastoyksiköiden vaikutusalueita, joten kirjastoautot olivat jääneet mittausten ulkopuolelle. Mittaamalla sekä kiinteiden kirjastoyksiköiden että kirjastoautojen vaikutusalueen saa huomattavasti todellisemman kuvan Varsinais-Suomen asukkaiden kirjastonkäytöstä.

3.1 Käytettävissä oleva data

Turun pääkirjastolla, kuten muillakin yleisillä kirjastoilla, on käytössään oman järjestelmänsä tallentamien käyttäjätietojen lisäksi kunnan ja eri instituuttien tarjoamaa tietoa käyttäjäkunnastaan. Turun pääkirjastolla itsellään on varsinaisen kirjastojärjestelmän lisäksi muitakin tiedonkeruu keinoja: Tietokoneiden käyttöä seuraa eBooking-järjestelmä, ovilla on laskurit, jotka mittaavat kävijämääriä ja lisäksi kirjaston palvelutiskeillä pidetään tasaisin välein niin kutsuttu palvelusuoriteotanta. Myös tapahtumista tilastoidaan melko paljon tietoja.

Tässä opinnäytteessä keskityn kuitenkin lähinnä kirjastojärjestelmästä saatavaan anonyymiin lainadataan ja valtion muiden instituuttien tarjoamiin datapaketteihin sekä niiden vertailemiseen. Nämä datapaketit on tallennettu helposti hyödynnettävään muotoon ja saatavilla valmiiksi anonymisoituna. Lisäksi niiden hyödyntäminen onnistuu ilmaisilla avoimen lähdekoodin ohjelmistoilla.

3.1.1 Kirjastojärjestelmän lainadata

Turun pääkirjastolla on käytössään Axiell Aurora -kirjastojärjestelmä. Kirjastot.fi lähettää kirjastolle neljä kertaa vuodessa Aurora-järjestelmästä saatavan otannan. Otanta on taulukkomuotoon tallennettu jonkin hetken läpileikkaus järjestelmästä: Siinä näkyvät kaikki sillä hetkellä järjestelmään kirjatut niteet luettelointitietoineen. Taulukossa on 42 saraketta ja niin monta riviä kuin on järjestelmässä niteitä. Vaski-kirjastokimpan tapauksessa tämä tarkoittaa 2 840 136 riviä eli nidettä.

Tämä lainadata on valmiiksi anonymisoitua, eli se ei sisällä asiakkaan henkilötietoja. Otannan ottamishetkellä ulkona olevien lainojen kohdalla on asiakkaan tiedoista jäljellä paikannustieto, ikä sekä sukupuoli. Nämä tiedot löytyvät otannasta lainassa olevien niteiden kohdalla, muiden niteiden kohdalla nämä sarakkeet ovat tyhjiä. Näin ollen vain asiakkaat, joilla on otannan ottamishetkellä jotain lainassa, näkyvät datapaketissa.

Alkuperäinen lainadata, jossa on asiakkaiden täydetyt tiedot, tuhotaan heti kun anonymisoitu data on saatu tallennettua. Tämän toteuttaa Kirjastot.fi -verkkopalvelu. (Varila, K. sähköpostihaastattelu, 8.11.2017.)

3.1.2 Yhteiskuntarakenne ja kartat

Kaupungin tarjoama Tilastokeskuksen tietoihin perustuva yhteiskuntarakennetta kuvaava datapaketti on käytännössä kartalla käytettävää tietoa: Paketissa on jaettu kunnan alue 250*250 metrin kokoiisiin ruutuihin. Jokaisesta ruudusta ilmoitetaan asukasmäärä sekä ikä- ja sukupuolijakauma. Kirjastojärjestelmän lainadatassa ulkona olevien lainojen korttelitunniste on tällaisen ruudun tunniste. Niiden avulla lainat voidaan liittää oikeaan ruutuun kartalla ja saadaan havainnollistettua vaikkapa jonkin kirjastoyksikön vaikutusalueita.

Näiden ohella hyödynnetään karttoja. Niitä käytetään oikeastaan vain muun datan havainnollistamiseen: Ilman karttaa muu data näkyisi näytöllä vain erikokoisina ruutuina ja pisteinä. Lisäksi kartan avulla on helppo huomata mahdolliset virheet työssä: Jos jokin laina paikallistuu esimerkiksi keskelle merta, voi päätellä, että jossain vaiheessa on tullut virhe.

Yhteiskuntarakennedata ja kartat ovat kirjastolle hyödyllistä pohjatietoa kunnasta: Kirjaston käyttäjätiedot ovat hyödyllisiä lähinnä, jos on jotain mihin niitä voi verrata. Suoraan on esimerkiksi melko helppo tarkistaa kuinka iso osa kunnan asukkaista käyttää kirjastoa.

3.2 Datan hyödyntäminen

Kirjastolla on saatavillaan melko paljon tietoa käyttäjistään ja ympäröivästä kunnasta. Jo yksistään lainadatasta voidaan tarkastella lainamääriä, mitä lainataan mistäkin sekä vaikkapa mitkä asiasanat esiintyvät useimmiten lainoissa. Kun tähän lisätään kaikki muu data, mitä kunnasta on saatavilla, ovat mahdollisuudet erittäin monipuoliset.

Kuitenkin datan hyödyntäminen vaatii taitoa ja erinäisiä ohjelmia. Lisäksi useimmat saatavilla olevat tulokset ovat hyödyllisiä vain, jos mittaamista voidaan jatkaa tasaisin väliajoin; Yksittäisestä otannasta otettu yksittäinen luku ei välttämättä vielä kerro kovin paljon, mutta jos luku voidaan päivittää neljästi vuodessa niin pitkään kuin halutaan, alkaa tuloksella olla merkistystä.

3.2.1 Käytetyt ohjelmistot

Tätä opinnäytetyötä varten käytin kahta ohjelmistoa: **R** sekä **Qgis**. Molemmat ovat ilmaisia, avoimen lähdekoodin projekteja.

R on tilastolliseen laskentaan ja grafiikkaan erikoistunut kieli ja ympäristö. Sen avulla voidaan käsitellä valtavia datamääriä ja luoda grafiikkaa, esimerkiksi erilaisia taulukoita. R:n erilaiset käyttömahdollisuudet ovat valtavat, sillä se on laajennettavissa oleva kokonaisuus. Käyttäjä voi muokata omalla tietokoneellaan olevaa R:ää: Siihen voidaan lisätä käskyjä ja ominaisuuksia joko asentamalla toisten tekemiä laajennuksia tai luomalla omia. (r-project.org 2017).

Tässä opinnäytetyössä R:ää on hyödynnetty oikeastaan taulukkotyökaluna ja laskimena: Lainadatasta voi poimia R:n avulla haluamiaan tietoja ja laskea niistä haluamiaan tuloksia. Esimerkiksi sen tarkistaminen, kuinka iso osa kokoelmasta on lainassa, on suhteellisen yksinkertaista (Kuva 1, Taulukko 1). Toisaalta R:llä voi luoda haluamistaan tiedoista uusia tiedostoja, joita voi sitten hyödyntää muissa ohjelmissa.

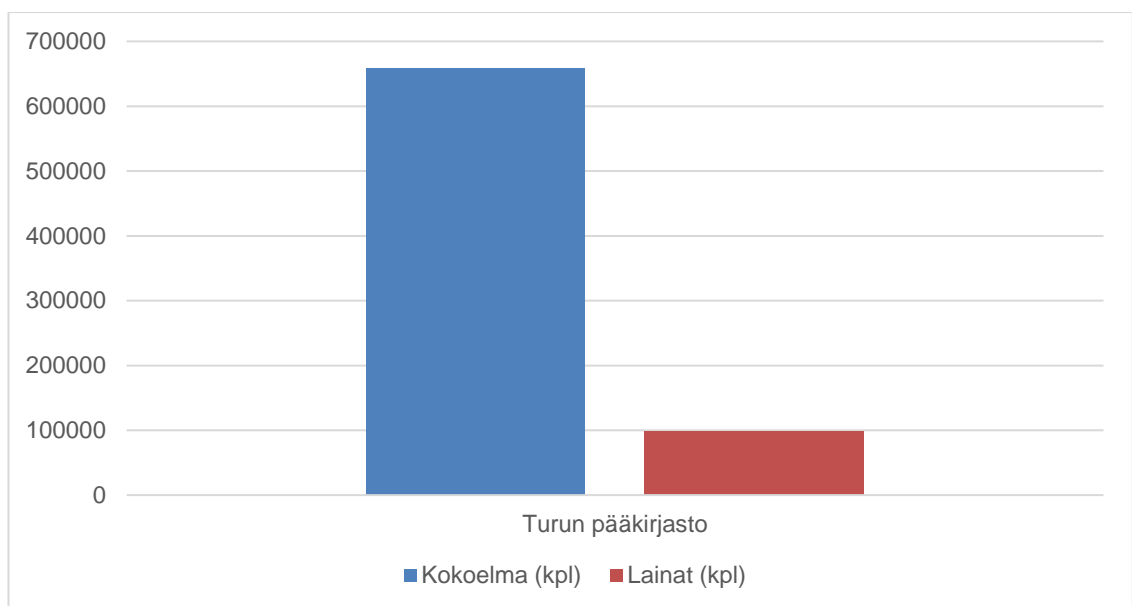
```

Turun pääkirjasto
98002
> x <- subset(a, Niteen.kirjastoyksikkö == "Turun pääkirjasto")
> x1 <- subset(x, Lainaüksikkö == "Turun pääkirjasto")
> dim(x)
[1] 658768    42
> dim(x1)
[1] 98002    42

```

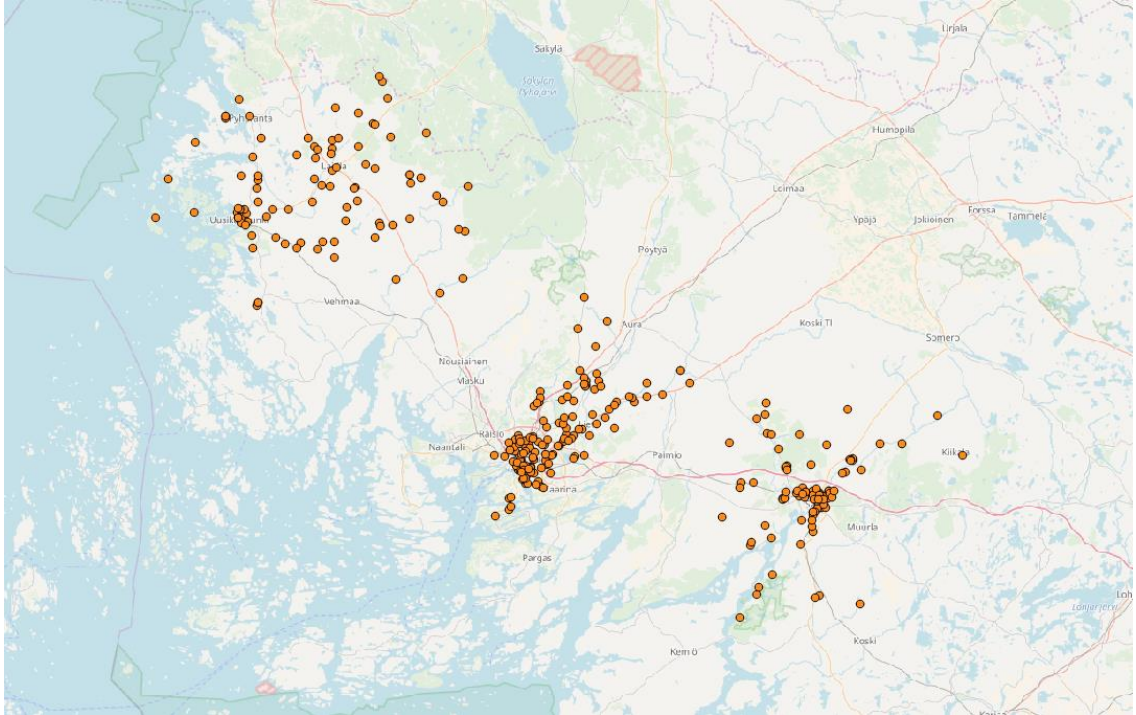
Kuva 1. R:llä erotetut Turun pääkirjaston kokoelmat ja lainat.

Taulukko 1. Turun pääkirjaston kokoelmat ja lainat.



Qgis on maantieteellisen datan käsittelyyn erikoistunut ohjelmisto. GIS tulee sanoista *Geographical Information System*, Q puolestaan on tämän ohjelmiston nimi. Sen avulla voidaan käsitellä monenlaisia tietoja maantieteellisesti: Esimerkiksi listan koordinaattipareja voi näyttää kartalla ja laskea niiden etäisyyksiä toisistaan. Myös Qgis on hyödynnettävissä erittäin monin eri tavoin, sillä myös sitä on mahdollista laajentaa ja muokata omien tarpeidensa mukaisesti. (Qgis.org 2009.)

Qgis:n avulla voi myös viedä R:llä poimittuja tietoja kartalle. Usein R:llä voidaan erottaa esimerkiksi lista koordinaatteja, jotka sitten Qgis-ohjelmiston avulla on helppo asettaa kartalle (Kuva 2).



Kuva 2. Vaski-alueen kirjastoautopysäkit kartalla.

3.2.2 Työvaiheet

Kirjastoautopysäkkien vaikutusalueen mittaamisessa ensimmäinen työvaihe oli erotella R:n avulla lainadatasta kirjastoautojen lainat erilleen. Tämä tapahtui erottamalla kaikista niteistä ne, joissa lainauksikkonä oli jokin kirjastoauto. Niiden niteiden riveillä, jotka eivät ole lainassa, lainausyksikkö sarake on tyhjä.

Kirjastoautoista lainatuista niteistä tehtiin tämän jälkeen oma listansa, josta poistettiin kaikki muut tiedot paitsi niteen nidetunnus, nimike, lainaajan korttelitunniste, ikä ja sukupuoli sekä lainausyksikkö.

Seuraavaksi hankittiin lista kirjastoautopysäkkien osoitteista. Näille osoitteille haettiin Google Mapsin avulla koordinaatit, jotta niitä voidaan hyödyntää Qgis:ssä.

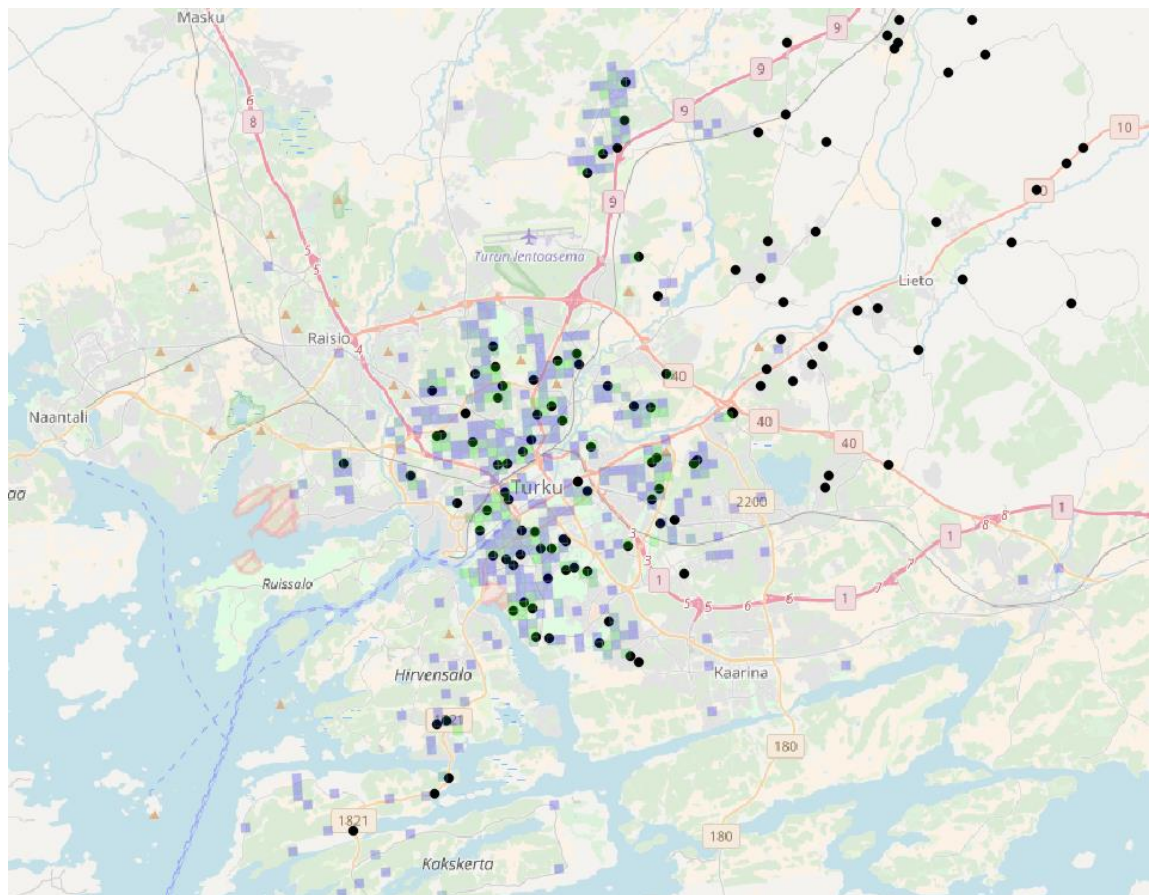
Seuraavaksi Qgis:n vietiin yhteiskuntarakennedata, kirjastoautopysäkkien koordinaatit ja kirjastoautojen lainat sekä kartta. Yhteiskuntarakennedata näkyy kartan päällä koko Varsinais-Suomen kattavana ruudukkona, kirjastoautopysäkit koordinaattien mukaan sijaitsevana pisteinä. Lainat puolestaan liitetään yhteiskuntarakennedatan ruutuihin.

Tämän jälkeen erotetaan eri lainausyksiköt eli kirjastoautot toisistaan (esimerkiksi Turun kirjastoauto tai Liedon kirjastonauto). Tämän jälkeen voi valita, minkä auton lainat haluaa kartalla nähdä.

3.3 Tulokset

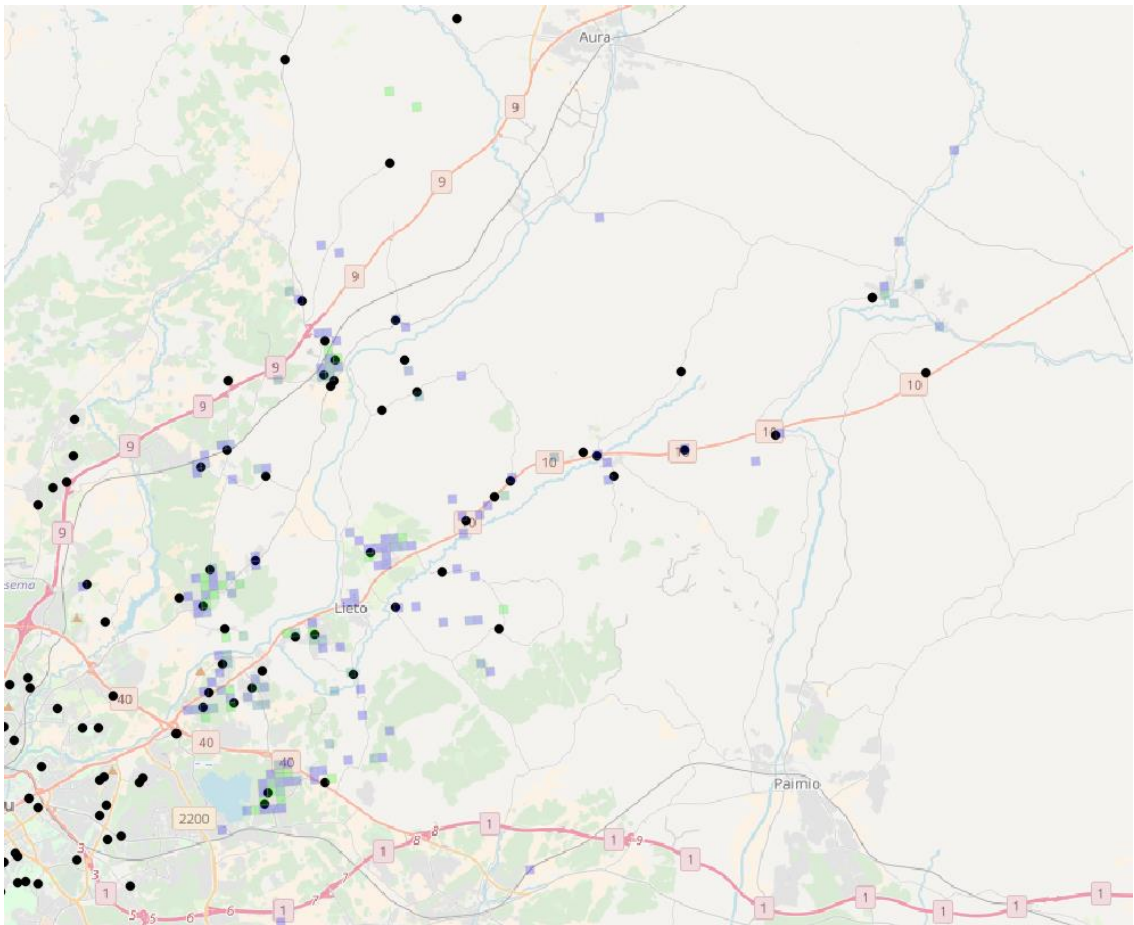
Vaski-alueen kirjastoautojen vaikutusalueet ovat erittäin vaihtelevia. Joidenkin pysäkkien ympärille muodostuu selkeitä vaikutusalueita, toisaalta jotkin lainat näyttävät olevan pitkän matkan päässä mistään pysäkistä.

Aktiivisimmin kirjastoautoa käytetään Turun alueella (Kuva 3). Kuvan liiloissa ruuduissa on lainoja 1-10 kappaletta, vihreissä yli 20. Ruudun lainat voivat olla usealla asiakkaalla. Kartalla mustat pisteet edustavat kirjastoautopysäkkejä.



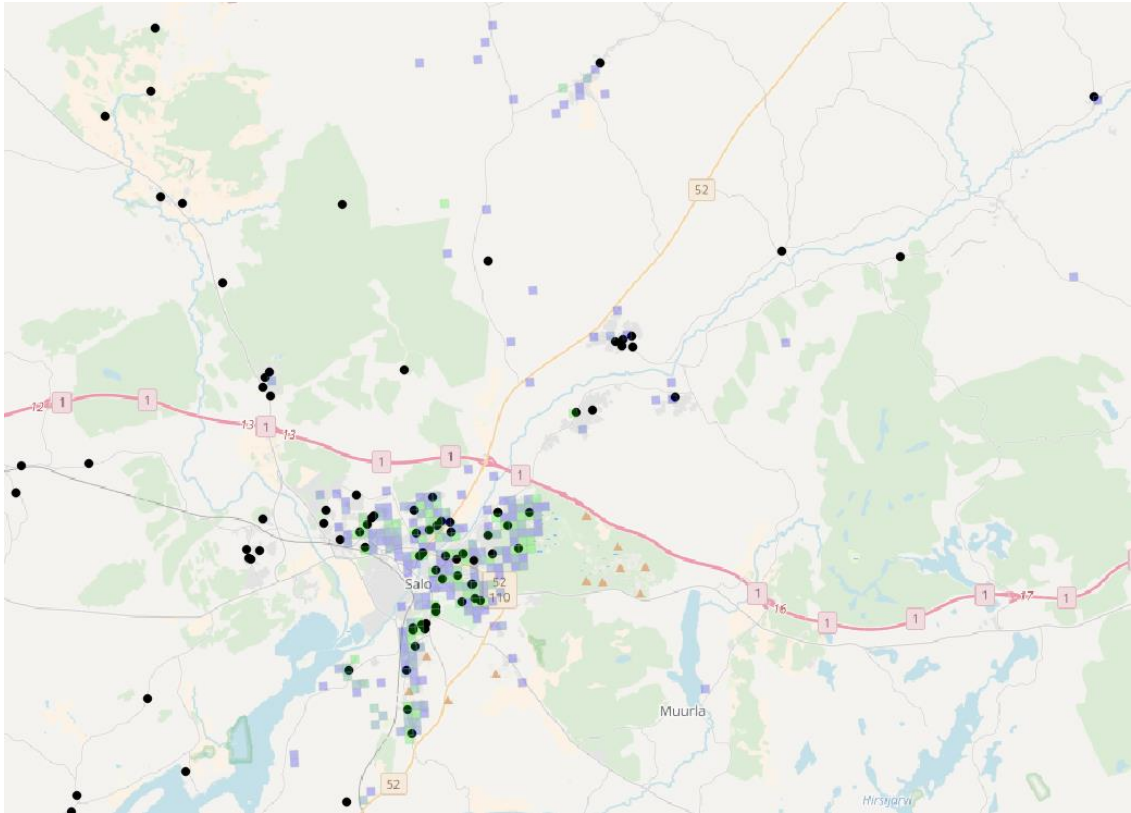
Kuva 3. Turun kirjastoauton vaikutusalue.

Sen sijaan esimerkiksi Liedon kirjastoauton vaikutusalue on huomattavasti hajanaisempi ja lainoja on myös selkeästi vähemmän (Kuva 4). Toisaalta vaikuttaisi siltä, että pysäkeillä käydään pitkienkin matkojen päässä ja melko monet ruuduista ovat vihreitä eli lainoja on yli 20.



Kuva 4. Liedon kirjastoauton vaikutusalue.

Salon kirjastoauton käyttö vaikuttaisi olevan kiintynyt keskustan alueelle ja muutaman kauempana keskustasta olevan pysäkin ympärille (Kuva 5). Käyttö on selvästi aktiivisempaa kuin esimerkiksi Liedon autolla ja hajontaa on vähemmän. Tästä voi päätellä auton onnistuneen pysäkkisijoittelussaan melko hyvin.



Kuva 5. Salon kirjastoauton vaikutusalue.

Verrattuna kiinteiden kirjastojen vaikutusalueisiin, ovat kaikkien autojen vaikutusalueet hajanaisia ja melko harvoja. Kuitenkin autojen vaikutusalueet yltävät kauas kuntien keskustoista ja lainamäärät ovat siellä täällä suuriakin. Jotkin pysäkit vaikuttavat tämän otannan perusteella melko käyttämättömiltä, mutta toisaalta alueita ja niiden kehittymistä pitäisi seurata, jotta tulkinnot olisivat luotettavia. Kuitenkin tämä metodi on osoittautunut hyväksi tavaksi tarkkailla eri kirjastoautojen vaikutusalueita ja lainamääriä.

4 POHDINTAA

Kaikilla Suomen yleisillä kirjastoilla on lähitulevaisuudessa käytettävissään valtavasti enemmän ja tarkempaa tietoa käyttäjistään ja heidän lainakäyttäytymisestään anonyymien lainadatan muodossa. Lainadata mahdollistaa palveluiden kehittämisen asiakaslähtöisempään suuntaan rikkomatta kuitenkaan käyttäjien yksityisyyden suojaa. Kunkin kirjastoyksikön tai -kimpan käyttäjäkunnan tarpeista saadaan aiempaa tarkempaa ja monipuolisempaa tietoa, jolloin näiden tarpeiden täyttäminen on helpompaa.

Palveluissa tapahtuvien muutosten seurauksia voidaan ennakoida ja tarkastella huomattavasti tarkemmalla tasolla. Koska kirjasto joutuu uudistumaan jatkuvasti, on tärkeää saada tietoa uudistusten vaikutuksista käyttäjiin ja heidän käyttäytymiseensä. Lainadatan avulla voidaan melko nopeasti päätellä, kannattiko jokin uudistus vai ei.

Toisaalta kirjasto joutuu jatkuvasti perustelemaan toimintansa tarpeellisuutta lainsäätäjille ja muille päättäjille sekä toki tavallisille kansalaisille. Lainadataa yhdistelemällä muihin valtion keräämiin tietoihin asukkaistaan, voidaan osoittaa kirjaston pystyvän sille annettuun osallistavaan tehtävään yhteiskunnassa. Datan avulla voidaan mitata ja osoittaa kirjaston vaikutus kansalaisten elämässä.

Hyödyt kirjastolle ovat ilmeiset, joten kirjastojen kannattaisi alkaa ottaa anonyymi lainadata laajemmin huomioon toimintasuunnitelmissaan. Toki monet lainadataa hyödyntävät projektit tulevat tuomaan sen kirjastolaisten tietoisuuteen laajemmin. Tulevaisuudessa voisi kuvitella lainadatan olevan osa lähes joka kirjaston arkea.

LÄHTEET

Chant, Ian. "The Numbers Game: The Art and Science of Using Data to Improve Public Library Services.(DATA-DRIVEN LIBRARIES)." *Library Journal* 139, no. 8 (2014): 28

Dando, Priscille. *Say It With Data: A Concise Guide to Making Your Case and Getting Results*. Chicago: ALA Editions, 2014.

Kirjastot.fi 2017. Viitattu 4.12.2017 <https://www.kirjastot.fi/> > Kirjastopro > Tilastot

Kuntaliitto. Hyötyä, tietoa, elämyksiä – Kirjastojen vaikuttavuuden ulottuvuuksia. Helsinki: Suomen Kuntaliitto, 2016. Saatavilla sähköisesti osoitteessa <http://vaikuttavuus.kirjastot.fi/>

Laki yleisistä kirjastoista 1492/2016. Annettu Helsingissä 29.12.2016. Saatavilla sähköisesti osoitteesta <http://www.finlex.fi/fi/laki/ajantasa/2016/20161492>

Qgis.org 2009. Viitattu 20.10.2017 <http://docs.qgis.org/2.18> > User guide/manual (QGIS 2.18) > Foreword.

R-project.org 2017. Viitattu 25.10.2017 <https://www.r-project.org> > About R.

Virtuaali ammattikorkeakoulu 2007. Viitattu 5.12.2017 <http://www2.amk.fi/digma.fi/www.amk.fi/opintojaksot/0709019/1193463890749.html> > Kvantitatiivisen analyysin perusteet.