

Opinnäytetyö (AMK)

Diakin viestinnän koulutusohjelma

Journalismi

2014

Esa Töykkälä

PAIKALLISTA DATAA

– datajournalismin tekeminen pienessä
paikallistoimituksessa



TURUN AMMATTIKORKEAKOULU
TURKU UNIVERSITY OF APPLIED SCIENCES

Esa Töykkälä

PAIKALLISTA DATAA

Datajournalismi yleistyy maailmalla ja Suomessa. Suomessa erityisesti Helsingin Sanomat ja Yle ovat tehneet isoja datajournalistisia projekteja. Suurissa toimituksissa on suuret resurssit, mutta miten datajournalismia pystyy tekemään pienessä paikallismedian toimituksessa?

Se on keskeinen kysymys, johon opinnäytetyö etsii vastausta. Lisäksi opinnäytetyö pyrkii selvittämään esimerkiksi esteitä, joihin paikallistoimittaja voi datajournalismia tehdessään törmätä ja sitä, miten datajournalistinen työprosessi paikallistoimituksessa etenee. Millaista datajournalismia paikallistoimituksessa ei pystyisi tekemään? Miten datajournalismia voi soveltaa paikallisjournalismiin, ja miksi se on tärkeää? Mikä datajournalismissa on vaikeinta?

Opinnäytetyötä varten on haastateltu turkulaisen Aamuset-kaupunkilehden toimitussihteeri Teemu Peltolaa. Hän on toimittanut lehdelle noin parikymmentä uutisjuttua, joiden tekemisessä on hyödynnetty datajournalistisia työkaluja.

Opinnäytetyön tuoteosa on Aamuset-lehdelle tehty datajournalistinen uutisjuttu, joka käsittelee Turun lentoaseman matkustajamääriä vuosina 1998 – 2013. Opinnäytetyö kuvailee tuoteosan työprosessin eri vaiheita. Viimeinen luku käsittelee tuoteosan tekemisen herättämiä huomioita ja pohtii, mitä hyötyä niistä voisi olla muille paikallistoimituksille, jotka harkitsevat datajournalismin tekemistä.

Haastattelun ja tuoteosan prosessista voidaan päätellä, että datajournalismi on mahdollista myös pienessä paikallistoimituksessa, mutta tietyin rajoituksin. Erityisen tärkeää on ottaa huomioon, että pienessä toimituksessa datajournalisti työskentelee hyvin todennäköisesti yksin. Datajournalismin tekeminen voi riippua jossain määrin esimerkiksi toimittajan omasta osaamisen tasosta, mutta ei ainakaan työkaluista, sillä monet niistä ovat käytettävissä internetissä ilmaiseksi.

ASIASANAT:

Datajournalismi, paikallisjournalismi, avoin data

BACHELOR'S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Media and Communications | Journalism

2014 | 28

Samuel Raunio

Author(s)

LOCAL DATA

Data journalism is becoming more common in the whole world and in Finland. In Finland especially Helsingin Sanomat and Yle have produced big datajournalism projects. In big news desks the resources are big, but how does one produce data journalism in a small local media news desk?

That is an essential question to which this bachelor's thesis attempts to find an answer. In addition to that the bachelor's thesis tries to find out obstacles that a local journalist might face when producing data journalism and clarify the working process of data journalism in a local media news desk. What kind of data journalism could not be done in a local media news desk? How can data journalism be applied to local journalism and why is it important? What's the most difficult aspect in data journalism?

Teemu Peltola, who is a subeditor in a local free newspaper called Aamuset in Turku, was interviewed for the bachelor's thesis. For the newspaper he has produced approximately twenty news stories that have somehow utilized data journalism tools.

The product part of the thesis is a data journalism newspiece made for Aamuset which is focused on the amount of passengers at the airport of Turku in 1998 - 2013. The bachelor's thesis describes the different phases of the product part's process. The last chapter deals with the observations that have come from making the product part and discusses the possibility of them being useful to other local news media desks that might be considering producing data journalism.

From the interview and the product part it can be deduced that it is possible to produce data journalism in a small local news media desk, but with certain restrictions. It's particularly important to take into consideration that in a small news desk a data journalist most likely works alone. The making of data journalism can to some extent depend on for example the level of the journalist's own skills, but not at least on the tools, since most of them can be used on the internet for free.

KEYWORDS:

Data journalism, local journalism, open data

SISÄLTÖ

1 JOHDANTO	5
2 MITÄ DATAJOURNALISMI ON?	7
2.1 Mistä dataa saa?	8
2.2 Miksi datajournalismi on tärkeää?	9
3 AAMUSET-KAUPUNKILEHTI	13
3.1 Datajournalismi Aamuset-lehdessä	13
3.2 Ensimmäinen datajournalistinen juttu Aamuset-lehdessä	15
3.3 Datajournalismin ongelmat ja esteet Aamuset-lehdessä	15
4 DATAJOURNALISTINEN JUTTU AAMUSET-LEHDELLE	18
4.1 Datan löytäminen ja käsitteleminen	18
4.2 Tiedonhankinta datan saamisen jälkeen	22
5 JÄLKIPOHDINTA: DATA AUTTAA IDEOINNISSA	24
5.1 Datajournalismin tekeminen paikallistoimituksessa	25
LÄHTEET	27

LIITTEET

Liite 1. Opinnäytetyön kirjallisuuskatsaus.

KUVIOT

Kuvio 1. Turun lentoasemalla kotimaan lennoilla matkustaneiden määrät.	20
Kuvio 2. Turun lentoasemalla ulkomaan lennoilla matkustaneiden määrät.	21

1 JOHDANTO

Opinnäytetyön keskeinen tavoite on tutkia, millaisia mahdollisuuksia pienellä paikallismedian toimituksella on datajournalismia hyödyntävien uutisjuttujen tekemistä ja kokeilua varten. Paikallismedioiden toimitukset ovat usein pieniä, joten merkittävä osa opinnäytetyöstä keskittyy siihen, kuinka hyvin paikallismedian toimittaja kykenee työskentelemään datajournalismin parissa yksin. Tarkastelussa ovat datajournalistisen työn vaiheet, jotka voidaan tiivistäen jakaa kolmeen osaan: datan hankkiminen, sen käsitteleminen sekä datajournalistisen jutun julkaisu.

Datajournalismi on yleistynyt viime vuosina paitsi Suomessa, myös maailmalla. Suomessa datajournalismin tekemiseen ovat erikoistuneet kenties näkyvimmin Helsingin Sanomien kesällä 2012 perustettu datatoimitus sekä tammikuussa 2013 perustettu Ylen Plus-toimitus. Tällaisessa työskentely-ympäristössä datajournalistilla on tukena työtovereiden apu, mikäli hän kohtaa työssään esteitä vaikkapa datan hankkimisessa; esimerkiksi Ylen Plus-toimituksessa työskentelee graafikkoja, tuottajia ja yksi datajournalisti. (Koskenranta 2013, 7.)

Opinnäytetyö pyrkii selvittämään muun muassa sitä, minkä laajuisia data-aineistoja paikallisjournalisti pystyy käsittelemään työssään paikallismedian resurssit ja ajankäytön huomioon ottaen. Toisin sanoen erityisenä tarkastelun kohteena on kartoittaa sellaisia konkreettisia tilanteita, joissa paikallistoimittaja ei enää pysty jatkamaan datajournalismin tekemistä – tämä voi johtua esimerkiksi datan suuresta koosta liian tiiviin työskentelyaikataulun puitteissa tai datan hankalasta saatavuudesta.

Opinnäytetyö keskittyy turkulaisen Aamuset-kaupunkilehden tapaan tehdä datajournalismia. Aamuset on julkaissut noin kaksikymmentä juttua, joissa datajournalistiset prosessit ovat olleet jollain tapaa osa toimittajan työskentelyä. Opinnäytetyön tuoteosa on Aamuset-lehdessä julkaistu datajournalistinen juttu, jonka vaiheita analysoidaan tarkemmin opinnäytetyön tutkimustekstissä: Esimerkiksi mistä ideat syntyvät datajournalismissa, ottaako toimittaja ensin

käsiteltäväkseen datan ja saa idean vasta sen jälkeen tai millaisiin esteisiin datajournalismia tehdessä voi törmätä.

Opinnäytetyön tutkimusmenetelmänä on käytetty haastattelua. Aamuset-lehden datajournalismia käsittelevät tiedot on saatu haastatteleamalla lehden toimitussihteeri Teemu Peltolaa. Datajournalismin käsitteleminen paikallisjournalismin näkökulmasta on kiinnostavaa ja ajankohtaistakin, sillä kuten luvun 2.2 lopussa todetaan, datajournalismilla uskotaan olevan merkittäviä vaikutuksia paikallistason journalismiin.

Datajournalismin käsitteleminen opinnäytetyönä on hankalaa, sillä kyse on verrattain uudesta ilmiöstä (osittaiset syyt tähän perustellaan luvussa 2.1), joka on ainakin toistaiseksi keskittynyt Suomessa tiettyihin ydintoimituksiin, kuten juuri Ylelle ja Helsingin Sanomiin. Toisaalta juuri tästä syystä opinnäytetyön näkökulma rajautui datajournalismin ja paikallisjournalismin yhdistelmään, joka tuntuu olevan paikallismedioiden toimituksissa vielä varsin harvinainen.

Datajournalismin yleistymisen esteenä ei ole pelkästään toimittajien osaamisen puute, vaan myös julkisen tiedon avaaminen Suomessa. Julkisen datan avaamiseen liittyvää aktiivista yhteisötoimintaa on esimerkiksi Tampereella¹, mutta Turussa hyvin vähän tai ei miltei ollenkaan. Opinnäytetyön tuoteosana on Aamuset-lehdessä julkaistu datajournalistinen juttu, jota varten käsitelty data on saatu käyttöön Lounaispaikka-tilastopalvelusta. Opinnäytetyössä on pyritty käsittelemään Aamuset-lehden kaltaiselle paikallismedialle tehdyn datajournalistisen uutisprosessin vaiheet mahdollisimman tarkasti. Viimeisessä luvussa analysoidaan prosessin herättämiä ajatuksia ja sitä, mitä kenties muut paikallistoimitukset voisivat oppia siitä.

¹ Esimerkkisivustoja: <http://www.tampere.fi/tampereinfo/avoindata.html>
<http://www.hermiagroup.fi/opendatatre/>

2 MITÄ DATAJOURNALISMI ON?

Datajournalismi on journalismin osa-alue, joka hyödyntää suurten tietokantojen ja datamäärien käsittelyä uutistyössä. Datajournalismille keskeisiä työkaluja ja tietokoneohjelmia tarvitaan silloin, kun käsiteltävän datan määrä kasvaa niin suureksi, ettei sitä ole enää mahdollista analysoida ja suodattaa käsin ilman tietoteknistä apua. Esimerkiksi The Guardianin entinen datajournalisti ja nykyään Twitterin dataosastolla työskentelevä Simon Rogers käsitteli WikiLeaksille vuodetuista Afganistan-asiakirjoista kertovaa uutista varten Excel-tiedostoa, joka sisälsi 92 201 rivin verran dataa (Rogers 2011, 252).²

Data voi toimia joko journalistisen jutun lähteenä, välineenä, jonka avulla juttu kerrotaan, tai se voi olla molempia (Gray ym. 2012, 274). Datajournalismia kuvaavat määrittelyjen sijaan paremmin konkreettiset työskentelytavat ja työkalut, joiden avulla sitä tehdään.

Yksi tärkeimpiä datajournalistisia työkaluja on taulukkolaskentaohjelma, kuten Microsoft Excel, mutta vastaavia työkaluja pystyy käyttämään maksuttomasti esimerkiksi Google Docs -palvelun kautta. Esimerkiksi suuri osa The Guardianissa tehtävästä datajournalismista tapahtuu Excelin avulla (Rogers 2011, 1084). Myös Yle on käyttänyt maksuttoman Google Fusion Tables -ohjelman apua erityisesti datajournalismissa, joka hyödyntää karttatietoja (Sjöholm 2013). Koska monet datajournalismiin liittyvät työkalut ovat ilmaisia ja helposti saatavilla, on varsin epätodennäköistä, että paikallistoimittajan työ datajournalismin parissa pysähtyisi välineiden puutteeseen.

Eastern Illinois Universityn dosentti Bryan Murley näkee datajournalistisessa työssä kolme vaihetta: Datan keräämisen, analysoimisen sekä esittämisen. Esittämistapoja on monia, mutta hänen mukaansa pelkkä datasta kirjoittaminen

² Opinnäyteyössä käytetyt lähteet Gray ym, Reimold ja Rogers ovat e-kirjalähteitä. Sivunumeron sijaan näihin teoksiin viittaaviin lähdeviittauksiin on kirjoitettu sijainti (englanniksi location). Lähdeteokseen viitattu kohta löytyy syöttämällä lähdeviitteen sijaintinumero e-kirjan lukulaitteeseen.

ei riitä. (Reimold 2013, 3497.) Simon Rogersin mielestä taas datajournalismin esittäminen voi hyvin toimia tekstinäkin. (Rogers 2011, 143).

Iso osa datajournalistisesta prosessista on datan käsittelemistä ja puhdistamista. Datan taulukkoa voi joutua järjestelemään esimerkiksi postinumeroiden mukaan, jotta kokonaisuus hahmottuu paremmin. Bryan Murley neuvookin käyttämään paljon aikaa siihen, ettei datasta tehtäviin grafiikoihin tai statistiikoihin pääse lipsahtamaan virheitä. Monet yksilöt ja ryhmät haluavat manipuloida dataa, jotta se näyttäisi julkisuudessa parempana, minkä takia datan taustat ja lähteet tulee tarkistaa huolellisesti. Tästä esimerkkinä Murley käyttää yhdysvaltalaisen yliopiston kampusalueilla tapahtuvista rikoksista kertovia tilastoja. Koulujen työntekijät ja kampusten turvallisuudesta vastaavat viranomaiset pystyvät Murleyn mukaan muokkaamaan lukuja monella tapaa niin, että rikosten todellinen laajuus ei näyttäydy lopullisissa tilastoissa. Tällöin kampusalueen rikostilastoja kannattaa verrata esimerkiksi saman kaupungin poliisille tulleiden rikosilmoitusten määrään. (Reimold 2013, 3507, 3547.) Datajournalismissa tulee muistaa osoittaa tarpeellista kriittisyyttä lähteenä toimivaa dataa kohtaan, aivan kuten esimerkiksi yritystenkin tiedotteisiin suhtaudutaan. Datajournalistin tärkein ja mahdollisesti vaikein tehtävä onkin esittää datalle oikeita kysymyksiä (Vehkoo 2011, 107). Mikäli jokin data ei tunnu täsmäävän, se voi jopa johdattaa kokonaan uuden uutisjutun lähteelle ja herättää kysymyksiä siitä, miksi dataa on alun perin haluttu vääristellä.

Journalistien tehtävänä on aina ollut hankkia tietoa lähteiltä ja esittää se yleisölle helpommin ymmärrettävässä muodossa esimerkiksi uutisena. Tässä mielessä datajournalismissa ei ole mitään uutta, mutta lähteiltä saatavan ja käsiteltävän tiedon määrä sekä käsittelytavat ovat datajournalismissa poikkeuksellisen suuria.

2.1 Mistä dataa saa?

Simon Rogersin mukaan datajournalismi on noussut yleiseen tietoisuuteen niin sanotun läpinäkyvyyssajattelun maailmanlaajuisen leviämisen myötä. Datan

saatavuus ja sen helppokäyttöisyys sekä yleinen mielenkiinto datan visualisoimista kohtaan ovat lisääntyneet, kuten myös laajat uutisjutut, joiden tekemiseen on käytetty suuria määriä dataa. (Rogers 2011, 112.)

Englantilainen The Guardian on ollut datajournalismia tehdessään edullisessa asemassa, sillä datan avaaminen julkiseen käyttöön on ilmiö, joka on levinnyt erityisesti Isossa-Britanniassa sekä Yhdysvalloissa. Myös Suomessa hallitus teki vuonna 2011 periaatepäätöksen julkishallinnon datan avaamisesta. (Vehkoo 2011, 96.) Mitä enemmän ja helpommin dataa on saatavilla, sitä helpommaksi myös datajournalistin työ muodostuu. Suomessa muun muassa Ilmatieteen laitos, Maanmittauslaitos ja Helsinki Region Infoshare ovat avanneet näkyvästi dataa julkiseen käyttöön.³

Datajournalismin voidaan katsoa yleistyneen kahdesta syystä: Datan saatavuus on parantunut ja datan käsittelyä varten tarvittavat ohjelmat ovat nykyään saatavilla kenelle tahansa, jolla vain on käytössään tietokone. Datajournalismi on journalismia, joka hyödyntää molempien ilmiöiden yleistymistä. Datajournalismi tällaisenaan on verrattain uusi ilmiö, sillä esimerkiksi Yhdysvaltain avoimen datan sivusto <http://www.data.gov/> aukesi vasta vuonna 2009 (Rogers 2011, 112).

Datan avaamiseen liittyvät oleellisesti myös taloudelliset seikat. Läpinäkyvyyden ja demokratisoivan vaikutuksen lisäksi avoin data voi luoda uutta liiketoimintaa. Euroopan komission mukaan datan avaamisesta voi koitua jopa 140 miljoonan euron vuosittaiset hyödyt. Toisaalta Maanmittauslaitos menettää vuodessa noin 1,5 miljoonaa euroa juuri sen takia, että se on avannut datansa maksuttomaan käyttöön. (Salminen 2012.)

2.2 Miksi datajournalismi on tärkeää?

Arvioiden mukaan maailmaan tuotetaan joka toinen päivä noin viiden eksabitin eli runsaan viiden miljardin gigabitin verran informaatiota. Ihmiset tuottavat osan informaatiosta esimerkiksi sosiaaliseen mediaan tekstipäivityksinä ja kuvina.

³ Ilmatieteen laitoksen avoin data: <https://ilmatieteenlaitos.fi/avoin-data>
Maanmittauslaitoksen avoin data: <http://www.maanmittauslaitos.fi/avoindata>
Helsinki Region Infoshare: <http://www.hri.fi/fi/>

(Vehkoo 2011, 95.) Datajournalismi auttaa jäsentelemään ja analysoimaan alati kasvavaa datan määrää. Yksi journalismin perustehtävistä on aina ollut pyrkiä selittämään maailmaa. Nyt datajournalismia tarvitaan selittämään, mitä miljardit gigabitit voivat kertoa maailmasta.

Saksalaiselle Bertelsmann-säätiölle työskennellyt Tom Fries puhuu ”informaation epäsymmetriasta”: Uutta tietoa syntyy niin paljon ja niin nopeasti, että se vaikuttaa ongelmallisella tavalla jo ihmisten päivittäisiin päätöksiin ja toimintoihin. Hyvin tehty datajournalismi helpottaa ihmisiä informaatiotulvan keskellä. (Gray ym 2012, 366.) Googlen pääekonomisti Hal Varianin mukaan ongelmana ei olekaan enää datan puute, vaan ”dataähky” (Vehkoo 2011).

World wide webin keksijä Tim Berners-Lee uskoo vahvasti datajournalismin olevan olennainen osa journalismin tulevaisuutta ja väittää, että toimittajien täytyy opetella ymmärtämään tietokantoja ja koodia (Vehkoo 2011, 107). Datajournalismiin erikoistuva toimittaja voi olla nyt ja tulevaisuudessa edullisessa asemassa hakiessaan töitä. Toisaalta tietualan osaajille riittää kysyntää muutenkin: Deutsche Wellen toimittaja Mirko Lorenz huomauttaa, että datajournalistisista taidoista voi olla hyötyä muuallakin kuin toimittajan töissä, esimerkiksi viestintä- ja tiedotusosalalla, sillä moniin yrityksiin ja järjestöihin kaivataan paljon työntekijöitä, jotka osaavat muokata datasta ymmärrettävää (Gray ym. 2012, 302).

Simon Rogers kuitenkin huomauttaa, että toimittajan ei tarvitse olla koodari tai ohjelmoija pärjätäkseen datajournalismissa. Hän painottaa, kuinka tärkeää on pikemminkin tarkastella dataa journalistin eikä analyytikon näkökulmasta (Rogers 2011, 74). Media Standards Trust -järjestön johtajan Martin Mooren mielestä toimittajien täytyy tulevaisuudessa ymmärtää koodin ja algoritmien merkityksiä, mutta silti hänen mielestään ammattitaitoisen journalistin ja koodia osaavan ohjelmoijan yhdistelmä on harvinainen. Mooren mukaan toimitukset tulevat tavitsemaan ryhmiä, jotka koostuvat molempien osa-alueiden osaajista. (Vehkoo 2011, 110.) Suomessa tällaisia ryhmiä on ainakin edellä mainittujen Helsingin Sanomien ja Ylen datatoimituksissa.

Hollantilaisen NU.nl- uutissivuston datajournalisti Jerry Vermanen väittää, että datajournalismi tulee entisestään vahvistamaan journalismin ”vallan vahtikoiran” roolia erityisesti taloudellisesti vaikeina aikoina. Samaa mieltä on datajournalismiin erikoistuneen Journalism++-verkoston toimittaja Nicolas Kayser-Bril. Hänen mukaansa juuri journalistien kyvyttömyys ymmärtää numeroita johti siihen, miksi esimerkiksi yhdysvaltalaisen Enron-energiayhtiön kirjanpitoskandaalia⁴ sekä Bernard Madoffin tapausta⁵ käsittelevistä uutisista puuttui aikoinaan tietty journalistinen oivalluskyky. (Gray ym. 2012, 355, 377.)

Myös The Seattle Times -lehden toimittaja Cheryl Phillips on samaa mieltä. Hänen mukaansa datajournalismi paljastaa sellaisia asioita, jotka jäisivät tavalliselta journalismilta piiloon, kuten ympäristöä haittaavat poliittiset päätökset tai huume politiikan aiheuttamat kuolemat (Gray ym. 2012, 443). Data voi kätkeä taakseen paljon erilaisia uutisten aiheita.

Datajournalismilla on myös hyvin ajankohtainen tehtävä. Uutisorganisaatiot kilpailevat uskottavuudesta kansalaisjournalistien ja bloggaaajien kanssa. Tämä on johtanut siihen, että uutistoimistot pyrkivät olemaan yhä avoimempia uutisten lähteiden suhteen. (Rogers 2011, 173.) Valittujen Palojen ja sen sisarlehtien julkaiseman tutkimuksen mukaan vain joka neljäs suomalaisista luottaa toimittajiin (Aro 2012). Koska uutisten kuluttajista on tullut aiempaa mediakriittisempiä ja luottamus toimittajia kohtaan on vähentynyt, datajournalismilla on mahdollisuus ansaita lukijoiden luottamus numeroihin ja dataan perustuvalla tiedonvälityksellä. Datajournalistiset jutut kertovat lukijoille avoimesti, mistä tieto on peräisin. Kuten aiemmin on mainittu, journalistien tärkeimpänä tehtävänä on analysoida dataa kriittisesti itse jutuntekovaiheessa ja päättää, millaisessa muodossa se esitetään lukijoille. Virheellisesti esitetty data voi sekin murentaa julkaisun luotettavuutta.

Datajournalisti Jerry Vermasen mukaan datajournalismilla on ratkaiseva rooli paikallisjournalismissa. Vermasen mielestä journalistien tulee osata etsiä,

⁴ Yhdysvaltalaisen Enron-energiayhtiön valtava talousskandaali paljastui vuonna 2001 ja johti yhtiön konkurssiin samana vuonna.

⁵ Yhdysvaltalainen Bernard Madoff on NASDAQ-pörssin entinen johtaja ja sijoitushuijari, joka tuomittiin 150 vuoden vankeuteen kaikkiaan 11 talousrikoksesta.

analysoida ja visualisoida paikallistason dataa, sillä paikallisjulkaisuilla ja niiden uutisilla on suora vaikutus alueensa asukkaisiin. (Gray ym. 2012, 366.)

Vermasen näkemys on perusteltu: Suomessa kunnallispolitiikalla on usein paljon suurempi vaikutus asukkaan arkielämään esimerkiksi terveydenhuollon, koulutuksen tai muiden kunnallispalvelujen kannalta kuin eduskunnassa tehtävillä päätöksillä. Koska tiedon määrä kasvaa kovaa vauhtia, journalistien tulee osata esittää kysymyksiä myös paikalliselle datalle paikallisjournalismin keinoin.

3 AAMUSET-KAUPUNKILEHTI

Aamuset on Turussa ja sen lähiseudulla ilmestyvä kaupunkilehti, joka kuuluu TS-Yhtymään. Aamuset-lehden painos on noin 100 100 – 140 100 ja se ilmestyy keskiviikkoisin ja lauantaisin. Aamuset-lehti ilmestyi ensimmäisen kerran vuonna 1981, mutta se alkoi muokata journalistisia sisältöjään kunnolla verkkoon keväällä 2012. Aiempina vuosina Aamuset oli julkaissut verkkosivuillaan lähinnä joitakin printtilehdessä julkaistuja uutisia ja artikkeleita suoraan. Vuonna 2012 tapahtuvien muutosten aikana Aamuset-lehden juttujen verkkoversioiden otsikoita alettiin muuttaa erilaiseen muotoon verkkosivuja varten ja artikkeleita ruvettiin muutenkin stilisoimaan paremmin verkkoon sopiviksi. Lisäksi osa jutuista saattaa olla nykyään verkossa pidempiä, koska verkossa ei tarvitse huomioida printtilehden taiton asettamia merkkirajoituksia. Ennen kevättä 2012 Aamuset-lehteä pystyi lukemaan verkkosivuilta näköislehdestä. Aamuset-lehden vakituisia työntekijöitä journalistisessa työssä ovat päätoimittaja Lasse Virtanen, toimitussihteeri Teemu Peltola ja heidän lisäksi yksi toimittaja. Varsin suuri osa lehden sisällöstä tuotetaan avustajavoimin. (Peltola, Teemu, henkilökohtainen tiedonanto 13.11.2013.)⁶

3.1. Datajournalismi Aamuset-lehdessä

Aamuset-lehti on tehnyt noin parikymmentä uutisjuttua, joiden tekemisessä on hyödynnetty datajournalistisia työskentelytapoja. Suurimmaksi osaksi uutisista on vastannut Teemu Peltola. Peltola kertoo kiinnostuneensa datajournalismista ensimmäistä kertaa osallistuttuaan Tampereen yliopiston järjestämälle kurssille, jolla opeteltiin muun muassa käyttämään datajournalismin yleisimpiä alustoja ja työkaluja. (Peltola 2013.)

Pääasiassa Aamuset-lehden datajournalistiset jutut ovat olleet lähtöisin aiheista, jotka ovat sisältäneet paljon lukuja ja tilastoja. Tilastoja on havainnollistettu graafeilla, jotka ovat usein verkkoversiossa jollain tapaa interaktiivisia, mutta printtilehdessä hieman suppeampia. Monet Aamuset-

⁶ Jatkoissa (Peltola 2013).

lehden datajournalismia hyödyntävistä uutisista ovat lähtöisin esimerkiksi Terveystiedon ja hyvinvoinnin laitoksen tilasto- ja indikaattoripalvelu SOTKANetin kaltaisista verkkosivustoista. Tilastoja on uutisjutuissa sovellettu niin, että ne keskittyvät Aamuset-lehden jakelualueen kuntiin. (Peltola 2013.)

Osa Aamuset-lehden datajournalismista on syntynyt joukkoistamisella (Peltola 2013). Tällöin apuna on käytetty erilaisia nettipohjaisia lukijakyselyitä, jollaisiin esimerkiksi Google tarjoaa ilmaisen alustan. Monet tällaisista uutisjutuista ovat syntyneet runsaasti julkista keskustelua herättäneistä aiheista, kuten Turun Kupittaaalla patsaspuiston purkamisesta, josta Aamuset-lehti uutisoi ensimmäisenä (Virtanen 2013).

Jatkojuttua varten Aamuset-lehti kehitti kyselyn, jossa lukijoilta kysyttiin, oliko patsaspuiston purkaminen Turun kaupungilta oikea vai väärä ratkaisu. Lisäksi lukijoilta kysyttiin, miten he kehittäisivät Seikkailupuistoa, josta patsaat raivattiin pois. (Peltola 2013a.) Kyselyn tuloksena syntyi uutisjuttu ”Seikkikseen halutaan lisää leikkipaikkoja”, joka julkaistiin Aamuset-lehdessä 18.10.2013. Juttu kertoi infografiikkaa hyödyntäen muun muassa, että 61 prosenttia kyselyyn vastanneista piti purkamista oikeana ratkaisuna ja että 80 prosenttia haluaisi Seikkailupuistoon lisää leikkipaikkoja. (Peltola 2013b.) Peltola toteaaakin, että tällainen journalistinen menetelmä on paljon toimivampi ja vaivattomampi kuin perinteiset tavat.

Perinteisen katugallupin sijaan datajournalisti pystyttää nettiin kyselyn, jossa on sekä rajoitettuja monivalintakysymyksiä ja avokysymyksiä. Tällä tavalla saadaan vaivatta käyttöön jopa satojen ihmisten mielipiteet sen sijaan, että käytetään kahdeksan tuntia vanhanaikaiseen katugallupiin, jonka tuloksena on ehkä 10-20 omalla naamallaan kameran eteen uskaltaneen ihmisen kommentti. (Peltola 2013.)

Kupittaaan patsaspuiston purkaminen ei kenties ollut kovinkaan merkittävä kunnallispoliittinen päätös, mutta se oli äärimmäisen paikallinen aihe ja turkulaiset kokivat sillä olevan merkitystä heidän arkeensa. Vaikka patsaspuiston tapauksesta syntynyt datajournalismi ei saanutkaan alkunsa esimerkiksi tuhansista tilastoriveistä, sitä voidaan silti verrata luvussa 2.2 mainitun Jerry Vermasen näkemykseen (Gray ym. 2012, 366), jonka mukaan

juuri paikalliset asiat ja niiden käsitteleminen voivat tehdä datajournalismista merkittävän työskentelytavan paikallistoimittajille.

3.2 Ensimmäinen julkaistu datajournalistinen juttu Aamuset-lehdessä

Aamuset-lehden ensimmäinen datajournalistinen juttu ”Koulukiusaamisessa isot erot Turun seudulla” julkaistiin 21.3.2013 (Peltola 2013c). Uutista varten Peltola oli hankkinut dataa Terveyden ja hyvinvoinnin laitokselta, joka oli toteuttanut valtakunnallisen 8. ja 9. luokan koululaisille suunnatun koulukiusaamiskyselyn. Kyseinen uutista varten käytetty kyselydata on vuodelta 2011.

Datan käsittelyä varten Peltola käytti Infogr.am-palvelua, joka auttaa Excel-muodossa olevan datan muuntamista helposti ymmärrettävään visuaaliseen muotoon. Datan analysoinnin aikana ilmeni, että esimerkiksi Maskussa lähes 10 prosenttia THL:n kyselyyn vastanneista koki tulleen kiusatuksi koulussa vähintään kerran viikossa, kun taas esimerkiksi Turussa vastaava prosenttiluku oli 5,7. (Peltola 2013.) Tämä havainnollistaa hyvin sitä, miten datan analysointi voi toisinaan johdattaa uutisidean jäljille. Hyvä toimittaja osaa löytää datasta kiinnostavia seikkoja ja esittää tilastoille kysymyksiä, ja kuten tässä tapauksessa, toimittajan tulee selvittää, voiko kiusaamistilastojen paikkakunta-kohtaiselle vaihtelulle löytyä jokin syy.

Toisinaan datan taustalta voi löytyä jokin konkreettinen uutinen. Juuri näin kävi Aamuset-lehden ensimmäisen datajournalistisen jutun kohdalla. Dataa analysoituaan Peltola jatkoi tiedonhankintaa ja sai selville, että Maskun yläkoulun eli Hemmingin koulun oppilaiden määrä kasvoi kuntaliitosten vuoksi puolella vuonna 2011. Tuolloin kouluun tuli uusia oppilaita Maskun Askaisten ja Lemun kouluista, mikä Hemmingin koulun opettajien mukaan näkyi kiusaamisena ja osaltaan selittää korkeat luvut THL:n koulukiusaamiskyselyssä. (Peltola 2013.)

3.3 Datajournalismin ongelmat ja esteet Aamuset-lehdessä

Aina tilastopoikkeamiinkaan ei löydy selitystä. Näin kävi Aamuset-lehden uutisjutussa, joka kertoo alle 25-vuotiaiden naisten aborttien olevan yleisempiä

kuin muualla Suomessa (Peltola 2013d). Peltola ei tiedonhankintaa tehdessään onnistunut lukuisista yrityksistä huolimatta selvittämään, mistä kyseinen piikki tilastoissa johtuu. Lopputuloksena syntynyt uutinen lähinnä havainnollistaa tilastot, muttei esitä syitä niille. Uutinen oli Peltolan mukaan eniten aikaa ja vaivaa vaatinut Aamuset-lehdessä julkaistu datajournalismia hyödyntävä juttu, vaikka hän toteaaakin sen olevan vain ”tilastojuttu”. (Peltola 2013.)

Peltola näkee erityisesti Turkuun kohdistuvan paikallisen datajournalismin esteenä sen, että Turussa ei ole syntynyt samanlaista avoimeen dataan keskittyntä aktiivisten henkilöiden verkostoa kuten esimerkiksi Helsingissä. Esimerkiksi Helsinki Region Infoshare -palvelu tarjoaa erilaisia avoimia pääkaupunkiseutuun liittyviä aineistoja, mutta Turussa ei ole mitään vastaavaa sivustoa. Peltola uskoo vakaasti, että mikäli samanlainen datapankki syntyisi joskus Turkuunkin, se vaikuttaisi myönteisesti paikallisen datajournalismin määrään. (Peltola 2013.)

Peltola kertoo myös, että Aamuset-lehdessä oman osaamisen puute voi toisinaan olla esteenä syventävälle datajournalismille.

En toistaiseksi osaa koodata omia ohjelmia, joten todelliset nuuskivan journalismin jutut ovat jääneet tekemättä. Esimerkiksi jossain Keski-Suomessa oli keissi, jossa parkkivuolan rakentamisen tarpeellisuutta perusteltiin olemassaolevien parkkivuolien korkealla käyttöasteella. Joku datajournalisti koodasi ohjelman, joka nuuski olemassaolevien parkkivuolien kotisivuja reaaliajassa kahden viikon ajan. Sivulla näet kerrottiin reaaliaikainen tyhjien paikkojen määrä. Tämä data purkitettiin jutuksi, jolloin ilmeni, että parkkivuolien käyttöaste ei ollut alkuunkaan niin korkea mitä jotkut olivat väittäneet. (Peltola 2013.)

Joskus Aamuset-lehden teettämät joukkoistamista hyödyntävät kyselyt eivät ole keränneet tarpeeksi materiaalia, jolloin uutisjutut ovat jääneet yksinkertaisesti tekemättä. (Peltola 2013.)

3.4 Haastattelu opinnäytetyön tutkimusmenetelmänä

Haastattelu valikoitui opinnäytetyön tutkimusmenetelmäksi erityisesti siksi, että tietooni ei ole tullut erityisen paljon sellaisia paikallistoimituksia, joissa datajournalismia käytettäisiin aktiivisesti osana normaalia toimitustyötä. Olisi toki mahdollista haastatella esimerkiksi Helsingin Sanomien datatoimituksen ja

Ylen uutisten Plus-deskin työntekijöitä, mutta tämä opinnäytetyö keskittyy paikallisjournalismiin. Kaikkiin Suomessa työskenteleviin datajournalisteihin keskittynyt kvantitatiivinen tutkimus olisi joka tapauksessa saattanut jäädä suppeaksi, sillä vuonna 2013 suomalaisia datatoimittajia arvioitiin olevan noin kymmenen (Koskenranta 2013, 7).

Haastattelu sopii opinnäytetyön tutkimusmenetelmäksi siksikin, että datajournalismi ja erityisesti datajournalismi paikallistoimituksessa on varsin kartoittamaton aihe, ja oli työn kannalta tärkeää saada laajoja ja syventäviä vastauksia. Nämä nähdään haastattelun etuina tutkimusmenetelmää päätettäessä (Hirsjärvi & Hurme 2000, 35).

Peltolaa on haastateltu opinnäytetyötä varten sähköpostitse, ja hän vastasi laajasti kaikkiaan 11 kysymykseen. Tämä vaikutti oikealta ratkaisulta siksi, että olen toiminut Aamuset-lehden avustajana vuodesta 2011 ja käynyt siitä lähtien lukuisia journalistiseen työhön liittyviä keskusteluja Peltolan kanssa niin kasvotusten, puhelimitse kuin sähköpostinkin välityksellä. Uskoakseni haastattelemisen esimerkiksi kasvotusten tuskin olisi muuttanut Peltolan vastausten laajuutta tai sisältöä merkittäväällä tavalla. Mikäli haastateltava olisi ollut ennestään täysin tuntematon, en olisi valinnut sähköpostia haastatteluvälineeksi, sillä tällöin vastausten tyylistä, pituudesta tai edes vastausajasta olisi ollut mitään takeita ennalta.

4 DATAJOURNALISTINEN JUTTU AAMUSET-LEHDELLE

Tämän opinnäytetyön tuoteosaa varten tarkoituksena oli tehdä Aamuset-lehdelle jokin uutisjuttu, joka hyödyntäisi tavalla tai toisella datajournalismia. Vaikka datajournalismin teoria on aiemmin ollut suurilta osin tuttua, en ollut ennen tätä opinnäytetyötä koskaan tehnyt julkaistua datajournalistista uutisjuttua. Kenties juuri tästä syystä aiheen ideoiminen ja tuoteosan aloittaminen osoittautuivat yllättävän hankaliksi, sillä uutisen muoto ja esitystapa olivat jo määrätty ennalta, mikä rajoittaa ideointia varsin paljon. Olen tottunut tekemään journalistisia juttuja verkkoon niin, että ensin kehittelen idean, ja vasta sen jälkeen tai tiedonhankinnan aikana pohdin mahdollista esitystapaa.

Tuoteosan idean kehittelyvaiheessa on käyty läpi runsaasti erityisesti nuoriin suomalaisiin liittyviä erilaisia data-aineistoja THL:n SOTKANet-palvelussa. Työn alkuvaiheessa lukuisia Excel-muodossa olevia tilastoja ladattiin ja muunneltiin erilaisiksi graafeiksi. Monista tilastoista löytyi kiinnostavia poikkeamia, mutta useista, kuten nuorten alkoholin kulutuksen vähentymisestä oli jo tehty uutisjuttuja muualla mediassa. Monet alkoholin käyttöön liittyvät tilastot uutisoidaankin usein jo samana päivänä kun ne ilmestyvät netissä yleiseen käyttöön. Lisäksi uutisen aiheen kehittämissä ongelmana oli se, että kaikkia SOTKANet-palvelun kiinnostavia tilastoja ei ole saatavilla paikkakunta- tai edes maakuntakohtaisesti. Tällöin juttua ei kannata ruveta tekemään, mikäli lopputuloksella ei ole paikallista merkitystä Aamuset-lehden lukijoille.

4.1 Datan löytäminen ja käsitteleminen

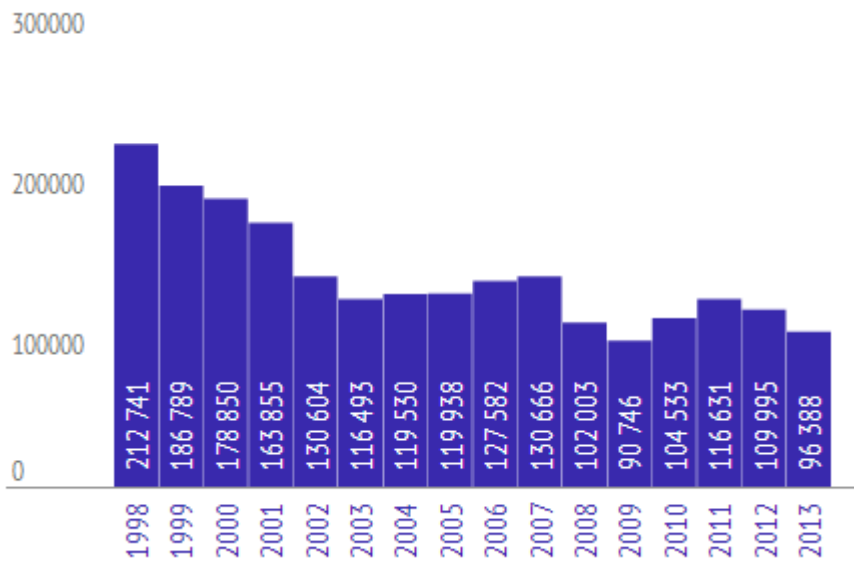
Vaikka Turun alueelta ei löydy Helsinki Region Infosharen kaltaisia datapalveluja, Lounais-Suomen paikkatietokeskus Lounaispaikka⁷ tarjoaa jotakin alueeseen liittyvää avointa dataa, kuten paikkatietoja Varsinais-Suomen veneenlaskupaikoista ja yritystonteista. Aineistosta löytyivät myös Finavian lentoliikennetilastot, josta ilmenivät Turun lentoaseman matkustajatilastot,

⁷ <http://www.lounaispaikka.fi/>

liikenneilmailun laskeutumiset ja ulkomaan sekä kotimaan tavaraliikenteen määrä. Datamäärä itsessään ei ole erityisen valtaisa, sillä tilastot ovat vuosilta 1998 – 2012. Jotta jutun näkökulmaa voisi rajata sopivasti, tuntui kannattavalta keskittyä vain lentoaseman matkustajamääriin. Vuoden 2013 tilastot löytyivät vielä erikseen Finavian verkkosivuilta.

Infogr.am-palvelu sopi datan analysointiin hyvin, sillä se havainnollistaa luvut ymmärrettävään visuaaliseen muotoon, jolloin tilastoja on kätevä tarkkailla ja mahdolliset poikkeamat nousevat esiin. Ennen tätä tilaston rivit ja sarakkeet tuli vaihtaa taulukkolaskentaohjelmassa oikeanlaiseen järjestykseen, jotta ne näkyvät Infogr.am-sivustolla halutulla tavalla. Koska matkustajamäärätilastoja on vain 15 vuoden ajalta, käsittely Excelissä ei vaatinut erityisen paljon työtä. Aamuset-lehti on usein käyttänyt datajournalistisissa jutuissaan Infogr.am-palvelun esitystapaa, mutta dataa kannattaa analysoida palvelussa, vaikka sen julkaisusta ei vielä olisikaan varma. Dataa voi visualisoida vaikka vain itseään varten, mikä helpottaa sen mahdollista jatkokehittelyä uutiseksi.

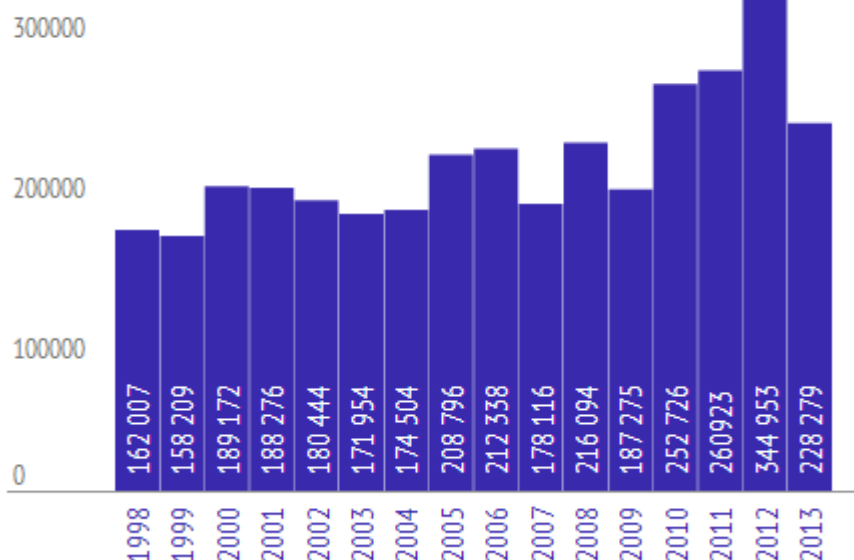
Tuoteosassa käsiteltyihin Turun lentoaseman matkustajatilastoihin lasketaan mukaan saapuvat ja lähtevät lennot sekä vaihtomatkustajat. Tilastoissa ei ole otettu huomioon sylilapsia eikä transitmatkustajia eli matkustajia, jotka eivät vaihda lentokonetta, vaan jatkavat matkaansa. Kun halutuksi muokattu data lopulta avautui Infogr.am-palvelussa, nousi esille havainto siitä, että kotimaan lentoliikenteen matkustajamäärät ovat olleet pienessä laskussa vuodesta 1998 (Kuvio 1.)



Kuvio 1. Turun lentoasemalla kotimaan lennoilla matkustaneiden määrät.

Visuaalisessa muodossa tilastot herättävät erilaisia kysymyksiä. Miksi juuri kotimaan lennoilla matkustavien määrä on laskenut Turun lentoasemalla juuri tuolla tavalla? Tapahtuiko jotain merkittävää vuonna 1998, jolloin lasku lähti käyntiin, vai jo kenties sitä ennen? Miltä kotimaan lentojen tilastot mahdollisesti näyttävät tulevaisudessa? Tuleeko kotimaan matkailu lentäen kenties kasvamaan, mikäli ulkomaille matkustaminen muuttuu tulevaisuudessa kalliimmaksi? Kuten kuviosta voi nähdä, kotimaan lennoilla matkustavien määrä Turun lentoasemalla on pysynyt pitkään suurin piirtein samana, mutta jyrkkä lasku vuodesta 1998 on erityisen silmiinpistävä.

Kun kotimaan lennoilla matkustavien data oli käsitelty, vuorossa oli vastaavien tilastojen muokkaaminen ulkomaisten lentojen aineistosta. Ulkomaan lennoilla matkustaneiden määrät näyttävät kotimaan tilastoihin verrattuna tyystin erilaisilta. Paitsi että määrät ovat olleet vuodesta 1998 maltillisessa kasvussa, tilastoissa vaikuttaa olevan myös paljon satunnaisia heilahteluja (Kuvio 2).



Kuvio 2. Turun lentoasemalla ulkomaan lennoilla matkustaneiden määrät.

Erityisen suuri tilastopiikki on havaittavissa vuoden 2012 tilastopalkissa, jonka luku on miltei 350 000. Tämäkin johdattelee erilaisen kysymysten äärelle, joskin hieman eri tavalla kuin kotimaan lentojen matkustajamäärien kohdalla. Mistä johtuu, että matkustajamäärät ovat olleet kasvussa, kuitenkin hieman vaihtelevasti? Mistä vuoden 2012 korkea luku johtuu? Entä voivatko sekä kotimaan että ulkomaan lentojen matkustajamäärät ja vaihtelut selittyä joillakin yleisemmällä lentomatkustamiseen liittyvillä trendeillä?

Tässä vaiheessa uutisen suunta alkoi hahmottua enemmän, ja lisäksi jutun laajuus tuntui sopivalta paikallisuutisia tarjoavaa Aamuset-lehteä varten. Ensimmäistä datajournalistista juttua varten ei olisi muutenkaan kannattanut ottaa liian suurta tilastomäärää käsiteltäväksi. Jutun tekemisen aikana oli koko ajan tärkeää huomioida, että työskentelyn rajoina tulisivat olemaan sekä oma varsin vähäinen datajournalistinen osaaminen että Aamuset-lehden tyyli ja painotus paikallisuudessa. Tämän jälkeen tarkemmat kysymykset alkoivat hahmottua ja tiedonhankinta alkoi edetä.

4.2 Tiedonhankinta datan saamisen jälkeen

Toisinaan erilaisissa datajournalistisissa jutuissa on esitelty tilastoja, jotka ovat kaikkien vapaasti ladattavissa, mutta joita mikään muu media ei vain ole keksinyt aiemmin käyttää jutuissaan. Vaikka Turun lentoaseman matkustajamäärätilastot ovatkin juuri sellaisia, pelkät tiedot tilastoista itsessään eivät ole uutinen. Tämän vuoksi oli tärkeää miettiä jutulle toisenlaista näkökulmaa. Kysymyksiä herättivät muun muassa Turun lentoaseman rooli koko Suomen lentomatkustusliikenteessä ja halpalentoyhtiöiden yleinen merkitys Turun lentoasemasta lähtevälle ja sille saapuvalla lentoliikenteelle.

Uutisen tiedonhankinnan aikana juttua varten haastateltavaksi valikoitui Finavian verkostojohtaja. Kotimaan lennoilla matkustavien määrän väheneminen johtuu hänen mukaansa pääosin siitä, että hiljalleen mahdollisuudet junamatkustamiseen ja moottoriteillä ajamiseen ovat tulleet paremmiksi. Tarkempaa syytä vuonna 1998 alkaneeseen laskuun on varsin vaikeaa määritellä. Verkostojohtajan mukaan vuoden 2012 tilastopiikki puolestaan on ilmiselvää seuraus halpalentoyhtiö Ryanairista, joka aloitti kyseisenä vuonna suorat lennot Turun lentoasemalta neljään eri kohteeseen. Tällainen selitys tilastopoikkeamalle oli pitkälti arvattavissa, mutta on luonnollisesti tärkeää saada asiaan aina vahvistus viralliselta tietolähteeltä. Ryanair siis lisäsi merkittävästi matkustuskysyntää Turun lentoasemalla, mutta se ei ole uutta tietoa, sillä asia esimerkiksi mainittiin jo useissa uutisissa, jotka kertoivat Ryanairin lopettavan lennot Turusta jo samana vuonna 2012 (Peltoniemi 2012). Lopettamisen vuoksi matkustajamäärät laskivat vuonna 2013 lähes edellisten vuosien tasolle (Kuvio 2).

Tiedonhankintaa tehdessä kävi siis ilmi se, että Turun lentoaseman tilastot itsessään eivät ole uutinen eivätkä toisaalta kaikki tilastoja (ja erityisesti tilastopiikkiä) selittävät syytkään. Kuitenkin Ryanairin toimintatapa ja matkustajamääriin liittyvät tilastot herättivät yleisiä kysymyksiä lentomatkustamisesta ja sen tulevaisuudesta sekä halpalentoyhtiöiden merkityksestä Turun lentoasemalle. Siksi uutisen näkökulmaksi muodostui lopulta se, kuinka valtavan suuri merkitys halpalentoyhtiöiden suorilla reiteillä on

Turun lentoaseman toiminnalle. Lopputuloksena syntynyt uutinen on enemmänkin katsaus Turun lentoaseman nykytilanteeseen ja siihen, millä tavalla se on riippuvainen nykyajan lentomat kustustrendeistä.

Lopullinen julkaistu tuoteosa on luettavissa Aamuset-lehden sivuilta: <http://www.aamuset.fi/naista-puhutaan/uutiset/halpalentoyhtioiden-tarjonta-elintarkeaa-turun-lentokentalle>

5 JÄLKIPOHDINTA: DATA AUTTAA IDEOINNISSA

Tämän opinnäytetyön tuoteosa oli ensimmäinen käytännön kosketukseni datajournalismiin. Osittain siksi juttua varten käsiteltävä data ei koostunut tuhansista riveistä, mutta toisaalta onnistunut lopputulos on osoitus siitä, ettei datajournalismin aina tarvitsekaan olla erityisen suuripiirteistä. Koska paikallisjournalismi ja tietyt puitteet asettavat datajournalismille omat rajansa, poikkeuksellisen laaja juttu ei lehden printtiin tai verkkoon välttämättä mahtuisikaan, jos näkökulma ei pysy tarpeeksi paikallisena.

Tuoteosan aloittaminen oli yllättävän vaikeaa, sillä tilanne oli hieman keinotekoinen: tehtävänä oli ideoida uutisjuttu tiettyyn muotoon eli tehdä siitä jollain tapaa datajournalistinen. Tällainen lähestymistapa toimii omien kokemusteni mukaan uutistyössä huonosti, sillä journalismin tulisi alkaa hyvästä uutisideasta, minkä jälkeen vasta pohditaan parasta mahdollista tapaa esittää uutinen lukijalle. Kun esitystapa on tietyllä tapaa rajattu jo ennalta ja tiedonhankinnan täytyy noudattaa tiettyä kaavaa (datan hankkiminen), tulee ideoinnista hyvin hankalaa. Osaltaan tähän vaikuttaa myös se, että tuoteosan tekemisessä kyse oli minulle aiemmin kokonaan tuntemattomasta työskentelytavasta.

On kuitenkin tärkeää huomioida, että datajournalismin ei tarvitse olla monimutkaista. Lopputuloksena syntyneessä uutisjutussa on kaksi yksinkertaista graafia, mutta varsinainen sisältö on itse uutisen tekstissä. Koko uutisen tehtävänä ei ole tukea graafeja, vaan nimenomaan graafit tukevat uutisen sisältöä.

Kaikkein olennaisinta syntyneessä tuoteosassa ja sen aiheuttamassa pohdinnassa on kuitenkin se, että datajournalistiset työskentelytavat auttoivat löytämään Lounaispaikka-tietopalvelusta uutisidean, joka tuskin olisi tullut löydetyksi muualta toisin keinoin. Finavian data ei ole erityisen suuressa roolissa, mutta sillä ei ole väliä, koska pelkkä data itsessään ei edusta uutista. Onkin tärkeämpää miettiä, mitä hyötyä datajournalismista on ollut Aamuset-

lehdelle tehdyn tuoteosan tekemisessä kuin yrittää määritellä tarkasti, millaista datajournalismia lopputulos mahdollisesti edustaa.

5.1 Datajournalismin tekeminen paikallistoimituksessa

Luvussa 2 kerrottiin siitä, miten The Guardianin toimituksessa suurin osa datajournalistisesta työstä tapahtuu pitkälti Excel-ohjelman parissa (Rogers 2011, 1084). The Guardiania pidetään yleisesti datajournalismissa edistyneenä suunnannäyttäjänä, mutta ainakin yhdessä asiassa Aamuset-lehti toimii täsmälleen samalla tavalla: kaikki Aamuset-lehdessä datajournalismin yhteydessä käytettyä dataa on ainakin jossain vaiheessa käsitelty Excelissä. (Peltola 2013). Kenen tahansa datajournalismiin pyrkivän toimittajan kannattaakin ehdottomasti opetella Excelin tai muun vastaavan taulukkolaskentaohjelman käyttö. Esimerkiksi monia datajournalismin alkeita opettavilla kursseilla on käynyt ilmi, ettei toimittajien ole aiemmin tarvinnut työssään tutustua datan käsittelyyn juuri ollenkaan (Koskenranta 2013, 13).

Kuten Aamuset-lehden Peltolakin osittain toteaa, paikallista dataa tarjoavat palvelut sekä niin kutsutun data-ajattelun edistyminen Turun alueella saivat hyvin todennäköisesti monet paikalliset toimittajat kiinnostumaan datajournalismista helpommin. Avointa dataa saa kyllä käyttöön esimerkiksi edellä mainituista palveluista sekä julkisuuslakiin perustuvan, viranomaisille esitettävän tietopyynnön avulla, mutta kovin harva datajournalismia aloitteleva edes osaa etsiä dataa kyseisellä tavalla.⁸

Datajournalismin tekeminen on mahdollista pienessä paikallismediassa, mutta tietyin rajoittein. Sen tulee olla tarpeeksi paikallista, jotta sillä on merkitystä lukijoille. Liian suurilta tuntuvia datamääriä ei kannata lähteä purkamaan, mutta niistä voi aina yrittää etsiä esimerkiksi maakunta- tai kuntakohtaisia tilastoja. Lisäksi toimittajan tulee varautua siihen, että dataa hankittuaan ja analysoituaan hän saattaa jossain vaiheessa havaita siihen, että on tullut tehneeksi turhaa työtä. Toisaalta tällaista saattaa tapahtua jatkuvasti muussakin uutistyydyssä ja journalismin tekemisessä.

⁸ Tietopyynnössä ja julkisuuslain selkiyttämisessä auttaa verkkosivu <http://tietopyynto.fi/>

Yksi hankaloittava tekijä paikallistoimituksessa tehtävässä datajournalismissa on se, ettei toimittaja voi etukäteen täysin tietää, kuinka paljon aikaa jutun tekeminen tulee viemään. Näin voi käydä erityisesti silloin, kun toimittaja vasta etsii data-aineistosta mahdollisia uutisideoita vailla minkäänlaista käsitystä esimerkiksi siitä, mitä aihealuetta hänen uutisensa edes tulisi käsittelemään. On mahdollista, että mitä enemmän datajournalistisia juttuja toimittaja on tehnyt, sitä paremmin hän osaa arvioida, kuinka paljon työaikaa datajournalistisen uutisidean etsiminen vaatii. Luvussa 2.2 mainittiin ”dataähky” (Vehkoo 2011), ja Aamuset-lehden Peltolan mukaan hankalin asia datajournalismissa onkin juuri kaiken datan määrä.

Hankalin asia on aineistojen valtava määrä. Jo nykyäänkin Suomessa saatavilla olevaa avointa dataa voisi louhia päiväkausia pelkästään aiheita etsien, datan Excelissä pyörittelystä puhumattakaan. Lisäksi uutta matskua julkaistaan koko ajan, joten resurssini eivät yksinkertaisesti riitä kaiken mielenkiintoisen tutkimiseen. Eivätkä riittäisi sittenkään, vaikka olisin täysipäiväinen datajournalisti. (Peltola 2013.)

LÄHTEET

Aro, J. Vain joka neljäs luottaa toimittajaan. Journalisti. Viitattu 4.11.2013
<http://www.journalistiliitto.fi/journalisti/lehti/2012/14/uutiset/vain-joka-neljas-luottaa-toimitt/>

Gray, J; Chambers, L. & Bounegru, L. 2012. The Data Journalism Handbook. Sebastopol: O'Reilly Media.

Hirsjärvi, S. & Hurme, H. 2000. Tutkimushaastattelu: teemahaastattelun teoria ja käytäntö. Helsinki: Yliopistopaino.

Koskenranta, R. 2013. Toimittaja, data tarinaksi – Datajournalismi uutistoimituksen työkalupakkiin. Opinnäytetyö. Mediatuottamisen koulutusohjelma. Helsinki: Metropolia Ammattikorkeakoulu.
https://www.theseus.fi/bitstream/handle/10024/59105/130516_Koskenranta_R_opinnaytetyo_final.pdf

Peltola, T. Henkilökohtainen tiedonanto. 13.11.2013.

Peltola, T. 2013a. Nettikysely: Mihin suuntaan kehittäisit Seikkailupuistoa? Aamuset-kaupunkilehti. Viitattu 4.5.2014 <http://www.aamuset.fi/naista-puhutaan/kaupunki/nettikysely-mihin-suuntaan-kehittaisit-seikkailupuistoa>

Peltola, T. 2013b. Seikkikseen halutaan lisää leikkipaikkoja. Aamuset-kaupunkilehti. Viitattu 4.5.2014 <http://www.aamuset.fi/naista-puhutaan/kaupunki/nettikysely-mihin-suuntaan-kehittaisit-seikkailupuistoa>

Peltola, T. 2013c. Koulukiusaamisessa isot erot Turun seudulla. Aamuset-kaupunkilehti. Viitattu 4.5.2014 <http://www.aamuset.fi/naista-puhutaan/uutiset/koulukiusaamisessa-isot-erot-turun-seudulla>

Peltola, T. 2013d. Alle 25-vuotiaiden naisten abortit Turun seudulla merkittävästi muuta maata yleisempiä. Aamuset-kaupunkilehti. Viitattu 4.5.2014 <http://www.aamuset.fi/naista-puhutaan/uutiset/alle-25-vuotiaiden-naisten-abortit-turun-seudulla-merkittavasti-muuta-maata>

Peltoniemi, P. 2012. Ryanair lopettaa Turun lennot. Yle Uutiset. Viitattu 4.5.2014 http://yle.fi/uutiset/ryanair_lopettaa_turun_lennot/6345289

Reimold, D. 2013. Journalism of Ideas: Brainstorming, Developing, and Selling Stories in the Digital Age. New York: Routledge.

Rogers, S. 2011. Facts are Sacred. Lontoo: Guardian Books.

Salminen, J. 2012. Avoin tieto tuo miljardeja ja uudistaa demokratiaa – näin se tehdään. Suomen Kuvalehti. Viitattu 4.11.2013 <http://suomenkuvalehti.fi/jutut/kotimaa/avoin-tieto-tuo-miljardeja-ja-uudistaa-demokratiaa-nain-se-tehdaan>

Sjöholm, P. 2013. Maksu, jota et kilpailuta – kartta näyttää isot erot sähkön siirtohinnoissa. Yle. Viitattu 4.11.2013 http://yle.fi/uutiset/maksu_jota_et_kilpailuta_kartta_nayttaa_isot_erot_sahkon_siirtohinnoissa/6869685

Vehkoo, J. 2011. Painokoneet seis! Kertomuksia uuden journalismin ajasta. Helsinki: Teos.

Virtanen, L. 2013. Kupittaa patsaspuisto aiotaan purkaa. Aamuset-kaupunkilehti. Viitattu 4.5.2014 <http://www.aamuset.fi/naista-puhutaan/uutiset/kupittaa-patsaspuisto-aiotaan-purkaa>

Opinnäytetyön kirjallisuuskatsaus

Tarkastelussa datajournalismi – mitä se on ja miksi se on tärkeää journalismille?

Mitä datajournalismi on?

Journalisti Paul Bradshaw'n mukaan data voi toimia joko journalistisen jutun lähteenä tai välineenä, jonka avulla juttu kerrotaan, tai se voi olla molempia (Gray ym. 2012, 274). Tietokannat ovat olennainen osa datajournalismia, mutta olisi liian yksinkertaistavaa kutsua datajournalismia vain journalismiksi, joka käsittelee dataa. Siksi yksi tapa hahmottaa datajournalismin tarkoitus paremmin on tutustua siihen, kuinka datajournalismia käytännössä tehdään.

Datajournalismi tulee tarpeelliseksi silloin, kun toimittajan tulee käsitellä ja analysoida niin isoja tietokantoja, ettei niitä pysty suodattamaan käsin. The Guardian entinen datajournalisti ja nykyään Twitterin dataosastolla työskentelevä Simon Rogers käsitteli WikiLeaksille vuodetuista Afganistan-asiakirjoista kertovaa uutista varten Excel-tiedostoa, joka sisälsi 92 201 rivin verran dataa (Rogers 2011, 252).

Opinnäytetyöni keskeisenä näkökulmana on se, millaiset mahdollisuudet pienen paikallismedian toimittajalla on datajournalismin tekemiseen. Tätä ajatellen kannattaa ottaa huomioon, että varsin suuri osa erilaisista datajournalismia varten käytettävistä työkaluista on saatavilla ilmaiseksi internetistä.

Yksi tärkeimpiä työkaluja on taulukkolaskentaohjelma, kuten Microsoft Excel, mutta vastaava ohjelma on saatavilla ilmaiseksi esimerkiksi Google Docs-palvelun kautta. Simon Rogersin mukaan suuri osa The Guardianissa tehtävästä datajournalismista tapahtuu Excelin avulla (Rogers 2011, 1084). Näin ollen datajournalismin tekemiseen ryhtyvä paikallistoimittaja tuskin törmää ainakaan työkalujen puutteeseen.

Rogers arvelee, että datajournalismi on noussut yleiseen tietoisuuteen eräänlaisen läpinäkyvyysajattelun maailmanlaajuisen leviämisen myötä. Datan

saatavuus ja sen helppokäyttöisyys sekä yleinen mielenkiinto datan visualisoimista kohtaan ovat lisääntyneet, kuten myös laajat uutisjutut, joiden tekemiseen on käytetty suuria määriä dataa. (Rogers 2011, 112.)

Englantilainen The Guardian on ollut datajournalismia tehdessään edullisessa asemassa, sillä datan avaaminen julkiseen käyttöön on ilmiö, joka on edennyt erityisesti Britanniassa sekä Yhdysvalloissa. Myös Suomessa hallitus teki vuonna 2011 periaatepäätöksen julkishallinnon datan avaamisesta. (Vehkoo 2011, 96.)

Myös tämä on opinnäytetyön kysymyksen kannalta oleellinen seikka: mitä enemmän ja helpommin dataa on saatavilla, sitä helpommaksi datajournalistinen työ muodostuu, ja silloin myös juttuaiheita on saatavilla enemmän.

Datajournalismi on siis yleistynyt käytännössä katsoen kahdesta syystä: datan saatavuus on parantunut ja datan käsittelyä varten tarvittavat ohjelmat ovat nykyään saatavilla kenelle tahansa, jolla vain on käytössään tietokone. Datajournalismi on journalismia, joka hyödyntää näitä molempia. Ilmiönä se on vasta verrattain uusi, sillä esimerkiksi Yhdysvaltain avoimen datan sivusto data.gov aukesi vasta vuonna 2009 (Rogers 2011, 112).

Eastern Illinois Universityn dosentti Bryan Murley näkee datajournalistisessa työssä kolme vaihetta: datan keräämisen, analysoimisen sekä esittämisen. Esittämistapoja on monia, mutta hänen mukaansa datasta kirjoittaminen ei riitä. (Reimold 2013, 3497.) Simon Rogersin mielestä taas datajournalismin esittäminen voi hyvin toimia tekstinäkin. (Rogers 2011, 143).

Sinänsä ajatuksessa ei ole mitään uutta: journalistit ovat aina etsineet ja käsitelleet tietoa sekä muokanneet sitä helpommin ymmärrettävään muotoon tekstiksi, kuviksi tai infografiikoiksi. Kuitenkin datajournalismin ollessa kyseessä niin tiedon määrä kuin käsittelytavat ovat poikkeuksellisia.

Iso osa datajournalistisesta prosessista on datan käsittelyä ja puhdistamista. Esimerkiksi datan taulukkoa voi joutua järjestelemään postinumeroiden

mukaan, jotta kokonaisuus hahmottuu paremmin. Bryan Murley neuvoo myös käyttämään paljon aikaa siihen, ettei datasta tehtäviin grafiikoihin ja statistiikoihin pääse lipsahtamaan virheitä. Monet tahot haluavat datan näyttävätyvän julkisuudessa todellisuutta parempana, joten siksi datan taustat ja lähteet tulee tarkistaa. (Reimold 2013, 3507, 3547.) Mikäli jokin data ei täsmää, se voi jopa johdattaa kokonaan uuden uutisjutun alkulähteelle ja herättää kysymyksiä siitä, miksi dataa on alun perin vääristelty.

Aivan kuten tiedotteitakin muokataan lähteelle edulliseen sävyyn, voidaan dataakin muokata samalla tavalla. Siksi myös datajournalismissa tulee muistaa osoittaa tarpeellista kriittisyyttä lähteitä kohtaan. Datajournalistin tärkein ja mahdollisesti vaikein tehtävä on esittää datalle oikeita kysymyksiä (Vehkoo 2011, 107).

Miksi datajournalismi on tärkeää journalismille? Miten journalismi hyötyy siitä?

Arvioiden mukaan maailmaan tuotetaan joka toinen päivä noin viiden eksabitin eli runsaan viiden miljardin gigabitin verran informaatiota. Ihmiset tuottavat osan informaatiosta esimerkiksi sosiaaliseen mediaan tekstipäivityksinä ja kuvina. (Vehkoo 2011, 95.) Datajournalismi auttaa jäsentelemään ja analysoimaan alati kasvavaa datan määrää. Yksi journalismin perustehtävistä on aina ollut pyrkiä selittämään maailmaa. Nyt datajournalismia tarvitaan selittämään, mitä miljardit gigabitit voivat kertoa maailmasta.

Saksalaiselle Bertelsmann-säätiölle työskennellyt Tom Fries puhuu ”informaation epäsymmetriasta”: uutta tietoa syntyy niin paljon ja niin nopeasti, että se vaikuttaa ongelmallisella tavalla jo ihmisten päivittäisiin päätöksiin ja toimintoihin. Hyvin tehty datajournalismi helpottaa ihmisiä informaatiotulvan keskellä. (Gray ym 2012, 366.) Googlen pääekonomisti Hal Varianin mukaan ongelmana ei olekaan enää datan puute, vaan ”dataähky” (Vehkoo 2011, 95).

Bryan Murley toteaa, että journalistien ei kannata turhaan pelätä numeroita, sillä numeroissa piilee ideoita uutisjutuille, joita sanat eivät aina kykene selittämään (Reimold 2013, 3488). Numeroilla Murley viittaa bitteihin ja binääreihin, joista loppujen lopuksi kaikki julkinen data muodostuu.

World wide webin keksijä Tim Berners-Lee uskoo vahvasti datajournalismin olevan olennainen osa journalismin tulevaisuutta ja väittää, että toimittajien täytyy opetella ymmärtämään tietokantoja ja koodia (Vehkoo 2011, 107). Datajournalismiin erikoistuva toimittaja voi olla nyt ja tulevaisuudessa edullisessa asemassa hakiessaan töitä. Toisaalta tietualan osaajille riittää kysyntää muutenkin: Deutsche Wellen toimittaja Mirko Lorenz huomauttaa, että datajournalistisista taidoista voi olla hyötyä muuallakin kuin toimittajan töissä, esimerkiksi viestintä- ja tiedotusallalla, sillä moniin yrityksiin ja järjestöihin kaivataan paljon työntekijöitä, jotka osaavat muokata saapuvasta datasta ymmärrettävää (Gray ym. 2012, 302).

Simon Rogers kuitenkin huomauttaa, että toimittajan ei tarvitse olla koodari tai ohjelmoija pärjätäkseen datajournalismissa; hän painottaa moneen kertaan, kuinka tärkeää on pikemminkin tarkastella dataa journalistin eikä analyytikon näkökulmasta (Rogers 2011, 74). Media Standards Trust -järjestön johtajan Martin Mooren mielestä toimittajien täytyy tulevaisuudessa ymmärtää koodin ja algoritmien merkityksiä, mutta silti hänen mielestään ammattitaitoisen journalistin ja koodia osaavan ohjelmoijan yhdistelmä on harvinainen. Mooren mukaan toimitukset tulevat tavitsemaan ryhmiä, jotka koostuvat molempien osa-alueiden osaajista. (Vehkoo 2011, 110.) Suomessa tällaisia ryhmiä on ollut ainakin Helsingin Sanomien ja Ylen toimituksissa.

Tämä on kiinnostava näkemys opinnäytetyöni tutkimuskysymyksen näkökulmasta: miten paikallismedian toimittaja kykenee tekemään datajournalismia yksin? Millainen datajournalismi on mahdollista toteuttaa yksin, millainen työ taas vaatii Mooren mainitsemien ryhmien apua? Missä kohtaa tulee vastaan raja, jolloin datajournalisti ei enää pärjää yksin?

Hollantilaisen NU.nl- uutissivuston datajournalisti Jerry Vermanen väittää, että datajournalismi tulee entisestään vahvistamaan journalismin ”vallan vahtikoiran” roolia erityisesti taloudellisesti vaikeina aikoina. Samaa mieltä on datajournalismiin erikoistuneen Journalism++-verkoston toimittaja Nicolas Kayser-Bril. Hänen mukaansa juuri journalistien kyvyttömyys ymmärtää numeroita johti siihen, miksi esimerkiksi yhdysvaltalaisen Enron-energiayhtiön

kirjanpitoskandaalia sekä Bernie Madoffin tapausta käsittelevistä uutisista puuttui aikoinaan tietty journalistinen oivalluskyky. (Gray ym. 2012, 355, 377.)

The Seattle Times -lehden toimittaja Cheryl Phillips on samaa mieltä. Hänen mukaansa datajournalismi paljastaa sellaisia asioita, jotka jäisivät tavalliselta journalismilta piiloon, kuten ympäristöä haittaavat poliittiset päätökset tai huumepolitiikan aiheuttamat kuolemat (Gray ym. 2012, 443). Data voi kätkeä taakseen paljon erilaisia uutisten aiheita.

Datajournalismilla on myös hyvin ajankohtainen tehtävä. Uutisorganisaatiot kilpailevat uskottavuudesta kansalaisjournalistien ja bloggaajien kanssa. Tämä on johtanut siihen, että uutistoimistot pyrkivät olemaan yhä avoimempia uutisten lähteiden suhteen. (Rogers 2011, 173.) Koska uutisten kuluttajista on tullut aiempaa mediakriittisempiä ja luottamus toimittajia kohtaan on vähentynyt, datajournalismi voi lisätä julkaisun uskottavuutta ja luotettavuutta sitä kohtaan. Datajournalistiset jutut kertovat lukijoille erittäin avoimesti, mistä tieto on peräisin. Kuten aiemmin on mainittu, journalistien tärkeimpänä tehtävänä vain on analysoida dataa kriittisesti itse jutuntekovaikheessa ja päättää, millaisessa muodossa se esitetään lukijoille.

Datajournalisti Jerry Vermasen mukaan datajournalismilla on ratkaiseva rooli paikallisjournalismissa. Vermasen mielestä journalistien tulee osata etsiä, analysoida ja visualisoida paikallistason dataa, sillä paikallisjulkaisuilla ja niiden uutisilla on suora vaikutus alueensa asukkaisiin. (Gray ym. 2012, 366.)

Tämä tuo jälleen uuden tulokulman, jota voi sivuta opinnäytetyössä. Miten datajournalismin mahdollisesti korostuva rooli on otettu huomioon paikallistoimituksissa? Mitä mieltä paikallismedioiden toimituksissa ollaan Vermasen näkemyksestä? Näitä voi sivuta lyhyesti, mutta kysymykset eivät saa viedä liikaa tilaa alkuperäisiltä tutkimuskysymyksiltä.

Lähteet

Gray, J; Chambers, L. & Bounegru, L. 2012. The Data Journalism Handbook. Sebastopol: O'Reilly Media.

Reimold, D. 2013. Journalism of Ideas: Brainstorming, Developing, and Selling Stories in the Digital Age. New York: Routledge.

Rogers, S. 2011. Facts are Sacred. Lontoo: Guardian Books.

Vehkoo, J. 2011. Painokoneet seis! Kertomuksia uuden journalismin ajasta. Helsinki: Teos.

Huomioitavaa

Kaikki kirjallisuuskatsauksessa mainitut teokset lukuunottamatta Vehkoon Painokoneet seis! -kirjaa ovat e-kirjoja. Koska kyseisillä e-kirjoilla ei ole merkittyjä sivuja, kirjallisuuskatsaukseen merkittyihin lähdeviitteisiin on lisätty sivunumeron tilalle viitatus tekstin sijainti (location). Viitatus kohdan löytää kustakin lähdeteoksesta syöttämällä sijaintiluvun e-kirjan lukulaitteeseen.

Vehkoon Painokoneet seis! -teokseen viitatessa on käytetty normaalisti sivunumeroita.