

TAMPEREEN YLIOPISTO

Anna-Sofia Nieminen

UUDEN AJAN JOURNALISMIA  
Datajournalismin määritelmä, merkitys ja tila Suomessa

Tiedotusopin pro gradu -tutkielma

Toukokuu 2015

TAMPEREEN YLIOPISTO

Viestinnän, median ja teatterin yksikkö

NIEMINEN, ANNA-SOFIA: Uuden ajan journalismia. Datajournalismin määritelmä, merkitys ja tila Suomessa

Pro gradu -tutkielma, 79 s. + 2 liites.

Tiedotusoppi

Toukokuu 2015

---

Selvitän pro gradu -työssäni, miten suomalaiset datajournalismin tekijät määrittelevät datajournalismin ja mitä merkitystä he näkevät sillä olevan journalismille, yleisölle ja yhteiskunnalle. Lisäksi kartoitan, minkälaista datajournalismia Suomessa on tehty syksyyn 2014 mennessä.

Tutkimukseni on laadullinen. Selvitin kahdeksan koodaajan tai toimittajan näkemyksiä datajournalismista teemahaastatteluilla. Haastateltavat ovat osallistuneet datajournalististen juttujen tekemiseen Yle Uutisissa, Helsingin Sanomissa tai STT-Lehtikuvassa. Lisäksi kävin vapaamuotoisia keskusteluja datajournalismin tilasta 13 eri toimituksen edustajan kanssa. Analyysini jakautuu kuvailevaan osioon ja teemahaastatteluille tehtyyn sisällönanalyysiin.

Tutkimuksessa ilmeni, että tekijöiden mielestä datajournalismia määrittää ennen kaikkea datan käyttö prosessissa eikä esitystapa. Kuitenkin datajournalismiin liittyy usein visuaalisia tai interaktiivisia elementtejä. Datajournalismia tehdään ensisijaisesti verkkoon. Ehdotan täältä pohjalta datajournalismin määritelmäksi seuraavaa: *Datajournalismi on journalistinen prosessi, jossa tuotetaan datan avulla uutta tietoa verkkoon tai verkko edellä ja johon liittyy usein visuaalisia ja interaktiivisia elementtejä.*

Datajournalismin merkitys näyttää kumpuavan sen määritelmästä. Merkitys kytkeytyy ennen kaikkea datasta löytyvään uuteen tietoon, monesti siis tutkivaan journalismiin. Tärkeää on myös verkonomaisen kerronnan kehittäminen ja sitä kautta uuden yleisön ja rahan houkuttelu.

Tutkimus vahvisti käsitystä siitä, että datajournalismi on Suomessa lapsenkengissä. Sen tekeminen on pitkälti Yleisradion ja Helsingin Sanomien harteilla. STT-Lehtikuva panostaa verkkovisualisointeihin. Maakuntalehdissä, MTV:ssä, iltapäivälehdissä, talousmedioissa ja ruotsinkielisessä mediassa suhtautuminen datajournalismiin on vaihtelevaa, eikä sen tekeminen ole säännöllistä. Haastatellut tekijät kuitenkin uskovat datajournalismin kehitykseen ja aseman vahvistumiseen.

Asiasanat: Datajournalismi, verkkojournalismi, tutkiva journalismi, visualisointi, data, Yleisradio, Helsingin Sanomat, STT-Lehtikuva, teemahaastattelu, laadullinen sisällönanalyysi

# Sisällys

1	Johdanto.....	1
1.1	Kohti tutkimusta.....	1
1.2	Tutkimuskysymykset .....	4
1.3	Tutkimuksen rakenne .....	5
2	Tutkimuksellinen viitekehys .....	7
2.1	Aiempi tutkimus .....	7
2.2	Keskeiset käsitteet ja kontekstit .....	9
2.2.1	Journalismi internet-aikana .....	9
2.2.2	Kontekstina verkkojournalismi ja osallistava journalismi.....	10
2.2.3	Kontekstina tietokoneavusteinen journalismi ja tutkiva journalismi .....	12
2.2.4	Data, avoin data.....	14
2.2.5	Datajournalismi .....	17
3	Aineisto ja menetelmät.....	25
3.1	Laadullinen tutkimus.....	25
3.2	Teemahaastattelu.....	26
3.3	Ensisijainen aineisto.....	28
3.4	Tausta-aineisto.....	32
3.5	Analyysin toteuttaminen .....	33
4	Suomalaisen datajournalismin tila vuonna 2014 .....	36
4.1	Datajournalismi Suomessa.....	36
4.2	Helsingin Sanomat .....	37
4.3	Yleisradio .....	39
4.4	STT-Lehtikuva .....	41
4.5	Maakuntalehdet.....	42
4.6	MTV .....	44
4.7	Talousmediat.....	45

4.8 Iltapäivälehdet .....	47
4.9 Ruotsinkieliset lehdet.....	48
4.10 Yhteenveto .....	49
5 Datajournalismin ilmiön selittämistä .....	51
5.1 Haastateltavat ilmiön arvioijina .....	51
5.2 Kuka on datajournalisti? .....	52
5.3 Kohti datajournalismin määritelmää.....	53
5.3.1 Verkonomaisuus, interaktiivisuus ja visuaalisuus .....	57
5.3.2 Tutkiva journalismi .....	59
5.4 Kohti datajournalismin merkitystä .....	60
5.4.1 Merkitys journalismille .....	62
5.4.2 Merkitys yleisölle ja yhteiskunnalle.....	64
5.5 Yhteenveto .....	65
6 Pohdinta .....	67
6.1 Kohti johtopäätöksiä .....	67
6.2 Tulokset ja aiempi tutkimus .....	67
6.2.1 Oma määritelmä datajournalismille .....	70
6.3 Tutkimuksen kriittinen tarkastelu ja merkitys.....	71
6.4 Jatkoa varten.....	73
Lähdeluettelo .....	75
Liite: Teemahaastattelun runko.....	80

# 1 Johdanto

## 1.1 Kohti tutkimusta

Kun alkuvuodesta 2015 pidettiin historian toinen pohjoismainen datajournalismikonferenssi, Suomella ei ollut hurraamista menestyksessä. Palkinnot menivät Norjaan, Tanskaan ja Ruotsiin. Suomalainen datajournalismin tekijä Teemo Tebest summasi blogissaan konferenssin herättämiä ajatuksia: muihin Pohjoismaihin verrattuna suomalaisten edustus oli hävyttömän kapea, ja edelliseen vuoteen verrattuna Suomi tuntuu jääneen tekemisen tasossa muita Pohjoismaita jälkeen (Tebest 2015).

Datajournalismi on tuore ilmiö, joka ei ole vielä vakiinnuttanut asemaansa Suomessa. Sitä ei myöskään ole tutkittu kovin paljoa, ei Suomessa eikä muualla maailmalla. Datajournalismilla ei ole vakiintunutta määritelmää, eikä sen tekijöillä välttämättä ole yhtenäistä käsitystä datajournalismista. Siinä on yksi tärkeä syy, miksi tartun aiheeseen pro gradussani.

Alustavasti ymmärrän datajournalismin isojen datamassojen pohjalta tehdyksi journalismiksi, jota tehdään yleensä verkkoon. Usein datajournalismiin liittyy hienoja visualisointeja ja interaktiivisuutta, mutta ei välttämättä. Parhaimmillaan datajournalismi on näkemykseni mukaan tutkivaa journalismia, sillä erilaisia datoja yhdistämällä ja analysoimalla voidaan synnyttää uutta tietoa. Pureudun datan, journalismin ja datajournalismin käsitteisiin tarkemmin luvussa 2.2.

Miksi datajournalismin sitten pitäisi kiinnostaa toimittajia? ”Ihmisten luottamus journalismiin on romahtanut, mutta he luottavat dataan. Numeroiden julkaiseminen lisää luottamusta meihin, sillä näytämme suoraan, mistä tietomme ovat peräisin”, The Guardianin Datablogin vetäjä Simon Rogers kertoo mielipiteensä Johanna Vehkoon kirjassa Painokoneet seis! (Vehkoo 2011, 106–107). Media Standards Trustin johtaja Martin Moore sanoo samassa kirjassa, että data on olennainen osa journalismin tulevaisuutta, koska datajournalismi auttaa selviytymään tiedon yltäkylläisyyden ajassa (Vehkoo 2011, 109–110).

Saksalainen datajournalisti Mirko Lorenz kirjoittaa teoksessa The Data Journalism Handbook, että datan käyttäminen muuttaa toimittajan työn painopistettä: toimittaja voi keskittyä jonkin tapahtuman tai kehityskulun vaikutuksiin sen sijaan, että kilpailee esimerkiksi silminnäkijöiden ja bloggaajien

kanssa siitä, kuka kertoo uutisen nopeimmin. Hän huomauttaa, että datajournalistiset taidot ovat myös kilpailuvaltti aikana, jolloin mediayhtiöt karsivat kuluja ja vähentävät väkeä. (Lorenz 2012.)

Kiinnostuin datajournalismista keväällä 2012, kun kävin Tampereen yliopiston ja Tampereen teknillisen yliopiston yhteisen datajournalismin kurssin. Kurssille mennessä en tiennyt, mitä datajournalismi on. Vuotta myöhemmin jatkoin asian opiskelua Otavan Opiston kurssilla. En kuitenkaan ole töissä ollessani itse tehnyt datajournalismia.

Ensimmäisen kurssin aikoihin aloin seurata suomalaista datajournalismikenttää. Seuraaminen oli helppoa, sillä mitenkään valtavasti ei tapahtunut: kenttää hallitsivat Yleisradio ja Helsingin Sanomat. Myös Suomen Kuvalehti teki jonkin verran datajournalismia, mutta sieltä tekijät karkasivat muualle töihin (ks. luku 4.1). Pikkuhiljaa datajournalistisia kokeiluja on alkanut näkyä muuallakin. Esimerkiksi STT-Lehtikuva teki ensimmäiset datajournalistiset juttunsa kesällä 2013.

Monessa ulkomaisessa mediassa datajournalismin tekeminen on kehittynyt nopeammin. Nykymuotoisen, 2000-luvun datajournalismin edelläkävijämaina pidetään yleisesti Yhdysvaltoja ja Iso-Britanniaa. Kun puhutaan datajournalismista, esiin nousee kerta toisensa jälkeen alan pioneerit, kuten The Guardian, The New York Times ja ProPublica. Myös saksalaisten Der Spiegelin ja Die Zeitin katsotaan usein kuuluvan tähän joukkoon. (Esim. Lehtonen 2011; Aitamurto, Sirkkunen & Lehtonen 2011; Gray, Chambers & Bounegru 2012.)

Journalistien keskuudessa datajournalismin mahdollisuuksia ovat avanneet esimerkiksi jutut, joita The New York Times, Der Spiegel ja The Guardian tekivät vuoden 2010 Afganistanin sotaa koskevista tietovuodoista (Aitamurto et al. 2011, 7). Yksi merkkipaalu oli The Guardianin vuonna 2009 lanseeraama Datablog, joka on keskittynyt datajournalismiin. Jo vuonna 2006 The Washington Postin silloinen ohjelmoija Adrian Holovaty kirjoitti bloggauksen, jossa hän kehotti sanomalehtiä muuttamaan verkkosivujaan perustavanlaatuisesti ja ottamaan datan vakavasti (Holovaty 2006). Monet pitävät Holovatyä termin datajournalismi isänä. (Howard 2014, 7.)

Yhdysvaltalaisen Alexander Howardin mukaan datajournalismista tuli vuonna 2014 valtavirtaa, kun uudet sivustot FiveThirtyEight.com ja Vox.com julkaistiin ja esimerkiksi The New York Times aloitti uuden datajournalismiin keskittyvän The Upshot -hankkeen (Howard 2014, 7, 16, 24). Näin voi olla Yhdysvalloissa, mutta Suomesta käsin väite tuntuu liioittelulta. Datajournalismi on kuitenkin yleistynyt hurjasti myös Euroopassa.

Suomessa datajournalismi on edelleen pienen piirin tekemisen varassa. Eniten sitä tekevät Helsingin Sanomat ja Yleisradio. Harva toimittaja on eturintamassa opettelemassa datajournalististen työkalujen ja -tapojen käyttöä, sillä se vaatii paljon aikaa ja vaivaa. Viime vuosina on kuitenkin alettu järjestää datajournalismiaiheisia luentoja, koulutuspäiviä ja kursseja eri oppilaitoksissa ja alan työpaikoilla. Journalistien seminaareissa ja konferensseissa datajournalismista tuntuu tulleen esitysten trendiaihe samaan tapaan kuin sosiaalinen media oli pari vuotta sitten.

Suomen pieni datajournalismiverkosto kommunikoi esimerkiksi Facebook-ryhmissä ja Twitterissä. Kokemukseni perusteella verkoston jäsenet toivottavat uudet datajournalismista kiinnostuneet tervetulleiksi mukaan ja neuvovat mielellään apua tarvitsevia. Kuulun esimerkiksi Datajournalismin avoin tukiryhmä -nimiseen Facebook-ryhmään, jossa jaetaan dataa, pyydetään neuvoja ongelmatilanteissa ja vinkataan kiinnostavista artikkeleista. Datajournalismista kiinnostuneet järjestävät myös erilaisia tapahtumia. Esimerkiksi Hacks/Hackers Helsingin tapaamisissa puhutaan journalismista, teknologiasta ja niiden kohtaamisesta.

Suomessa on muutama julkisuudessa datajournalistiksi identifioitunut henkilö, vahvimmin Yle Uutisten Teemo Tebest, Helsingin Sanomien Esa Mäkinen ja nykyään enimmäkseen Ruotsissa työskentelevä Jens Finnäs. Esimerkiksi Finnästä kirjoitettiin pari vuotta sitten Journalistissa (Valkama 2012) tulevaisuuden tekijänä ja datavelhona. Hän kertoo haastattelussa, että datajournalismista on ollut hänelle valtavasti hyötyä.

*Jos lähetän freelancerina juttuehdotuksen päätoimittajalle ja kerron, että olen toimittaja ja pidän kirjoittamisesta, reaktio on olematon. Jos taas sanon, että olen toimittaja ja pidän excelistä, olen heti kiinnostavampi. (Valkama 2012.)*

Finnäs ja Tebest ovat avanneet omaa matkaansa datajournalisteiksi blogeissaan. Se kuuluu datajournalismin henkeen: ideoita, oivalluksia ja löydettyjä työkaluja ei varjella mustasukkaisesti, vaan niistä kerrotaan eteenpäin. Blogeissa on esimerkiksi tutoriaaleja, luentomateriaaleja ja linkkejä avoimen datan resursseihin. Myös Mäkinen alkoi vuoden 2013 alussa julkaista kotisivuillaan muun muassa ohjeita datajournalistisiin harjoituksiin. Hän opetti tuolloin datajournalismin kurssilla Metropolia-ammattikorkeakoulussa.<sup>1</sup>

Kolmen pioneerin lisäksi datajournalismi työllistää monia muitakin. Esimerkiksi Tebest hoitaa Yle

---

<sup>1</sup> Finnäs kirjoittaa blogia osoitteessa <http://dataist.wordpress.com/> ja Tebest osoitteessa <http://datajournalismi.blogspot.fi/>. Mäkisen kotisivut ovat osoitteessa <http://www.esamakinen.fi/e/>.

Uutisissa sovellusten teknisen toteutuksen. Hänen lisäksi verkonomaisia sisältöjä tekevässä Plusdeskissä työskentelee kaksi tuottajaa ja kaksi graafikkoa. Verkonomaisilla sisällöillä tarkoitan juttuja, jotka on tehty verkon ilmaisukeinoja käyttäen. Jutuissa voi olla esimerkiksi hyperlinkkejä tai interaktiivisia elementtejä – siis asioita, jotka eivät toimisi muilla julkaisualustoilla (ks. tarkemmin luku 2.2.2).

Juttujen journalistisen puolen, esimerkiksi haastattelujen tekemisen ja artikkelin kirjoittamisen, voi hoitaa kuka tahansa Yle Uutisten toimittaja. On kiinnostavaa, että datajournalistiksi ei ole identifioitunut toimittajia, vaan koodaava Tebest. Kuka oikeastaan on datajournalisti? Mitä datajournalisti osaa ja tekee?

Termien määrittelemineen on tärkeää, jotta tekijät puhuvat samasta asiasta. Monet datajournalismin tekijät korostavat yhteistyön ja yhteisen kielen merkitystä. Toimittajan ei siis tarvitse opetella koodaamaan, vaan puhumaan koodaajien kanssa samaa kieltä, jotta he voivat tehdä yhteistyönä datajournalismia. Samalla tavalla on tärkeää, että tekijöillä on yhteinen käsitys siitä, mitä datajournalismi on.

Aloittaessani tämän tutkimuksen tekemistä syksyllä 2013 näytti siltä, että datajournalismia tehtiin usein välineet ja hienot visualisoinnit edellä, sisällön kustannuksella. Se on ymmärrettävää, onhan kyseessä varsin uusi ilmiö. Nyt, kesän 2015 kynnyksellä, Helsingin Sanomat ja Yleisradio tekevät jo hienostuneempia juttuja. On mielenkiintoista nähdä, millaiseksi datajournalismi ja sen merkitys muodostuvat.

## 1.2 Tutkimuskysymykset

Pohdin tutkimuksessani datajournalismin määritelmää ja merkitystä. Lisäksi luon yleiskatsauksen datajournalismin tilaan ja tekijöihin Suomessa. Tutkimuskysymykseni ovat:

1. Miten suomalaiset datajournalismin tekijät määrittelevät datajournalismin keväällä 2014?
2. Mitä merkitystä he näkevät datajournalismilla olevan journalismille, yleisölle ja yhteiskunnalle keväällä 2014?
3. Mitä Suomessa on tehty datajournalismin saralla syksyyn 2014 mennessä?



Etsin vastauksia kysymyksiin vahvasti oman aineistoni perusteella, sillä aiheesta on vain vähän aiempaa tutkimusta ja muuta kirjallisuutta. Ensisijainen aineistoni koostuu suomalaisten datajournalismin tekijöiden teemahaastatteluista. Haastattelin niin koodaajia kuin toimittajia. Lisäksi keräsin tausta-aineiston tekemällä vapaamuotoisia puhelinhaastatteluja ja kokoamalla muutamia lehtiartikkeleita. Tausta-aineistoon haastattelin datajournalismista vastaavia toimitusten edustajia.

Kaksi ensimmäistä kysymystä ovat tutkimukseni pääkysymyksiä. Etsin niihin vastauksia analysoimalla ensisijaista aineistoa, siis teemahaastatteluja. Tärkeiksi osoittautuivat niin haastateltavien omat kokemukset datajournalismista kuin heidän näkemyksensä datajournalismin määritelmästä ja merkityksestä. Oman lisänsä aineistoon tuovat datajournalistiset jutut, joita pyysin haastateltavia valitsemaan ja analysoimaan. Jokainen haastateltava valitsi omasta mielestään parhaan datajournalistisen jutun, jota on ollut tekemässä. Lisäksi peilasin keräämäni aineistoa aiempaan tutkimukseen ja muuhun kirjallisuuteen.

Kolmas tutkimuskysymys on apukysymys, joka tukee tutkimuksen tekoa. Etsin siihen vastausta ennen kaikkea tausta-aineistosta eli eri toimitusten edustajien kanssa käymistäni keskusteluista. Selvitin näiden keskustelujen avulla, millaisia datajournalismin tekemiseen liittyviä käytäntöjä toimituksissa on ja millaista datajournalismia toimituksissa tehdään. Selvitystyössä oli apua myös siitä, että olen seurannut suomalaista datajournalismikenttää melko aktiivisesti keväästä 2012 lähtien. Suomalaisesta datajournalismikentästä saa paljon tietoa esimerkiksi Twitteristä ja Facebookin datajournalismiaiheisista ryhmistä. Myös muutamat lehtiartikkelit auttoivat kokonais kuvan luomisessa.

### **1.3 Tutkimuksen rakenne**

Luvussa 2 kerron lyhyesti, millaista tutkimusta ja kirjallisuutta datajournalismista on olemassa, ja perehdyn tutkimukseni tärkeimpiin käsitteisiin ja konteksteihin. Erityisesti paneudun datajournalismin käsitteeseen. Ensin hajotan kuitenkin käsitteen osiin ja määrittelen datan ja journalismin. Vasta sitten esittelen erilaisia määritelmiä, joita datajournalismille on annettu. Palaan datajournalismin määritelmään tutkimukseni analyysissä.

Samassa luvussa myös kontekstoin datajournalismia kahdesta vastakkaisesta kulmasta. Toisaalta datajournalismin voi nähdä tämän ajan ilmiönä, verkkojournalismin ilmentymänä. Toisaalta taas

datajournalismi kiinnittyy tietokoneavusteisen ja tutkivan journalismin perinteisiin.

Luvussa 3 esittelen tutkimuksen aineiston ja menetelmät. Tutkimus on laadullinen. Haastattelin tutkimusta varten Yle Uutisten, Helsingin Sanomien ja STT-Lehtikuvan entisiä tai nykyisiä koodaajia ja toimittajia, jotka ovat osallistuneet datajournalististen juttuprojektien tekemiseen. Lisäksi keräsin tausta-aineiston keskustelemalla toimitusten edustajien kanssa siitä, millainen rooli datajournalismilla on toimituksissa. Tausta-aineistoon kuuluu myös muutama lehtiartikkeli. Analyysimenetelmä on sisällönanalyysi.

Luvuissa 4 ja 5 esittelen analyysin tuloksia. Analyysini ensimmäinen osa on kuvaileva. Vastaan kolmanteen tutkimuskysymykseen eli kerron, minkälaista datajournalismia Suomessa on tehty. Tällaista katsausta ei tietääkseni ole Suomessa aiemmin tehty. Vastaan kysymykseen erityisesti tausta-aineiston perusteella, ja siksi käsittelen asiaa vasta aineiston ja menetelmän esittelyn jälkeen eli luvussa 4.

Katsaus pohjustaa analyysin analyttisempaa osaa, jonka tuloksia esittelen luvussa 5. Siinä etsin vastauksia ensimmäiseen ja toiseen tutkimuskysymykseen. Käyn läpi haastateltavien näkemyksiä datajournalismin määritelmästä ja merkityksestä ja etsin niistä yhdistäviä piirteitä. Teen aineiston pohjalta tulkintoja siitä, mitä datajournalismi merkitsee suomalaisille tekijöilleen ja miten eri tavoin he määrittelevät sen.

Luvussa 6 tarkastelen analyysiä kriittisesti ja arvioin käyttämiäni menetelmiä. Vertaan tutkimukseni tuloksia aiempaan tutkimukseen ja pohdin, mitä uutta tutkimukseni paljastaa datajournalismista. Erityisesti arvioin analyysin toisen osan tuloksia ja merkittävyyttä. Teen tutkimustulosten ja aiemman kirjallisuuden pohjalta oman ehdotuksen siitä, miten datajournalismin voi määritellä. Lisäksi esittelen muutaman jatkotutkimusidean.

## 2 Tutkimuksellinen viitekehys

### 2.1 Aiempi tutkimus

Datajournalismia on tutkittu Suomessa vasta vähän. Ensimmäisenä aiheeseen on perehdytty Next Media -hankkeessa. Hankkeen raporteissa käsitellään paitsi esimerkiksi avointa dataa ja sen pohjalta tehtyjä sovelluksia (Lehtonen 2011; Kinnari, Lindman & Rossi 2012) myös datajournalismin trendejä (Aitamurto, Sirkkunen & Lehtonen 2011). Datajournalismin trendejä käsittelevä raportti *Trends in Data Journalism* painottuu ulkomaille.

Jyväskylän yliopiston viestintätieteiden laitoksella on parhaillaan tekeillä tutkimus datajournalismin työkäytännöistä. Tutkimusta vetää Turo Uskali. Helsingin Sanomien Säätiö myönsi maaliskuussa 2013 tutkimukselle 200 000 euron apurahan. (Helsingin Sanomien Säätiön... i.v.). Säätiön sivulla tutkimusta kuvataan näin:

*Datajournalismissa uutisoidaan, taustoitetaan ja visualisoidaan laajoja data-aineistoja. Tutkimuksessa selvitetään, miten datajournalismin käytäntöjä voidaan soveltaa journalistisessa tiedonhankinnassa ja julkaisutoiminnassa. Tutkimus toteutetaan datajournalismin ns. edelläkävijämaissa Yhdysvalloissa ja Isossa-Britanniassa ja myös Suomessa. (Helsingin Sanomien Säätiön... i.v.)*

Lisäksi datajournalismia oli käsitelty Suomessa joissakin opinnäytetöissä kevääseen 2014 mennessä, jolloin etsin niitä käsiini. Ilkka Toikkanen (2014) tutkii journalistiikan pro gradussaan datajournalistisia työprosesseja Yle Uutisissa, Svenska Ylessä ja Helsingin Sanomissa. Hän teki pro gradunsa osana Uskalin johtamaa tutkimushanketta. Minna Rapeli (2013) keskittyy tietojärjestelmätieteen pro gradussaan datajournalismin tuotantoprosessiin, mutta tekee myös lyhyen katsauksen siihen, mitä Suomessa on datajournalismin saralla tapahtunut. Tomi Kinnari (2013) käsittelee tietojärjestelmätieteen pro gradussaan avoimen datan liiketoimintamalleja media-alalla ja sivuaa tutkimuksessa datajournalismia.

Suomessa on myös tehty joitakin datajournalismiaiheisia kandidaatin tutkielmia. Samuli Leivonniemi (2012) käsittelee tutkielmassaan datajournalismia osana sanomalehden uutistuotantoa. Vilma Ruokoski (2013) on puolestaan tehnyt kuvailevan tapaustutkimuksen vuorovaikutteisuudesta ja joukkoistamisesta Helsingin Sanomien verkkosivuston datajournalismin aihealueella julkaistuista uutissovelluksista.

AMK-opinnäytetöitä on ilmestynyt useita. Riikka Koskenranta (2013) otti opinnäytetyönsä tavoitteeksi tutustuttaa pieniä uutistoimituksia datajournalismin tekemiseen. Heidi Kähkönen (2013) kertoo opinnäytetyössään, miten Haaga-Helian datajournalismin kurssi rakentui ja toteutettiin. Mika Minkkisen (2013) opinnäytetyö keskittyy infografiikan verkkojulkaisemiseen ja perinteisen tilastografiikan uusiin muotoihin. Minna Ala-Kyyny ja Susanna Patrikainen (2013) toimivat avustajina Kuningaskuluttaja-ohjelman yhden jutun teossa, hyödynsivät siinä joukkoistamista ja datajournalismia ja raportoivat kokemuksista yhteisessä opinnäytetyössään.

Yksikään näistä tutkimuksista ei keskity siihen, mitä datajournalismi oikeastaan on ja minkälainen merkitys sillä on Suomessa. Moni tehdyistä tutkimuksista käsittelee jollakin tavalla työprosesseja tai paneutuu yksittäisiin tapauksiin, esimerkiksi yhden kurssin järjestämiseen. Kokonaiskuva ei hahmotu. Siksi haluan haukata omassa tutkimuksessani ison palan ja pureutua peruskysymyksiin.

Myös kansainvälistä tutkimusta datajournalismista oli vielä vähän, kun aloin tehdä tätä tutkimusta syksyllä 2013. Yksi merkittävimpiä on Cindy Royalin *The Journalist as Programmer: A Case Study of The New York Times Interactive News Technology Department* (2010). Vuoden 2014 aikana on kuitenkin ilmestynyt lukuisia datajournalismia käsitteleviä tutkimusartikkeleita. Yksi perusteellisimmista on Alexander B. Howardin *The Art and Science of Data-driven Journalism* (2014). Mark Coddington vertailee artikkelissaan *Clarifying Journalism's Quantitative Turn* (2014) käsitteitä datajournalismi, laskennallinen journalismi ja tietokoneavusteinen journalismi. Lähimmäksi omaa tutkimustani tulevat neljässä maassa samoihin aikoihin tehdyt tutkimukset, joissa on selvitetty puolistrukturoiduilla haastatteluilla datajournalismin kehitystä toimituksissa. Maat ovat Belgia (De Maeyer, Libert, Domingo, Heinderyckx & Le Cam 2014), Norja (Karlsen & Stavelin 2014), Ruotsi (Appelgren & Nygren 2014) ja Yhdysvallat (Fink & Anderson 2014).

Ensimmäinen ja varsin perusteellinen käytännönläheinen datajournalismiopas on Jonathan Grayn, Lucy Chambersin ja Liliana Bounegrin *The Data Journalism Handbook* (2012). Sittemmin on ilmestynyt useita muita oppaita, esimerkiksi Claire Millerin *Getting Started with Data Journalism. Writing data stories in any size newsroom* (2013). Datajournalismista on kirjoitettu myös joissakin oppaissa, joissa käsitellään laajemmin journalismia ja verkkojournalismia. Tällainen on esimerkiksi Paul Bradshaw'n ja Liisa Rohumaan *The Online Journalism Handbook* (2011).

Palaan osaan näistä tutkimuksista ja teoksista myöhemmin.

## 2.2 Keskeiset käsitteet ja kontekstit

Keskeisin käsitteeni on datajournalismi. Määrittelen sen alustavasti kirjallisuuden pohjalta tässä luvussa. Aihetta käsittelevää kirjallisuutta on kuitenkin melko vähän, eikä käsitteelle ole vakiintunutta määritelmää. Palaan käsitteen määrittelemiseen uudestaan tutkimukseni lopussa, kun olen tehnyt teemahaastattelut.

Jotta saan paremman otteen datajournalismista, hajotan käsitteen ensin osiin ja teen lyhyen katsauksen siihen, miten data ja journalismi tavataan määritellä. Erityistä huomiota kiinnitän avoimen datan käsitteeseen. Lisäksi kontekstoin datajournalismia muutaman käsitteen kautta. Ilmeisin konteksti lienee verkkojournalismi, sillä datajournalistisia juttuja tehdään ennen kaikkea verkkoon, ja tekniikan kehittyminen on tuonut mukaan uusia ja entistä helpokäyttöisempiä työkaluja datajournalismin tekemiseen. Yksi luonteva osa verkkojournalismia ja datajournalismia on yleisön osallistaminen, jonka määrittelen myös tässä luvussa.

Datajournalismia voi lähestyä myös vastakkaisesta kulmasta. Se ei ole pelkästään tämän ajan ilmiö, vaan luonteva jatkumo tietokoneavusteiselle ja tutkivalle journalismille. Nämä käsitteet muodostavat toisen kontekstin, jonka nostan esiin tässä luvussa.

### 2.2.1 Journalismi internet-aikana

Journalismiin on tavattu liittää joukko määreitä, jotka eivät ole yksiselitteisiä. Esimerkiksi Risto Kunelius (2003) määrittelee journalismin näin: “Journalismi on ajankohtaista ja faktapohjaista joukkoviestintää.” Lisäksi hän liittää journalismiin määreet itsenäinen ja edustava. Kunelius toteaa kuitenkin, että journalismin määrittelemisen on perin juurin ongelmallista määreiden suhteellisuuden takia. (Kunelius 2003, 21–25.)

Internet-aikana journalismin määrittelemisen on entistä vaikeampaa, koska kuka tahansa voi julkaista tekstiä, ääntä, kuvaa, videota ja niiden yhdistelmiä maailmanlaajuiselle yleisölle. Bloggaajat saattavat kirjoittaa teräviä uutisanalyyssejä, ja monesta uutistapahtumasta saa nopeiten tietoa Twitteristä. Internet ei kuitenkaan ole ensimmäinen väline, joka haastaa journalistit ja saa heidät tuntemaan ammattinsa uhatuksi. Janet Jonesin ja Lee Salterin (2012) mukaan journalistit ja mediatalot ovat perinteisesti suhtautuneet skeptisesti teknologisiin innovaatioihin, oli kyse sitten televisiosta tai internetistä – ja

lopulta mukautuneet niihin. Samalla uudet välineet ovat muuttaneet ja monipuolistaneet journalistisia käytäntöjä. Se tekee journalismin määrittelemisestä vaikeaa, ja määrittely-yritykset kaatuvat pyrkimykseen olla liian yleisiä. Teknologian lisäksi määrittelyä vaikeuttaa eri alueiden ja maiden perinteet, normit ja eettiset koodistot. Jonesin ja Salterin mukaan on kuitenkin tiettyjä periaatteita, jotka journalistit jakavat kaikkialla. Näistä he mainitsevat erityisesti journalistien velvollisuuden kertoa totuus ja vuorovaikutuksen yleisön kanssa. (Jones & Salter 2012, 1–15.)

Esa Väliverronen listasi vuonna 2009 neljä syytä, joiden takia puhe journalismin kriisistä on tullut muotiin: 1) journalismista on tullut yhä selvemmin liiketoimintaa, 2) mediakentän sisäinen kilpailu on kiristynyt uusien medioiden ja ennen kaikkea internetin myötä, 3) journalismia haastavat pr-toimistot ja muut suhdetoiminnan ammattilaiset, jotka tuottavat journalismin kentälle lähes valmista, uutisilta näyttävää aineistoa ja 4) tavalliset kansalaiset, bloggarit ja verkkoyhteisöt haastavat journalismia. Murroksessa on siis niin journalismin taloudellinen perusta kuin kulttuurinen ja tekninen kehitys. Myös keskustelu journalismin poliittisesta tehtävästä haastaa perinteisiä käsityksiä. (Väliverronen 2009, 13–15.)

Haasteiden listauksen jälkeen Väliverronen muistuttaa kuitenkin, että journalismin raaka-aine eli uutisaiheet eivät tule loppumaan. Journalismi ei siis käy tarpeettomaksi, mutta se voi muuttua. Journalismin tulevaisuuden kannalta on olennaista, miten uutisaiheita halutaan jalostaa, kuka maksaa tuosta jalostamisesta, ja voivatko journalistiset instituutiot säilyttää asemansa julkisuuden ja uutisvirtojen ohjailijoina. Millaisia uusia journalistisia käytäntöjä syntyy? (Emt., 31.) Juuri tässä näen datajournalismin mahdollisuuden. Datan käyttäminen uutisten lähteenä ja jalostamisessa voi olla yksi journalistinen käytäntö muiden rinnalla. Visuaalinen näyttävyyttä, interaktiivisuus ja uudenlaiset kerrontatavat voivat olla valtteja muuttuvassa mediamaailmassa.

### **2.2.2 Kontekstina verkkojournalismi ja osallistava journalismi**

Internetin sanomalehdille ja muille perinteisille medioille asettama haaste on niin taloudellinen kuin sisällöllinenkin. Digitaalisella aikakaudella journalismin ansaintalogiikka on hukassa, sillä internetin ilmaistarjonta uhkaa murtaa journalismin perinteisen rahoituspohjan. Internet kiristää medioiden välistä sisällöllistä kilpailua, koska uutiset ja muut sisällöt leviävät entistä laajemmalle yleisölle. Lisäksi internet muuttaa yleisösuhdetta. Yleisösuhteen kannalta internet tarjoaa uusia mahdollisuuksia,

koska vuorovaikutus on entistä helpompaa ja nopeampaa, mutta samalla se vaatii perinteisiä mediataloja näkemään vaivaa yleisösuhteen eteen. (Väliverronen 2009, 21.)

Datajournalismin kannalta tilanne on paradoksaalinen: toisaalta monet mediatalot eivät halua laittaa rahaa datajournalismin tekemiseen taloudellisesti haasteellisena aikana, ja toisaalta datajournalismi voi olla yksi tapa erottautua kilpailijoista. Esimerkiksi Yle Uutiset on viime aikoina panostanut syvällisiin verkonomaisiin sisältöihin, myös datajournalismiin. Sisältöjä tekee Plus-deski, jota kuvataan näin:

*Kokoamme tälle sivulle taustoittavia juttujamme, jotka syventävät Ylen uutis-, urheilu- ja ajankohtaisaiheita. Kokeilemme jutuissa usein erilaisia verkon ilmaisukeinoja. Kehitämme sivua ja erilaisia juttutyyppejä jatkuvasti. (Mikä on Plus? i.v.)*

Plus-deski oli mukana tekemässä kolmea viidestä Yle Uutisten vuoden 2013 suosituimmasta jutusta (Laakso & Pirhonen 2013). Kaikkein luetuin oli Pisa-testi, jossa sai kokeilla, pärjääkö matematiikassa ysiluokkalaisille. Plus-deski oli tekemässä myös toiseksi sijoittunutta painoindeksilaskuria ja neljänneksi sijoittunutta tulotietolaskuria. Kolmanneksi tullut juttu on kollaasi poliisin kuvista, joissa on kuvattu erikoisia teillä liikkuvia kuormia, ja viides juttu on perinteistä tekstiä, jonka aiheena oli maailman vaikeimmat kielet. Niiden tekemisessä Plus-deski ei ollut mukana.

Kaikki Plus-deskin jutut eivät välttämättä ole datajournalismia, vaan muita verkonomaisia sisältöjä. Näyttää siltä, että verkonomaisten sisältöjen ja datajournalismin kehitys kulkee usein käsi kädessä. Datajournalismin voi ajatella olevan yksi tapa tehdä verkonomaisia sisältöjä.

Mutta mitä kaikkea verkkojournalismi tai verkonomaiset sisällöt oikein ovat? Bradshaw ja Rohumaa (2011) luettelevat kahdeksan verkkojournalismin erityispiirrettä, jotka erottavat sen muusta journalismista: 1) journalistit käyttävät internetiä juttujen alustana, 2) hyperlinkit, 3) ei-lineaarinen kerronta, 4) multimedia, 5) journalistit voivat kertoa jutun reaaliaikaisesti ja julkaista jutun samanaikaisesti useilla sivustoilla ja alustoilla, kuten Twitterissä, blogissa ja uutissivustolla, 6) interaktiivisuus, 7) informaation helppo saatavuus missä vain ja mihin aikaan tahansa ja 8) läpinäkyvyys esimerkiksi hyperlinkkien avulla. (Bradshaw & Rohumaa 2011, 30–31.)

Verkkojournalismin tekeminen vaatii siis erityisosaamista. Journalistien pitää esimerkiksi osata käyttää hyperlinkkejä ja erilaisia alustoja. Jotta journalistin kirjoittama artikkeli löytyy verkosta, journalistin pitää ymmärtää hakukoneoptimointia ja kirjoittaa artikkeli sen mukaisesti. Multimedian tekeminen vaatii sitä, että osaa tuottaa niin tekstiä, kuvaa, liikkuvaa kuvaa kuin ääntä. (Emt., 30–31.)

Interaktiivisuus on verkkojournalismin erityispiirre, jota käytetään paljon myös datajournalismissa. Bradshaw'n ja Rohumaan mukaan yksi syy siihen, että interaktiivisuus on niin keskeistä verkkojournalismissa, on internetin käyttökulttuuri: ihmiset menevät verkkoon *tekemään* jotain (emt., 120). Toisin sanoen ihmiset eivät halua pelkästään lukea ja katsoa, vaan myös osallistua.

Jane B. Singer et al. (2011) käyttää käsitettä *participatory journalism* (suomeksi osallistava journalismi) kuvaamaan verkkolehtien käyttäjien panosta verkkosivuilla. Osallistuminen voi näkyä eri vaiheissa uutisprosessia, ja siihen voidaan käyttää eri työkaluja. Tässä määritelmässä osallistava journalismi sisältää niin verkkolehtien kommentit kuin kansalaisjournalismin (*citizen journalism*) ja muun käyttäjälähtöisen sisällön (*user-generated content*). (Singer et al. 2011, 206.)

Singerin et al. mukaan osallistava journalismi ei ole yhdensuuntaista vaan yhteistyössä tehtyä. Niin toimittajat kuin yleisö osallistuu verkkolehden luomiseen ja yhteisön rakentamiseen sen ympärille. (Emt., 2.) Sen voi nähdä uhkaavan toimittajien roolia tiedon ja uutisten portinvartijana, mutta toimittaja voi myös valjastaa yleisön avukseen. Joukkoistaminen (*crowdsourcing*) tarkoittaa käytäntöä, jossa toimittaja pyytää yleisöä osallistumaan juttuprosessiin haluamallaan tavalla (Hermida 2011, 20). Datajournalismissa toimittaja voi esimerkiksi kerätä dataa pyytämällä yleisöä vastaamaan kyselyyn tai merkkamaan jotakin interaktiiviseen karttaan (ks. joukkoistamisesta lisää luvusta 2.2.4).

### **2.2.3 Kontekstina tietokoneavusteinen journalismi ja tutkiva journalismi**

Vaikka internet on luonteva alusta datajournalismille, datajournalismi ei ole pelkästään tämän ajan ilmiö. Sen voi myös nähdä osana tutkivan journalismin perinnettä, sillä molemmissa pyritään parhaimmillaan synnyttämään uutta tietoa eikä tyydytä tiedon välittämiseen. Heikki Kuutti (2002) rajaa tutkivan journalismin omaksi toimitustyön alueeksi kolmella rajauksella. Ensinnäkin, tutkiva journalismi perustuu toimituksen omaan tutkimukseen. Muunlaisen journalismin taustalla voi olla muiden tahojen tekemää tutkimusta tai ei tutkimusta lainkaan. Toiseksi, tutkiva journalismi kohdistuu yleisesti ja yhteiskunnallisesti merkittäviin asioihin. Kolmanneksi, tutkiva journalismi käsittelee salattua tai kiusallista tietoa. (Kuutti 2002, 13–26.)

Kuutin mukaan tutkivalla journalistilla on kolme pääasiallista tiedonhankintakeinoa: dokumenttilähteet, henkilölähteet ja havainnointi (2002, 77). Erityisesti dokumenttilähteisiin perustuva tutkiva journalismi menee osin päällekkäin datajournalismin kanssa. Kuutti ennustikin vuonna 2002,



että “kairaus” (computer assisted investigative reporting, cair) mullistaa tutkivan journalismin. Kairauksella hän tarkoitti toimitusten omien tietokantojen atk-pohjaista analysointia. (Emt., 91.) Kairauksessa siis yhdistyy tietokoneavusteinen journalismi (computer assisted journalism, car) ja tutkiva journalismi.

Vaikka Kuutti käytti eri käsitteitä, kairaus osuu monin tavoin lähelle datajournalismia. Kuutti vertailee taulukossa tutkivan journalismin ja tiedon kairauksen piirteitä, ja monet tiedon kairauksen piirteet sopisivat yhtä hyvin datajournalismiin. Esimerkiksi käsiteltävä tietomäärä on tiedon kairauksessa “epäinhimillisen” mammuttimainen, kun tutkivassa journalismissa käsiteltävä tietomäärä on “inhimillisen” kokoinen. Tiedon kairauksessa tarvitaan teknistä tukea, tutkivassa journalismissa ei. (Emt., 95.)

Tutkiva journalismi sai alkunsa Yhdysvalloissa 1800- ja 1900-lukujen taitteessa. Suomessakin käsite on tunnettu laajalti 1950- ja 1960-lukujen vaihteesta asti, joskin se arkipäiväistyi Suomessa vasta 1990-luvulla. (Emt., 34–37.) Myöskään tietokoneavusteinen journalismi ei ole uusi ilmiö. Yhdysvaltalainen televisio- ja radioyhtiö CBS käytti sitä ensimmäisen kerran vuonna 1952 presidentinvaalien tulosten ennustamiseen. Datan käyttäminen journalismissa ulottuu vielä kauemmas menneisyyteen: Simon Rogers päivää *The Guardian* -lehden ensimmäisen esimerkin datajournalismista vuodelle 1821. Se on vuodettu taulukko Manchesterin kouluista, jossa listataan kouluittain oppilaiden määrät ja kulut. Rogersin mukaan taulukon avulla pystyi ensimmäistä kertaa näyttämään toteen ilmaisen koulutuksen saavien oppilaiden määrän, joka oli paljon virallisia lukuja suurempi. (Bounegru 2012, 18–19.)

Dataan pohjautuva journalismi on saanut aikojen saatossa monenlaisia nimityksiä, esimerkiksi precision journalism (suomeksi tarkkuusjournalismi). Käsitteen teki tunnetuksi toimittaja Philip Meyer, jonka kirja *Precision Journalism* ilmestyi jo vuonna 1973. Kirjasta ilmestyi vuonna 1991 uudistettu painos nimellä *The New Precision Journalism*, jossa Meyer kuvailee tarkkuusjournalismin olevan tieteellistä journalismia (scientific journalism): se vaatii tieteellisten metodien, objektiivisuuden ja ideaalien omaksumista koko journalistiseen prosessiin. Meyerin mielestä toimittajan pitää osata hallinnoida tietokantoja ja käsitellä ja analysoida dataa. (Meyer 1991, 1–6.) Hän itse käytti tarkkuusjournalismin metodeja jo 60-luvulla. Kuuluisassa työssään hän osoitti, että Detroitin vuoden 1967 mellakoihin eivät osallistuneet vain kouluttamattomat vaan yhtä lailla koulutetut. (Bounegru 2012, 18–19.)

Tunnetuin datajournalismin sukulaiskäsite lienee kuitenkin tietokoneavusteinen journalismi. Mutta onko nykymuotoinen datajournalismi ja tietokoneavusteinen journalismi sama asia? Joidenkin mielestä on, joidenkin mielestä ei, Bounegru vastaa kysymykseen. (Emt., 18–21.)

Joidenkin mielestä tietokoneavusteinen journalismi on datan keräämisen ja analysoimisen tekniikka, jolla parannetaan (yleensä tutkivaa) journalismia, kun taas datajournalismi huomioi tavan, jolla data istuu journalistiseen työkulkuun. Tässä mielessä datajournalismi kiinnittää yhtä paljon tai jopa enemmän huomiota dataan itseensä kuin datan käyttämiseen tarinan löytämisessä tai parantamisessa. Toinen ero on se, että ennen tutkivat journalistit käyttivät paljon energiaa datan etsimiseen ja keräämiseen. Nykyään dataa on enemmän saatavilla, jolloin sen prosessointi on entistä tärkeämpää. (Emt., 21.)

Toisten mielestä datajournalismilla ja tietokoneavusteisella journalismilla ei ole merkittävää eroa. Bounegru huomauttaa, ettei ole niinkään hedelmällistä pohtia, onko datajournalismi uutta vai ei. Sen sijaan pitäisi nähdä datajournalismi osana perinnettä, mutta samalla vastauksena uusiin olosuhteisiin. Hän näkee nykymuotoisen datajournalismin synnyn 2000-luvulla merkitsevän uuden vaiheen alkua. Vaiheelle on ominaista vapaasti saatavilla olevan datan runsas määrä, käyttäjäystävälliset työkalut ja joukkoistamisen keinot. Ne mahdollistavat sen, että yhä useampi ihminen voi työskennellä yhä suurempien datamäärien kanssa helpommin kuin koskaan ennen. (Emt., 21.)

#### **2.2.4 Data, avoin data**

Englantilaisen Paul Bradshaw'n (2012) mukaan data tarkoitti aikaisemmin kokoelmaa numeroita, jotka oli kerätty yleensä taulukkoon. Nykyisessä digitaalisessa maailmassa puolestaan melkein kaikki sisältö voidaan ja usein kuvailtaankin numeroina, ykkösinä ja nollina, siis digitaalisena informaationa. (Bradshaw 2012, 2–3; Heinonen 1999, 37.)

Data on digitaalisessa muodossa olevaa raaka-ainetta. Se on digitaalisesti tallennettua informaatiota, jota voi käsitellä tietokoneiden avulla. Datasta voidaan jalostaa merkityksellisempää informaatiota, ja informaatio puolestaan muuttuu tiedoksi, kun vastaanottaja tulkitsee sitä. Data on siis alimpana tiedon arvoketjussa: sen yläpuolella ovat järjestyksessä informaatio, tieto, tietämys ja viisaus. (Poikola, Kola & Hintikka 2010, 14.) Yksi keino jalostaa dataa on datajournalismi.

Mistä toimittajat sitten saavat dataa? Yksi vaihtoehto on avoin data. Jotta data olisi avointa, sen pitää täyttää tiettyjä vaatimuksia. Poikola (2013) asettaa luentokalvoissaan avoimelle datalle kolme vaatimusta: sen täytyy olla laillisesti, teknisesti ja maksutta uudelleenkäytettävää. Tämä on kuitenkin hyvin yksinkertaistettu määritelmä. Poikola, Kola ja Hintikka (2010, 34) esittävätkin Julkinen data -oppaassa isomman joukon kriteereitä, joiden pohjalta datan julkisuutta voi arvioida. Kriteerit ovat 1) löydettävyys, 2) kokonaisuus, 3) käyttöehtojen tasa-arvoisuus, 4) alkuperäisyys ja ajantasaisuus, 5) laillinen ja vapaa uudelleenkäytettävyys, 6) maksuttomuus, 7) koneluettavuus, 8) formaatin avoimuus ja 9) ymmärrettävyys. Määritelmä pohjautuu Open Knowledge Definition -määritelmään<sup>2</sup>, ja purkamalla se osiin voidaan arvioida yksittäisen datalähteen avoimuutta. (Poikola et al. 2010, 34.)

Avoin data on yksi datalähde, mutta ei suinkaan ainoa. Poikolan (2013) mukaan journalistit voivat saada dataa käyttöönsä myös tietopyynnöillä, ruudun raavinnalla ja joukkoistamisella. Selvennän hieman näitä kolmea käsitettä.

Tietopyyntö tarkoittaa sitä, että journalisti voi pyytää nähtäväkseen julkisen hallinnon edustajalta julkisia asiakirjoja, jotka eivät ole avoimesti saatavilla netissä. Laki viranomaisten toiminnan julkisuudesta säätelee viranomaisten hallussa olevien asiakirjojen julkisuutta ja salassapitoa, ja lain mukaan tietopyyntöön pitäisi vastata kahdessa viikossa tai enintään kuukaudessa. (Laki viranomaisten toiminnan... i.v.) Asiakirjojen eli datan ei tarvitse olla netissä, ja viranomaisen tekee työn, mutta tietopyyntöihin liittyy myös ongelmia: datasta saattaa joutua maksamaan, eivätkä kaikki viranomaiset suhtaudu kovin suopeasti tietopyyntöihin (Poikola 2013).

Ruudun raavinta (englanniksi screen scraping) tarkoittaa datan ”raapimista” tavallisilta nettisivuilta siihen tarkoitettun ohjelman tai koodaamisen avulla ja sen muuttamista koneluettavaan muotoon (Lindenberg 2012, 124–125). Otetaan esimerkki: Wikipediassa on luettelo Suomen kunnista ja niiden asukasluvuista, mutta se ei ole koneluettavassa muodossa. Kun asennan nettiselaimeni ”raapimen” (scraperin), saan sillä kopioitua luettelon, vietyä sen taulukkolaskentaohjelmaan ja tallennettua haluamassani muodossa. Sen jälkeen voin käyttää taulukkoa datana.

Aiemmin kerroin, että toimittaja voi joukkoistaa juttuprosessin pyytämällä yleisöä osallistumaan siihen esimerkiksi vastaamalla kyselyyn tai merkkäämällä jotakin interaktiiviseen karttaan. Wikipedian

---

<sup>2</sup> Alkuperäinen määritelmä löytyy osoitteesta <http://opendefinition.org/od/>. Siinä on hieman enemmän kriteereitä, ja ne ovat hieman eri tavalla ilmaistuja. Pidän Poikolan et al. määritelmää selkeämpänä.

mukaan joukkoistaminen tai joukkouttaminen (crowdsourcing) tarkoittaa ”yleisesti hajautettua ongelmanratkaisu- ja tuotantomallia, jossa toimeksiantaja hyödyntää yhteisön tietotaitoa rajattua tehtävää varten” (Joukkouttaminen i.v.). Toisin sanoen on mahdollista kerätä omaa dataa käyttämällä hyödyksi jotakin yhteisöä. Linaan tässä Wikipedian määritelmää, koska siihen viitataan niin usein – esimerkiksi Antti Poikolan luentokalvoissa ja Data Journalism Handbookissa. Kevyehkö esimerkki joukkoistamisesta on Helsingin Sanomien joulukuussa 2011 avaama Terveisiä maailmalta -palvelu, jonne ulkosuomalaiset voivat lähettää sähköisen postikortin ja kertoa, missä asuvat ja mitä siellä tekevät. Näin Helsingin Sanomat sai kerättyä dataa ulkosuomalaisista, ulkosuomalaisten itsensä avulla. (Terveisiä maailmalta 2011.)

Tietopyyntöjen, ruudun raavinnan ja joukkoistamisen tarkastelu osoittaa, että datajournalismi ei automaattisesti kytkeydy avoimeen dataan. Avoin data on kuitenkin yksi ilmeinen datajournalismin lähtökohta. Tuo lähtökohta tuo mukanaan myös kysymyksiä. Toimittajien lisäksi avoin data -aktiivit tekevät datajournalismiksi laskettavia tai sitä muistuttavia tuotoksia. Onko silloin kyse journalismista? Onko esimerkiksi alan eettinen koodisto tuttu avoimesta datasta innostuneelle koodaajalle? Voiko koodaustaitoista henkilöä, jolla ei ole journalistista koulutusta tai kokemusta, sanoa datajournalistiksi?

Kiinnostus avointa dataa kohtaan kasvaa jatkuvasti. Siitä ovat Suomessakin kiinnostuneet niin valtion viranomaiset ja liike-elämän vaikuttajat kuin yksittäiset kansalaisetkin. (Lehtonen 2011, 4.) Next Media -tutkimusohjelma esittelee joukon suomalaisia avoimen datan pohjalta tehtyjä sovelluksia, kuten Hahmota Oy:n kehittämän talousdataa visualisoivan veropuusovelluksen (Kinnari ym 2012, 14–15, 22–23). Tutkimusohjelman raportissa tiivistetään, että avoimen datan kaupallinen hyödyntäminen on Suomessa edelleen lapsenkengissä, mutta lupaavia esimerkkejä löytyy jo (Emt., 2).

Mikä on journalismin rooli tällä kentällä? Kinnari et al. (2012(2)) luovat katsauksen siihen toisessa raportissaan. He huomauttavat mediayritysten aseman olevan sikäli erityinen, että niillä on suora yhteys loppukäyttäjiiin, siis yleisöön. Niinpä mediayrityksillä on ainakin kolme mahdollisuutta hyödyntää dataa liiketoiminnassaan: 1) datan käyttäminen datajournalismissa, 2) kolmannen osapuolen data-analyysien käyttäminen ja 3) oman datan julkaiseminen. (Kinnari et al. 2012(2), 40–43.)

Taloudellisen puolen lisäksi datan julkisuuteen liitetään usein yhteiskunnallisia merkityksiä. Kinnari et al. perustelevat sitä, miksi mediayritysten pitäisi julkaista dataa, kustannussäästöjen ja tehokkuuden lisäksi läpinäkyvyydellä (2012(2), 46). Erityisesti avaamispainetta liittyy kuitenkin julkisiin

datavarantoihin. Poikola, Kola ja Hintikka näkevät julkisten datavarantojen avaamisen edistävän ainakin kolmea eri tavoitetta: demokratian ja hallinnon läpinäkyvyyden lisäämistä, uusien markkinoiden ja innovaatioiden syntymistä ja hallinnon sisäisen tehokkuuden lisäämistä (2010, 12). Datan avaamista ajavassa liikkeessä on siis samalla lailla yhteiskunnallisuutta kuin journalismissa. Erityisesti demokratian ja hallinnon läpinäkyvyys lienee yhteinen tavoite journalismille ja avoin data -aktiiveille.

Julkisten datavarantojen avaamisesta on viime vuosina käyty paljon keskustelua niin Suomessa kuin muualla maailmassa. Yhdysvalloissa ja Isossa-Britanniassa hallinnon datan avaaminen on otettu jo osaksi toimintapolitiikkaa ja strategioita. (Emt., 6.) Kansainvälinen kiinnostus avointa dataa kohtaan on innoittanut myös suomalaisia toimijoita ja julkishallintoa. Liikenne- ja viestintäministeriö on nostanut keskusteluun julkisten datavarantojen roolin ja pääsyn niihin, ja asia on nostettu kansallisen tietoyhteiskunnan strategiaan. Vuonna 2010 valtioneuvosto asetti periaatteeksi, että julkisen sektorin informaation pitäisi olla vapaasti saatavilla. (Lehtonen 2011, 4–5.) Datan avaamista ajavat useat organisaatiot ympäri maailmaa, esimerkiksi Open Data Institute ja Open Knowledge Foundation (Open data i.v.).

### 2.2.5 Datajournalismi

Datajournalismi on tavallaan hyvin helppo käsite: yksinkertaistettuna se on journalismia, jota tehdään datan pohjalta. Nähdäkseni kolme asiaa tekee datajournalismista kuitenkin monimutkaisen. Ensinnäkin, se on niin tuore käsite, että tällä hetkellä datajournalismiksi lasketaan hyvin sekalainen joukko erilaisia asioita. Toiseksi, myös muut kuin journalistit tekevät datajournalismin kaltaisia tuotoksia. Voidaanko ne laskea datajournalismiksi? Kolmanneksi, kuten tässä luvussa on jo käynyt ilmi, sekä data että journalismi ovat problemaattisia käsitteitä.

*What is data journalism? I could answer, simply, that it is journalism done with data. But that doesn't help much.*

*Both 'data' and 'journalism' are troublesome terms. (Bradshaw 2012, 2.)*

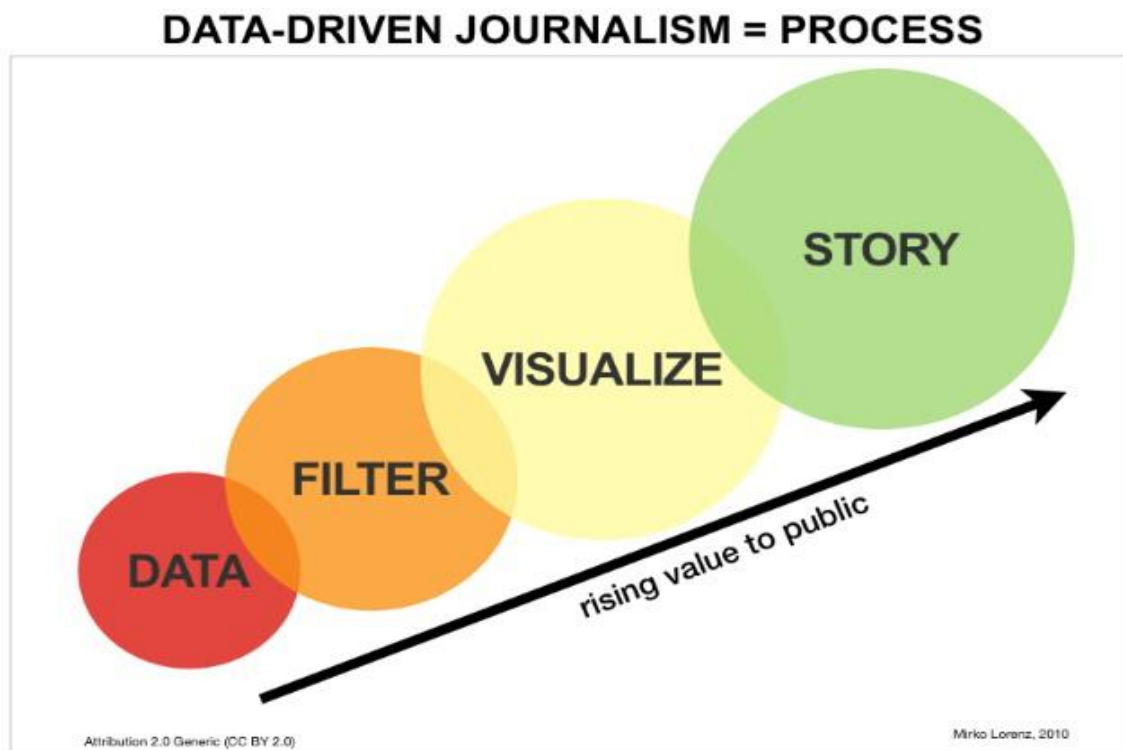
Bradshaw'n (2012) mukaan datajournalismin erottaa muusta journalismista ehkä uudet mahdollisuudet, jotka aukeavat, kun yhdistää perinteisen uutisen ja kyvyn kertoa juttu nykyisin saatavilla olevan informaation pohjalta. Nuo mahdollisuudet voivat näkyä missä vain journalistisen prosessin vaiheessa. Datajournalisti voi esimerkiksi ohjelmoida automaattiseksi prosessin, jossa

kerätään ja yhdistellään poliisilta saatuja tietoja. Tai datajournalisti voi käyttää ohjelmistoa löytääkseen yhteyksiä sadoista tai tuhansista dokumenteista. Bradshaw'n mukaan data voi olla datajournalismin lähde tai työkalu, jolla tarina kerrotaan, tai molempia. (Bradshaw 2012, 2–3.)

Mirko Lorenz määrittelee datajournalismin seuraavasti:

*Today data-driven journalism can be defined as a workflow, where data is the basis for analysis, visualization and – most importantly – storytelling. (Lorenz 2010, 10.)*

Hänen määritelmässään korostuu siis prosessi, jossa raakaa dataa muokataan siten, että siitä tulee yleisölle merkityksellistä (ks. kuvio 1).



*Kuvio 1: Data-driven journalism = process (Lorenz 2010, 13).*

Lorenzin mukaan juuri prosessi erottaa datajournalismin tietokoneavusteisesta journalismista: jälkimmäinen on enemmän tekniikka, kun taas edellinen on prosessi, joka vaikuttaa perustavanlaatuisesti koko journalistisen työn kulkuun. Datajournalismiin liittyy kuitenkin vielä avoimia kysymyksiä. Ensinnäkin, nykyisillä alustoilla on kyllä helppo kirjoittaa artikkeli, mutta ne eivät huomioi dataa kovin hyvin. Millaisilla alustoilla datan saisi paremmin näkyviin? Toiseksi,

millaisia liiketoimintamalleja datan ympärille voisi rakentaa? Miten toimittajat voivat saada elantonsa datasta? (Lorenz 2010, 10–11.)

Lorenzin määritelmä on melko yksinkertaistettu. On huomioitava, että Lorenz kirjoitti datajournalismista vuonna 2010. Sen jälkeen esimerkiksi datajournalistiset työkalut ja verkonomainen kerronta – Lorenz liittää datajournalismin nimenomaan verkkoon – ovat kehittyneet valtavasti. Nykyään on ehkä helpompi saada data paremmin esiin tekstin seasta, vaikka tekemistä riittää edelleen sillä saralla.

Toinen huomionarvoinen seikka on se, että Lorenz käyttää käsitettä data-driven journalism (ja sen lyhennettä ddj), eikä siis käsitettä data journalism. Molemmat kääntyvät suomeksi datajournalismiksi. Nähdäkseni käsitteissä on pieni vivahde-ero: data-driven journalism viittaa siihen, että journalismi on *datalähtöistä*, kun taas data journalism -käsitteessä datalla on hiukan suurempi painoarvo, eli data voi olla mukana missä tahansa journalistisen prosessin vaiheessa. Käsitteitä käytetään kuitenkin päällekkäin. En tässä tutkimuksessa tee eroa käsitteiden välille, vaan käytän yksinkertaisesti käsitettä datajournalismi.

Yhdysvaltalainen Cindy Royal (2010) tutki teknologian roolia jatkuvasti muuttuvassa uutistoimituksessa samoihin aikoihin, kun Lorenz teki ensimmäisiä datajournalismin määrittelyyrityksiä. Royal havainnoi ja haastatteli The New York Timesin Interactive News Technology -osaston tekijöitä kesäkuussa 2009. Osaston ensisijainen tehtävä on tehdä verkkoon uutissovelluksia, jotka saavat niiden käyttäjät osallistumaan ja joissa käytetään usein dataa. Osaston perustamisen tarkoitus oli saattaa yhteen joukko koodari-toimittajia. Osastolle palkattiin siis tekijöitä, joilla on vahvaa teknistä osaamista, mutta jotka samaan aikaan ovat intohimoisia toimittajia. Suurin osa tekijöistä kuvaili haastatteluissa olevansa itseoppineita. (Royal 2010, 2–12.)

Royal ei käytä tutkimuksessaan termiä datajournalismi, vaan puhuu datavetoisesta, interaktiivisesta uutisten esittämisestä (data-driven interactive news presentations). Hän mainitsee muutaman kerran myös tietokoneavusteisen journalismin. Tämä johtunee osaltaan tutkimuksen tekoajankohdasta: datajournalismi nousi muotitermiksi vasta vuoden 2010 jälkeen. Vaikuttaa kuitenkin siltä, että Royalin haastatteleminen koodari-toimittajien työssä on paljon elementtejä, joita yhdistetään nykyään usein datajournalismiin: dataa, koodaamista, toimittamista, yhteistyötä graafikoiden kanssa, interaktiivisuutta ja niin edelleen.

Royal kysyy tutkimuksessaan, mitä ilmenee, kun yhdistää journalistisen ja teknologisen kulttuurin. Hänen mukaansa Interactive News Technology -osaston kulttuurin juuret ovat luovuudessa ja innovatiivisuudessa, ja ne kumpuavat pitkälti avoimuutta painottavasta hakkerikulttuurista. Osasto tekee paljon yhteistyötä muiden osastojen, esimerkiksi grafiikan kanssa. Yhteistyön hedelminä syntyy asioita, joita kukaan muu ei ole tehnyt. Luovuus, innovatiivisuus, avoimuus ja yhteistyö ovatkin edellytyksiä osaston menestykselle. (Emt., 16–18.) Ne ovat myös ominaisuuksia, jotka on helppo liittää datajournalismiin. Samalla ne voivat muuttaa journalismin tekemisen tapaa. Royal ennusti vuonna 2010, että tutkimuksessa esiin tulleiden työtapojen ansiosta laskennallisesta journalismista (computational journalism) innostutaan uudestaan ja että erilaiset ihmiset kiinnostuvat journalismin koulutuksesta ja urasta. Nämä erilaisista taustoista ja kulttuureista tulevat ihmiset puolestaan vaikuttavat uutiskulttuuriin. (Emt., 22.) Näin näyttää jossain määrin käyneen. Suomessakin isojen mediatalojen toimituksiin on palkattu ohjelmointitaitoisia ihmisiä, jotka tekevät journalismia hieman erilaisista lähtökohdista kuin perinteisesti on tehty. Suomessa ei juuri puhuta laskennallisesta journalismista, vaan datajournalismista.

Useat tutkijat ja journalistit ovat yrittäneet saada datajournalismista otetta Lorenzin ja Royalin jälkeen. Yhdysvaltalaisen Alexander Howardin raportti (2014) on yksi uusimmista ja perusteellisimmista katsauksista datajournalismiin. Täytyy kuitenkin muistaa, että se on tehty yhdysvaltaisesta näkökulmasta. Esimerkiksi Howardin väite, että datajournalismista on tullut vuonna 2014 valtavirtaa, tuntuu Suomesta käsin liioittelulta (Howard 2014, 7, 16, 24). Raportista käy myös ilmi, että Euroopassa ja Yhdysvalloissa datajournalismi käsitetään usein eri tavalla. Tunnetun yhdysvaltalaisen datajournalistin Aron Pilhoferin mukaan Euroopassa datajournalismilla tarkoitetaan useimmiten visualisointeja, kun taas Yhdysvalloissa termillä viitataan joskus visualisointeihin, joskus tietokoneavusteiseen journalismiin. (Emt., 109).

Howardin mukaan vuonna 2014 kysymys ei ole, voivatko journalistit hyödyntää dataa, tietokoneita ja algoritmeja, vaan miten, milloin, missä, miksi ja kenen toimesta niitä hyödynnetään. Journalistit voivat käyttää dataa lähteenä ja haastatella sitä kuten ihmistä. Yhtäläillä kuin ihmislähteisiin, myös dataan pitää suhtautua kriittisesti, ei faktana. Tätä työtä Howard kutsuu datajournalismiksi, siis datan keräämiseksi, puhdistamiseksi, järjestämiseksi, analysoimiseksi, visualisoimiseksi ja julkaisemiseksi. Vielä ytimekkäämmin hän määrittelee datajournalismin olevan datatieteen soveltamista journalismiin. Perusmuodossaan datajournalismi koostuu kolmesta asiasta: 1) datan keräämisestä lähteeksi ja



validoinnista, 2) tilastojen haastattelemisesta ja 3) niiden visualisoimisesta. Joidenkin mielestä prosessiin kuuluu avoin data -hengessä neljäs osio, jossa toimittajat arkistoivat ja julkaisevat tutkimustensa taustalla olevan raakadatan ja sen analysointimenetelmän. Laajassa mielessä datajournalismi on juttujen kertomista numeroiden avulla tai juttujen etsimistä numeroista. (Howard 2014, 4–5, 19.)

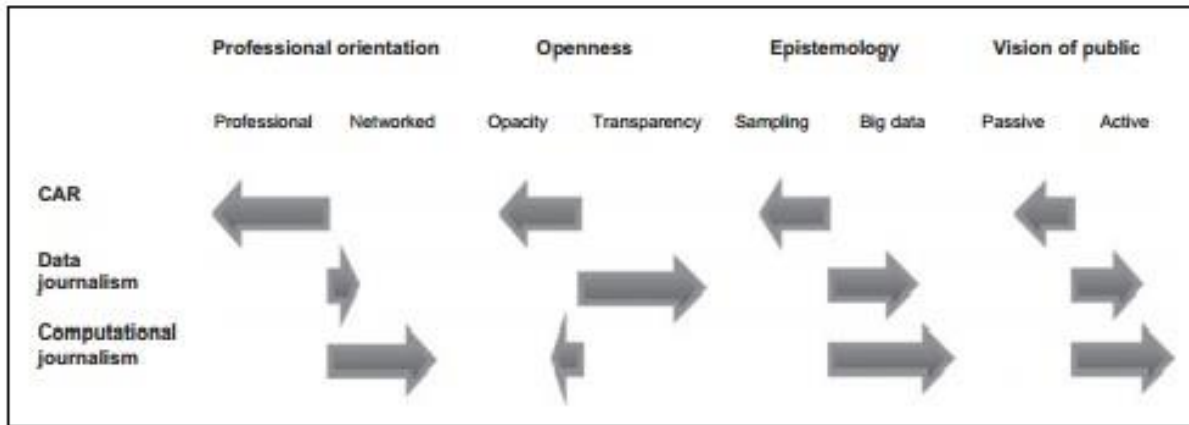
Howard näkee datajournalismin olevan looginen jatkumo tietokoneavusteiselle journalismille. Kun tietokoneavusteisessa journalismissa on perinteisesti keskitytty datan analysoimiseen, datajournalismi sisältää myös datan julkaisemisen, kierrättämisen ja käytettävyyden. ProPublican toimittajan Charles Ornsteinin mukaan toimittajat eivät voi enää vain sanoa, että luota meihin, olemme analysoineet datan, vaan toimittajien pitää yhä useammin julkaista myös raakadata yleisölle. (Emt., 16.)

Toinen, mikä on muuttunut, on kerrontatapa. Kun katsoo vaikka kymmenen vuotta taaksepäin, interaktiivisuus ja visualisoinnit olivat harvinaisia, huomauttaa The New York Timesin Interactive News -osaston vetäjä Chase Davis. Nykyään datajournalistit raportoivat ja kirjoittavat, tekevät interaktiivisia elementtejä ja visualisointeja, kehittävät uusia alustoja jutuille, laativat ennusteita, tekevät avoimen lähdekoodin ohjelmistoja ja paljon muuta. Howard huomauttaa, että samaan aikaan tietokoneavusteinen journalismi tuntuu jälleen ajankohtaiselta termiltä, koska Euroopassa tai Yhdysvalloissa on tuskin yhtäkään toimittajaa, joka ei käyttäisi tietokonetta työssään. (Emt., 12–13.)

Howardin raportin perusteella näyttää siltä, että tietokoneavusteinen journalismi, datajournalismi ja journalismi ylipäätään pyrkivät samaan, mutta eri keinoilla. Perinteisessä journalismissa esimerkiksi korruptiosta voitaisiin kertoa yksittäistapauksen kautta. Tietokoneavusteisessa journalismissa mukaan valjastettaisiin tilastot, joita analysoitaisiin ja joista tehtäisiin ehkä jonkinlainen grafiikka.

Datajournalismissa mentäisiin vielä pidemmälle, eli tilastoista tehtäisiin interaktiivinen visualisointi, ja lisäksi julkaistaisiin raakadata. Howard kuitenkin huomauttaa, että yhä useammalle datajournalistille datavisualisoinnit ja interaktiiviset kartat eivät ole itsetarkoituksellisia, vaan välineitä vaikka juuri korruption etsimiseen (emt., 14).

Mark Coddington (2014) on pyrkinyt selventämään käsitteiden datajournalismi, laskennallinen journalismi ja tietokoneavusteinen journalismi eroja. Hän vertailee näitä kolmea journalismin muotoa sen mukaan, miten ne suhteutuvat ammatilliseen suuntautumiseen, avoimuuteen, tiedon luonteeseen ja yleisöön (ks. kuvio 2).



Kuvio 2: A visualized typology of data-driven journalism forms (Mark Coddington 2014, 7).

Coddingtonin mukaan tietokoneavusteinen journalismi on kaikkein eniten sidoksissa ammatillisiin journalistisiin normeihin ja käytäntöihin, kun taas laskennallista journalismia tehdään usein yhteistyönä, ja sen lopputuote voi olla journalistisen jutun sijaan tuote tai alusta. Datajournalismi sijoittuu jonnekin näiden kahden väliin: siinä kyllä tehdään perinteisiä journalistisia juttuja, mutta ei välttämättä perinteisillä keinoilla. (Coddington 2014, 8–10.) Avoimuudessa datajournalismi pitää ykkössijaa. Datajournalismissa pyritään läpinäkyvyyteen esimerkiksi julkaisemalla juttujen lähteenä käytettyä dataa, mikä on osin avoin data -liikkeen perua. Tietokoneavusteisessa journalismissa ei perinteisesti ole pyritty läpinäkyvyyteen, ja laskennallisessa journalismissa läpinäkyvyyden tavoittelu on vaikeaa, koska algoritmien selittäminen yleisölle vaatisi ekstratyötä- ja rahaa – toisin kuin valmiin datan julkaiseminen. (Emt., 10–11.)

Kolmas vertailukohta liittyy tiedon keräämisen ja analysoimisen tapoihin. Yhteiskuntatieteistä vaikutteita saanut tietokoneavusteinen journalismi voi perustua otantaan, kun taas datajournalismissa ja etenkin laskennallisessa journalismissa pyritään hahmottamaan laajoja kokonaisuuksia isojen datamassojen avulla. (Emt., 11–12.) Datajournalismi ja laskennallinen journalismi muistuttavat siinäkin toisiaan, että niissä yleisö nähdään aktiivisena: yleisö kykenee itse analysoimaan dataa. Tietokoneavusteissa journalismissa yleisö käsitetään perinteisemmin passiiviseksi. (Emt., 12–13.)

Coddington toteaa, että datajournalismi pääsee näistä kolmesta journalismin muodosta lähimmäksi ammatillisen journalismin, avoimen datan ja laskennallisuuden (computational) yhdistämistä. Muut journalistit ottavat datajournalistit vakavasti, koska nämä myöntävät kerronnallisuuden ja perinteisen raportoinnin olevan tärkeää. Laskennallisella journalismilla on tiukempi side koodaamiseen ja siten

avoimeen dataan. Tietokoneavusteinen journalismi puolestaan lähti yrityksestä yhdistää yhteiskuntatieteet ammatilliseen journalismiin. (Emt., 14.)

Luen Coddingtonin tutkimustulosta siten, että datajournalismilla olisi parhaat mahdollisuudet yleistyä. Sen suosio onkin noussut monin paikoin. Belgiassa, Norjassa, Ruotsissa ja Yhdysvalloissa tehtiin samoihin aikoihin tutkimusta siitä, miten datajournalismi on kehittynyt toimituksissa (De Maeyer et al. 2014; Karlsen & Stavelin 2014; Appelgren & Nygren 2014; Fink & Anderson 2014). Vuoden 2014 lopussa ilmestyneet tutkimusraportit antavat varsin hyvän kuvan siitä, miten eri tavoin datajournalismi voidaan ymmärtää ja miten eri tavoin sitä tehdään. Norjaa koskevassa tutkimusraportissa puhutaan datajournalismin sijaan laskennallisesta journalismista (Karlsen & Stavelin 2014). Tutkijoiden mukaan se, mitä termiä käytetään, riippuu työprosessin painotuksista (emt., 35).

Kaikissa neljässä tutkimuksessa oli aineistonhankintamenetelmänä puolistrukturoitu haastattelu. Haastateltavia oli 7–23 per tutkimus. Suurin osa haastateltavista työskenteli eri rooleissa toimituksissa, mutta Belgiasta mukana oli myös muita datajournalismiin liittyviä henkilöitä, kuten kouluttajia. Ruotsissa tutkijat tekivät haastattelujen lisäksi lomakekyselyn suuremmalle joukolle toimittajia, ohjelmoijia ja muuta toimituksen väkeä. Missään tutkimuksessa ei ennalta määritelty datajournalismia, vaan pyrittiin selvittämään haastateltavien käsityksiä siitä. (De Maeyer et al. 2014; Karlsen & Stavelin 2014; Appelgren & Nygren 2014; Fink & Anderson 2014.)

Ruotsissa niin toimittajien kuin toimituspäälliköiden käsitykset datajournalismista olivat monenkirjavia. Kun yksi haastateltava näki datajournalismissa paljon mahdollisuuksia, toinen sanoi sitä typeräksi termiksi. Kourallinen kyselylomakkeeseen vastanneista ei nähnyt eroa datajournalismin ja journalismin välillä, ja osa huomautti, että juttuja on tehty aina datan ja tilastojen perusteella. Tyypillisesti tällaisia vastauksia antaneet eivät olleet tehneet paljoa datajournalismia. Ester Appelgren ja Gunnar Nygren huomasivatkin, että kokemuksen määrä korreloi positiiviseen asenteeseen datajournalismia kohtaan ja siihen, viihtyykö numeroiden ja tilastojen parissa. (Appelgren & Nygren 2014, 395–404.)

Myös Belgiassa huomattiin sama asia: vastaukset vaihtelivat datajournalistisen kokemuksen eikä niinkään vastaajan aseman mukaan. Jotkut eivät nähneet eroa datajournalismin ja journalismin välillä. Eräs vastaaja huomautti, että grafiikan tekeminen ei vielä tarkoita datajournalismin tekemistä. Määritelmät siis vaihtelivat huomattavasti, mutta myös yhteisiä nimittäjiä löytyi. Määritelmissä

korostui usein tekniikka, asioiden paljastaminen uusien keinojen avulla, tutkiva journalismi, innovointi ja visuaalisuus. Vastaajien mielestä on tavallista ja perinpohjaista datajournalismia. Toinen vie paljon aikaa ja vaatii taitoja, toinen on helppoa ja nopeaa. (De Maeyer et al. 2014, 8–10.)

Norjassa määritelmät olivat usein laiveita. Eräs vastaaja kuvaili, että laskennallinen journalismi on kaikkea excelin käytöstä sellaisten työkalujen käyttöön, joiden avulla saa tehtyä journalismia, jonka tekeminen olisi muuten mahdotonta. Datan analysoiminen nousi keskeiseksi asiaksi. (Karlsen & Stavelin 2014, 42.)

Yhdysvalloissa tutkimuksen tehneet Katherine Fink ja C. W. Anderson eivät tutkimuspaperissaan avaa, miten haastateltavat olivat määritelleet datajournalismin. Erilaiset määritelmät ovat kuitenkin luettavissa muista asioista. Esimerkiksi isojen mediatalojen edustajat kertoivat käyttävänsä datajournalismin tekemisessä eri koodikieliä, kun taas pienissä lehdissä käytössä ovat Googlen ilmaiset työkalut. Se kertonee, että eri mediataloissa datajournalismi tarkoittaa eri asioita. (Fink & Anderson 2014.)

Samankaltaisia seikkoja nousi esiin myös muiden maiden tutkimuksista. Neljässä maassa tehdyistä tutkimuksista on kaikkiaan luettavissa, että datajournalismi käsitetään monin tavoin, ja sen kehitys on eri paikoissa eri vaiheissa. Työkalut ja työtavat ovat kuitenkin monesti varsin samanlaisia. Usein ongelmaksi koetaan ajan ja taitojen puute. Esimerkiksi Appelgren ja Nygren saivat yhdeksi tutkimuksen tulokseksi sen, että datajournalismi on vielä lapsenkengissä Ruotsissa. Osa suhtautuu siihen innolla, osa skeptisesti. Appelgren ja Nygren näkevät kaksi yhteisöä, joista toinen haluaa kokeilla, leikkiä ja etsiä uusia työtapoja, kun taas toinen haluaa jatkaa tietokoneavusteisen journalismin perinnettä, tehdä datan pohjalta journalismia. Jälkimmäinen ryhmä ei pidä datajournalismista terminä tai ei tunne kaikkia uusia työtapoja ja työkaluja. Yhteinen nimittäjä on kuitenkin se, että datajournalismissa käytetään metodeja, joiden avulla datasta tehdään merkityksellistä journalismia. Yleisimpiä tutkimuksessa esiin nousseita metodeja ovat joukkoistaminen ja karttavisuaalisointien tekeminen. Suurimmiksi esteiksi datajournalismin tekemiselle nähtiin ajan ja taitojen puute. (Emt., 395–404.)

## 3 Aineisto ja menetelmät

### 3.1 Laadullinen tutkimus

Tutkimukseni on laadullinen. Kari Kiviniemen (2010) mukaan laadullista tutkimusta voi luonnehtia prosessiksi. Koska aineistonkeruun väline on tutkija itse, näkökulmat ja tulkinnat kehittyvät tutkijan tietoisuudessa vähitellen tutkimuksen edetessä. Laadullisen tutkimuksen prosessiluonteesta kertoo myös se, että tutkimuksen eri vaiheet eivät välttämättä ole etukäteen jäsennettävissä selkeisiin eri vaiheisiin, vaan tutkimuksen eri elementit kehittyvät joustavasti tutkimuksen edetessä. (Kiviniemi 2010, 70.) Näin kävi myös tässä tutkimuksessa. Tein esimerkiksi alustavaa analyysia jo ennen kuin olin kirjoittanut teorian loppuun, ja palasin sitten uudestaan teorian pariin. Vasta tutkimuksen edetessä minulle hahmottui, kuinka laaja analyysini kuvaileva osa eli katsaus suomalaiseen datajournalismiin on ja miten se sijoittuu muihin tutkimuksen osiin nähden.

Vaikka laadullinen tutkimus on aineistokeskeistä, tutkijan teoreettiset näkökulmat ja vähitellen käsitteellistyvät näkemykset tutkittavasta ilmiöstä suuntaavat osaltaan tutkimuksen kulkua. Aineisto ja teoria ovat vuorovaikutteisia, mutta tutkimuksessa ei ole kyse etukäteen hahmotetun teorian testauksesta. (Emt., 74–76.) Huomasin omassa tutkimuksessani, että aineisto vahvisti monta seikkaa, jotka nousivat esiin jo tutkimuksellista viitekehystä kerätessä. Valitsemani kontekstit osoittautuivat relevanteiksi ja varmasti myös ohjasivat tutkimuksen suuntaa.

Kiviniemen mukaan laadullisen tutkimuksen tekijää kiinnostaa se yksilöllinen merkityksenanto, jota tutkittavat henkilöt ilmiölle antavat (2010, 76). Anneli Sarajärven ja Jouni Tuomen (2009) mukaan “[l]aadullisessa tutkimuksessa ei pyritä tilastollisiin yleistyksiin. Niissä pyritään mm. kuvaamaan jotain ilmiötä tai tapahtumaa, ymmärtämään tiettyä toimintaa, antamaan teoreettisesti mielekäs tulkinta jollekin ilmiölle.” Siksi laadullisessa tutkimuksessa on tärkeää, että tutkittavat tietävät tutkittavasta ilmiöstä mahdollisimman paljon tai heillä on kokemusta asiasta. (2009, 85.) Lähdin keräämään tutkimukseni aineistoa tästä lähtökohdasta ja pyrin välttämään liiallisia yleistyksiä.

### 3.2 Teemahaastattelu

Tutkimusaiheittani voisi lähestyä monelta eri suunnalta. Olisin voinut esimerkiksi analysoida datajournalistisia juttuja ja tutkia, minkälaisen kuvan ne antavat suomalaisesta datajournalismista. Olisin voinut myös mennä toimituksiin havainnoimaan, millaista datajournalismin tekeminen käytännössä on. Minua kuitenkin kiinnostaa erityisesti tekijät datajournalismin takana, ja siksi tein haastattelututkimuksen. Halusin selvittää, millaisia näkemyksiä haastateltavilla on datajournalismista ja mitä merkityksiä he antavat sille, ja tuollaiset asiat selviävät parhaiten kysymällä.

Pidän lähtökohtaa hedelmällisenä, koska datajournalismi on sen verran tuore ilmiö, että sillä ei esimerkiksi ole vakiintunutta, yksiselitteistä määritelmää. Pyrin tutkimuksessani teoretisoimaan datajournalismia haastateltavien sanomisten ja niistä tekemieni tulkintojen perusteella.

Sirkka Hirsjärven ja Helena Hurmeen (2011, 35) mukaan haastattelu sopii hyvin tutkimusmenetelmäksi, jos tutkimuksessa halutaan korostaa tutkittavan merkityksiä luovaa ja aktiivista roolia, tutkimuskohde on vähän kartoitettu, haastateltavan puhe halutaan sijoittaa laajempaan kontekstiin, tutkimuksen aiheen tiedetään jo ennalta tuottavan monitahoisia vastauksia, halutaan selventää vastauksia, halutaan syventää saatavia tietoja ja halutaan tutkia arkoja tai vaikeita asioita. Viimeistä kohtaa lukuun ottamatta kaikki perustelut täyttyvät tutkimuksessani. Datajournalismia on tutkittu vasta vähän, ja siitä on olemassa monenlaisia näkemyksiä, joten tiesin jo ennalta tutkimuksen tuottavan monitahoisia vastauksia. Siten haastateltavien vastaukset vaativat myös selventämistä, syventämistä ja kontekstointia. Esimerkiksi lomaketutkimus ei sovi tutkimusmenetelmäksi tähän tutkimukseen, koska se ei salli asioiden täsmentämistä (emt., 36). Kyselylomakkeet sopivat parhaiten tutkimuksiin, joissa selvitetään melko konkreettisia ja yksiselitteisiä ilmiöitä, ja sellainen datajournalismi ei ole (emt., 37).

Haastattelun haittoiksi Hirsjärvi ja Hurme lukevat seuraavia asioita: haastattelijalta vaaditaan taitoa ja kokemusta, haastattelijan rooliin ja tehtäviin pitäisi kouluttautua, haastattelu vie aikaa, haastattelun katsotaan sisältävän monia virhelähteitä, haastattelusta aiheutuu kustannuksia, ja vapaamuotoisen haastatteluaineiston analysointi, tulkinta ja raportointi on usein ongelmallista valmiiden mallien puuttumisen takia (2011, 35). Omassa tutkimuksessani kustannukset ja aika eivät aiheuttaneet ongelmia. Haastattelun mahdolliset virhelähteet pyrin poistamaan tai ainakin tunnistamaan mahdollisuuksien mukaan. Koska tutkimus on opinnäytetyö, en pidä haastattelijan rooliin ja tehtäviin

kouluttautumista ongelmana, vaan yhtenä tutkimuksen päämääränä. Minulla ei ollut ennalta kokemusta haastattelututkimuksista, mutta perehtyneisyyteni datajournalismiin, aikaisempiin tutkimuksiin ja menetelmäkirjallisuuteen antoivat minulle hyvän pohjan haastattelurungon tekemiseen ja haastattelujen toteuttamiseen. Analysointi, tulkinta ja raportointi osoittautuivat haastaviksi, mutta eivät ylitsepääsemättömiksi. Niissä auttoivat hyvä ohjaaja ja kirjallisuus.

Tarkemmin valitsin tutkimusmenetelmäksi puolistrukturoidun teemahaastattelun. Hirsjärven ja Hurmeen (2011, 48) mukaan teemahaastattelussa olennaista on se, että haastattelu etenee yksityiskohtaisten kysymysten sijaan tiettyjen keskeisten teemojen varassa. Tämä tuo tutkittavien äänen kuuluviin. Teemahaastattelu ottaa huomioon sen, että ihmisten tulkinnat ja asioille antamat merkitykset ovat keskeisiä, sekä sen, että merkitykset syntyvät vuorovaikutuksessa. (Emt., 48.) Sarajärven ja Tuomen mukaan on makukysymys, pitääkö kaikille haastateltaville esittää kaikki suunnitellut kysymykset, pitääkö kysymykset esittää tietyssä järjestyksessä ja pitääkö kysymysten sanamuodot olla kaikissa haastatteluissa samat (2009, 75). Omassa tutkimuksessani pyrin kysymään kaikilta haastateltavilta samat kysymykset, mutta en ollut kovin tarkka kysymysten järjestyksestä ja sanamuodoista.

Teemahaastattelu antaa enemmän vapauksia kuin strukturoitu haastattelu eli lomakehaastattelu. Lomakehaastattelussa kysymysten ja väitteiden muoto ja esittämisjärjestys on täysin määrätty, eli vastausten syventäminen ja selventäminen ei ole mahdollista. (Emt., 44–45.) Sarajärven ja Tuomen mukaan lomakehaastattelu onkin käytännössä useimmiten määrällisen tutkimuksen aineistonkeruumenetelmä (2009, 74). Näistä syistä lomakehaastattelu ei sovi tutkimukseeni.

Teemahaastattelua vapaampi haastattelumuoto puolestaan on strukturoimaton haastattelu eli syvähaastattelu. Syvähaastattelussa käytetään avoimia kysymyksiä, ja haastattelijan pätehtävä on syventää haastateltavien vastauksia ja rakentaa haastattelun jatko niiden varaan. Syvähaastattelut vievät paljon aikaa, ja usein haastateltavia on vain muutama. (Hirsjärvi & Hurme 2011, 44–46.) En valinnut syvähaastattelua tutkimusmenetelmäksi, koska halusin selvittää useamman kuin vain muutaman toimittajan näkemyksiä datajournalismista. Lisäksi yksi teemahaastattelun etu syvähaastatteluun verrattuna on se, että teemat antavat selkeän pohjan aineiston analyysille.

On kuitenkin huomionarvoista, että teemahaastattelussa ei itse asiassa ole rajattu sitä, miten syväälle aiheen käsittelyssä mennään. Hirsjärven ja Hurmeen mukaan teemahaastattelusta tekee

puolistrukturoidun menetelmän se, että yksi haastattelun aspekti – aihepiirit tai teema-alueet – on kaikille sama. (Emt., 47–48.) Sarajärven ja Tuomen mukaan puolistrukturoitu haastattelu on lähellä syvähaastattelua (2009, 75). Myös Jari Eskola ja Jaana Vastamäki (2010) toteavat, että teemahaastattelun ja syvähaastattelun välistä rajaa voi olla vaikea vetää. (Eskola & Vastamäki 2010, 29.) Omassa tutkimuksessani en poikennut paljoa valmiiksi mietityistä kysymyksistä, mutta esitin harkinnan mukaan joitakin lisäkysymyksiä.

### 3.3 Ensisijainen aineisto

Haastattelin tutkimuksessani kahdeksaa toimittajaa, uutispäällikköä tai mediatekniikan harjoittelijaa, jotka ovat osallistuneet datajournalististen juttuprojektien tekemiseen. Rajasin tutkimukseni Yle Uutisissa, Helsingin Sanomissa tai STT-Lehtikuvassa työskenteleviin tai työskennelleihin henkilöihin.

Yle Uutiset ja Helsingin Sanomat ovat itsestään selvästi mukana tutkimuksessa, koska niissä on datajournalismiin tai verkonomaiseen journalismiin keskittyneet yksiköt, ja niissä datajournalismia tehdään säännöllisesti. Suurin osa suomalaisesta datajournalismista tehdään Ylellä tai Helsingin Sanomissa. STT-Lehtikuvassa puolestaan on tehty datajournalismia vasta kesästä 2013 lähtien, eikä sen tekeminen ole yhtä säännöllistä kuin Yle Uutisissa ja Helsingin Sanomissa. Halusin kuitenkin ottaa STT-Lehtikuvan mukaan tutkimukseen, koska se on toteuttanut useita datajournalistisia projekteja ja koska siellä on ollut datajournalismiin keskittyviä mediatekniikan harjoittelijoita. STT-Lehtikuvan tekemä datajournalismi on myös sikäli merkittävää, että sillä on mahdollisuus levitä maakuntalehtien kautta hyvin laajalle yleisölle. Muissa Suomen mediataloissa on tehty vasta hajanaisia datajournalistisia kokeiluja.

Olisin voinut ottaa mukaan tutkimukseeni myös tunnetun suomenruotsalaisen datajournalistin Jens Finnäsin, joka on tehnyt Suomessa töitä muun muassa Ylelle, Hufvudstadsbladetille ja Journalistille (Valkama 2012). Rajasin hänet kuitenkin tutkimukseni ulkopuolelle, koska hän työskentelee nykyään lähinnä Ruotsissa eikä siten vaikuta niinkään suomalaiseen datajournalismikenttään. Rajasin tutkimukseni ulkopuolelle myös Svenska Ylen, koska haluan tutkimuksessani keskittyä suomenkieliseen datajournalismiin. Kerron kuitenkin hieman Svenska Ylen datajournalismista luvussa neljä, sillä toimitus tekee melko paljon datajournalismia.



Seuraavaksi esittelen haastateltavani lyhyesti (haastattelujärjestyksessä). Esittelyt kuvaavat tilannetta maaliskuussa 2014, jolloin tein haastattelut.

1) Esa Koivuranta, 28, tekee tiedotusopin maisterin tutkintoa Tampereen yliopistossa ja on käynyt TaY:n ja TTY:n yhteisen datajournalismikurssin keväällä 2012. Hän on työskennellyt tarvittaessa töihin tulevana toimittajana STT-Lehtikuvassa tammikuusta 2013 lähtien. Datajournalismia hän teki STT-Lehtikuvassa kesällä 2013, kun oli kotimaan toimituksessa kesätöissä. Aiemmin hän on työskennellyt useissa sanomalehdissä, kuten Savon Sanomissa. Siellä hän muun muassa selvitti kansanedustajien läsnäoloa valiokunnissa ja eduskunnassa, mutta ei osaa sanoa, olivatko jutut datajournalismia.

2) Wolf Wikgren, 24, opiskelee mediatekniikan insinööriksi ammattikorkeakoulu Metropoliasissa. Hän suoritti kesällä 2013 opintoihinsa kuuluvan harjoittelun STT-Lehtikuvassa. Omien sanojensa mukaan hän oli "käytännössä datajournalistinen kesätoimittaja", joka hoiti datajournalististen juttujen teknisen toteutuksen ja pyöritteli exceliä. Wikgren on opiskellut myös tietojenkäsittelytiedettä.

3) Teemo Tebest, 29, on työskennellyt helmikuusta 2013 asti toimittajana Yle Uutisten Plus-deskissä. Hän käyttää itsestään leikkimielisesti nimitystä atk-toimittaja, sillä hänen toimenkuvaansa kuuluu enemmän koodaamista kuin perinteisiä toimittajan töitä. Aikaisemmin Tebest on työskennellyt web designerina Svenska Ylessä ja tutkijana Tampereen teknillisessä yliopistossa. Hän valmistui tietotekniikan diplomi-insinööriksi vuonna 2010 pääaineenaan hypermedia.

4) Tuomo Björksten, 31, aloitti työt Yle Uutisissa kesäkuussa 2013. Pääsääntöisesti hän työskentelee tv-uutisten toimittajana kotimaan toimituksessa. Ennen Yleä hän työskenteli lähes kuusi vuotta Aamulehdessä. Takana on töitä myös ohjelmoijana ja mainosalalla sekä freelancerina työskentelyä aikakauslehdille. Björksten on opiskellut viestintää ja puheviestintää Jyväskylän yliopistossa, mutta opinnot ovat kesken. Datajournalismiin hän on tutustunut oma-aloitteisesti.

5) Johan Jaakkola, 30, on työskennellyt Yle Uutisissa elokuusta 2012 lähtien. Tällä hetkellä hän työskentelee taloustoimituksessa. Ennen Yleä Jaakkola on tehnyt toimittajan töitä STT-Lehtikuvassa. Hän valmistui medianomiksi vuonna 2008 Oulun ammattikorkeakoulusta ja aloitti samana vuonna maisteriopinnot Jyväskylän yliopiston viestintälinjalla, mutta ei ole tehnyt opintoja loppuun. Varsinaista datajournalismikoulutusta hänellä ei ole, mutta hän on käynyt aiheeseen liittyvillä luennoilla esimerkiksi tutkivan journalismin seminaarissa.

6) Esa Mäkinen, 33, on Helsingin Sanomien datajournalismista vastaava uutispäällikkö. Hän aloitti tehtävässä heinäkuussa 2013, mutta on tehnyt käytännössä samaa työtä tuottajan nimikkeellä jo kesäkuusta 2012 lähtien. Mäkinen aloitti työt Helsingin Sanomien kulttuuritoimittajana vuonna 2006. Sitä ennen hän työskenteli Ylioppilaslehden päätoimittajana ja freelance-toimittajana. Mäkisellä ei ole datajournalismi- eikä journalismikoulutusta, vaan hän on valmistunut valtiotieteen maisteriksi Helsingin yliopistosta pääaineenaan sosiologia.

7) Jyri Hänninen, 36, siirtyi helmikuussa 2014 Helsingin Sanomien kaupunkitoimitukseen, mutta datajournalistisia projekteja hän on tehnyt pääosin taloustoimituksessa. Hän on työskennellyt Helsingin Sanomissa eri osastoilla useita vuosia. Yhteensä hän on tehnyt 12 vuotta toimittajan töitä Helsingin Sanomissa, Etelä-Saimaassa, freelancerina ja kaupunkilehti Lappeenrantalaisessa, mutta on välillä ollut muun alan töissäkin. Hänninen on käynyt Sanoman toimittajakoulun. Varsinaista datajournalismikoulutusta hänellä ei ole.

8) Tuomas Peltomäki, 32, on tehnyt toimittajan ja graafikon töitä ja harrastaa koodaamista. Hän on työskennellyt hs.fin toimitussihteerinä heinäkuusta 2013 lähtien. Sitä ennen hän on työskennellyt eri toimituksissa Helsingin Sanomissa, STT-Lehtikuvassa ja Ylellä. Peltomäki on opiskellut tiedotusoppia, mutta maisteriopinnot ovat kesken. Datajournalistina hän on itseoppinut.

Yle Uutisissa työskentelevä Teemo Tebest ja Helsingin Sanomien Esa Mäkinen ovat Suomen tunnetuimpia datajournalisteja, joten halusin heidät mukaan tutkimukseen. Tiesin myös entuudestaan, että Wolf Wikgren teki kesällä 2013 STT-Lehtikuvan ensimmäisiä datajournalismikokeiluja. Myöhemmin STT-Lehtikuvassa on ollut muita mediatekniikan harjoittelijoita, mutta Wikgren loi STT-Lehtikuvan datajournalismin tekemiselle pohjan esimerkiksi kehittämällä ensimmäiset työkalut, joten valitsin hänet haastateltavakseni.

Tebest, Mäkinen ja Wikgren ovat tutkimukseni avainhenkilöitä, sillä he ovat teknisiä osaajia ja hoitavat datajournalististen juttujen koodaamisen. Lisäksi halusin haastatella toimittajia, jotka ovat osallistuneet datajournalistisiin projekteihin ja niiden ideoimiseen, vaikka he eivät välttämättä olisikaan identifioituneet kovin vahvasti datajournalisteiksi. Näiden toimittajien etsimiseen käytin niin sanottua lumipallo-otantaa, eli pyysin valitsemiani avainhenkilöitä ehdottamaan yhtä tai kahta haastateltavaa, joiden kanssa he ovat tehneet eniten datajournalismia toimituksissa (Hirsjärvi & Hurme 2011, 59–60). Siten löysin loput haastateltavat. Haastatteluvaiheessa selvisi, että myös Tuomas

Peltomäki osaa toimittajan töiden lisäksi koodaata ja jossain määrin myös tehdä grafiikkaa. Lisäksi Tuomo Björksten on työskennellyt ohjelmoijana.

Rajasin tutkimukseni ulkopuolelle muut datajournalismin tekemiseen osallistuvat henkilöt, kuten uutispäälliköt, tuottajat ja graafikot, koska halusin haastatella juttuprojektien ydintyöntekijöitä. Poikkeuksena on tuottaja Esa Mäkinen, sillä nimikkeestään huolimatta hän tekee datajournalististen juttujen perustyötä, kuten koodaa ja joskus myös kirjoittaa.

Tein teemahaastattelut maaliskuussa 2014. Kaksi ensimmäistä haastattelua olivat harjoitushaastatteluja, joiden avulla testasin teemahaastattelurunkoa. Niiden perusteella lisäsin pari kysymystä runkoon ja muutin parin kysymyksen paikkaa. Suuria muutoksia en kuitenkaan tehnyt, ja runko pysyi perustaltaan samana. Siksi käytän ensimmäisiä haastatteluja normaalisti aineistona.

Kysyin kaikilta haastateltavilta, voinko käyttää heidän lausuntojaan tutkimuksessa omilla nimillä. Kaikki haastateltavat antoivat tähän luvan. Tein kaikki haastattelut kasvotusten, paitsi yksi haastattelu jäi kesken, koska haastateltavan piti lähteä juttukeikalle. Tein haastattelun loppuun puhelimitse. Nauhoitin kaikki haastattelut, ja haastattelunauhaa kertyi yli seitsemän tuntia. Litteroin kaikki haastattelut. Haastattelujen tekemisessä ei ollut teknisiä ongelmia. Kaksi ensimmäistä haastattelua tein Fazerin kahvilassa, joka oli niin hälyisä, että se vaikeutti hieman litterointia. Loput kuusi haastattelua tein rauhallisemmissa paikoissa, lähinnä haastateltavien työpaikkojen kahviloissa.

Teemahaastattelujen runko on tutkimuksen liitteenä. Haastattelujen neljä isoa teemaa ovat 1) haastateltavan omat kokemukset datajournalismista, 2) yhden datajournalistisen jutun synty, 3) datajournalismin määritelmä ja 4) datajournalismin merkitys. Ennen haastatteluja pyysin haastateltavia miettimään valmiiksi, mikä on paras datajournalistinen juttu, jota haastateltava on ollut tekemässä, ja linkkaamaan minulle jutun. Juttu toimi pohjana haastattelun toisen teeman kohdalla eli datajournalistisen jutun synnyn kuvaamisessa.

Analyysini perustuu litteroinnin avulla tuotettuun tekstiaineistoon, vaikka litteraatio onkin tulkinta analysoitavasta tilanteesta. Tällainen käytäntö on yleinen laadullisessa haastattelututkimuksessa, koska isojen kokonaisuuksien hahmottaminen nauhoitettua puhetta kuuntelemalla olisi vaikeaa tai jopa mahdotonta. (Ruusu vuori 2010, 427–428.) Aineiston litteroinnin sopiva tarkkuus riippuu tutkimusongelmasta ja -otteesta (Hirsjärvi & Hurme 2011, 139; Ruusu vuori 2010, 424). Koska en tutki vuorovaikutusta, vaan tutkimuksessani tärkeää on haastattelussa esiin tulevat asiasisällöt, en litteroinut

aineistoa ihan sanatarkasti (Ruusu vuori 2010, 424–423). Tällä tarkoitan, etten kirjannut jokaista “niinku”- tai “tota”-sanaa. En myöskään kirjannut puheen taukoja tai esimerkiksi naurahduksia. Sen sijaan pyrin täsmällisyyteen asiasisällön ja sanavalintojen osalta. Koska yksi tutkimukseni tavoite on määritellä datajournalismi, on oleellista huomioida, millä sanoilla haastateltavat siitä puhuvat.

On huomionarvoista, että litterointi ei ole vain analyysin perusta, vaan sitä voi pitää osana analyysiprosessia. Tutkijan teoreettinen esiyymmärrys ja oletukset tutkittavasta ilmiöstä heijastuvat siihen, millaisia ratkaisuja tutkija tekee aineistoa purkaessaan. Nuo ratkaisut puolestaan vaikuttavat siihen, millaisen kertomuksen tutkija rakentaa haastateltavista. (Nikander 2010, 432–435.) Siksi olen pyrkinyt tekemään läpinäkyväksi ne periaatteet, joiden mukaan olen litteroinut haastattelunauhat.

### **3.4 Tausta-aineisto**

Tutkimukseni tausta-aineisto koostuu ensisijaisesti vapaamuotoisista puhelinhaastatteluista. Haastattelin kolmentoista toimituksen edustajia, jotka ovat työssään tekemisissä datajournalismin kanssa. Puhelut on käyty lokakuussa 2014. Haastattelujen tarkoitus oli selvittää, millainen rooli datajournalismilla on suomalaisissa toimituksissa: ketkä sitä tekevät, kuinka paljon sitä tehdään, millaisia työskentelytapoja toimituksissa on ja niin edelleen. Lisäksi etsin muutamia lehtiartikkeleita, joissa kerrotaan datajournalismista suomalaisissa toimituksissa. Lehtiartikkeleiden etsiminen ei ollut järin järjestelmällistä, vaan tein muutamia verkkohakuja. Hakusanoina oli esimerkiksi “datajournalismi Yle” tai “datajournalismi MTV”.

Tausta-aineistoa kerätessäni haastattelin seuraavia henkilöitä: Helsingin Sanomien datadeskin uutispäällikkö Esa Mäkinen, Yle Uutisten Plus-deskin tuottaja Juho Salminen, Svenska Yle Uutisten toimittaja Peter Sjöholm, STT-Lehtikuvan tuottaja Harri Juutilainen, Keski-suomalaisen verkkoliiketoiminnan päällikkö Ville Grahn, Aamulehden verkkopalvelujen kehityspäällikkö Seppo Roth ja toimittaja Taneli Koponen, MTV:n tuottaja Hanna Takala, Taloussanomien uutispäällikkö Aleksis Moisio, Kauppalahden toimituspäällikkö Jari Saario, Talouselämän toimituspäällikkö Olli Ainola, Iltalehden verkon esimies Timo Melari, Iltä-Sanomien sosiaalisen median tuottaja Matti Markkola ja Hufvudstadsbladetin vastaava päätoimittaja Lena Skogberg.

Nämä henkilöt valikoituivat mukaan titteleidensä, oman harkintani ja vinkkien perusteella. Osaan

toimituksista soitin melko sattumanvaraisesti ja kysyin, kuka voisi olla oikea henkilö kertomaan minulle heidän tekemästään datajournalismista. Halusin käydä läpi kattavasti suomalaisen mediakentän, mutta rajasin ulkopuolelle paikallislehdet, koska ne eivät ole nousseet missään yhteydessä esiin datajournalismiin perehtyessäni. En myöskään katsonut tarpeelliseksi käydä läpi jokaista maakuntalehteä, koska asiaan tutustuessani on käynyt selväksi, että mikään niistä ei tee säännöllisesti datajournalismia. Valitsin mukaan Keski-suomalaisen ja Aamulehden sen perusteella, että STT-Lehtikuvan tuottaja Harri Juutilainen mainitsi niiden olevan jokseenkin aktiivisia datajournalismin suhteen. Moni muukin maakuntalehti on varmasti hyödyntänyt datajournalismia jossakin määrin. Aikakauslehdistä kelpuutin mukaan vain talousaiheisen Talouselämän ja maininnan oloisesti Suomen Kuvalehden. Muut aikakauslehdet eivät nousseet millään tavalla esiin asiaa kaivellessani.

Kirjoitin puheluista muistiinpanot, joiden pohjalta tein kuvailevan analyysin. Suurimman osan puheluista myös nauhoitin varmuuden vuoksi, mutta en katsonut tarpeelliseksi palata nauhoihin aineistoa kuvaillessani.

### **3.5 Analyysin toteuttaminen**

Valitsin analyysimenetelmäksi sisällönanalyysin. Tämä koskee ensisijaista aineistoa, jota käsitteletin analyttisesti. Sitä ennen kirjoitin tausta-aineiston pohjalta analyysin kuvailevan osion, minkä tekemiseen en tarvinnut erillistä analyysimenetelmää.

Sarajärven ja Tuomen (2009) mukaan sisällönanalyysi on perusanalyysimenetelmä, jota voidaan käyttää kaikissa laadullisen tutkimuksen perinteissä. Se voi olla joko yksittäinen menetelmä tai väljä teoreettinen kehys, ja tutkimuksessani se on molempia. Väljänä teoreettisena kehyksenä sillä tarkoitetaan kirjoitettujen, kuultujen tai nähtyjen sisältöjen analyysiä. (Sarajärvi & Tuomi 2009, 91.)

Sisällönanalyysin voi jakaa aineistolähtöiseen, teoriaohjaavaan ja teorialähtöiseen analyysiin (Emt., 95). Tutkimuksessani on piirteitä sekä aineistolähtöisestä että teoriaohjaavasta analyysistä: tulokset perustuvat pitkälti aineistosta nouseviin havaintoihin, mutta teoria ohjasi tekemistä. Luvussa 6 peilaan tuloksia aikaisempaan teoriaan.

Aineistolähtöisen analyysin tavoite on luoda tutkimusaineistosta teoreettinen kokonaisuus.

Tutkimuksessani lähimpänä tätä on datajournalismin määrittelemisen sen pohjalta, mitä haastateltavat puhuvat datajournalismista. Analyysiyksiköt valitaan tutkimuksesta käsin, eivätkä ne ole etukäteen harkittuja. Analyysi siis lähtee aineistosta, ei aikaisemmasta teoriasta. (Emt., 95–96.)

Myös teoriaohjaavassa analyysissä analyysiyksiköt valitaan aineistosta. Teoreettisia käsitteitä sen sijaan ei luoda aineistosta, vaan ne ovat valmiina olemassa. Näin on myös tutkimuksessani, eli en keksi käsitteitä itse. Analyysi etenee aineiston ehdoilla, mutta aikaisempi tieto ohjaa tai auttaa analyysiä. Teoriaohjaava analyysi ei ole aikaisempaa teoriaa testaava, vaan aukoo uusia ajatusuria. (Emt., 96–97, 117.)

Teorialähtöisessä analyysissä puolestaan testataan yleensä aikaisempaa tietoa uudessa kontekstissa. Analyysi nojaa johonkin teoriaan, malliin tai auktoriteetin esittämään ajatteluun, jonka pohjalta määritellään muun muassa tutkimuksessa kiinnostavat käsitteet. Käytännössä teorialähtöisen ja teoriaohjaavan analyysin lopputulos voi olla sama. (Emt., 97–98.) Datajournalismia on tutkittu sen verran vähän, että teorialähtöisen analyysin valitseminen olisi tuntunut vaikealta.

Aineistolähtöisen analyysin heikkous on siinä, että esimerkiksi tutkijan käyttämät käsitteet, tutkimusasetelma ja menetelmät ovat tutkijan valitsemia ja vaikuttavat tuloksiin. On siis vaara, että analyysi tapahtuu tutkijan ennakkoluulojen saattamana eikä haastateltavien ehdoilla. Ongelmaa on pyritty ratkaisemaan siten, että tutkija kirjoittaa auki ennakkokäsityksensä tutkittavasta ilmiöstä ja tiedostaa ne analyysin aikana. (Emt., 96.) Tähän pyrin myös omassa tutkimuksessani. Koska tutkimuksessani on myös teoriaohjaavan analyysin piirteitä, ongelma ei ole niin iso, vaan teoria kulkee jollain tavalla mukana läpi analyysin ja ohjaa valintojen tekemisessä.

Aloitin varsinaisen analyysin tekemisen käymällä läpi jokaisen haastattelun litteraatit. Merkitsin oleelliset kohdat värikoodauksella, joka mukaili haastattelun teemoja ja kysymysrunkoa. Esimerkiksi datajournalismin määritelmä -teemaa koskevat oleelliset kohdat merkitsin sinisen eri sävyillä sen mukaan, mihin kysymykseen ne vastasivat: esimerkit hyvästä datajournalismista merkitsin vaaleansinisellä, esimerkit huonosta datajournalismista hieman tummemmalla sinisellä ja niin edelleen. Joissain kohdin en noudattanut täysin kysymysrunkoa, vaan yhdistelin kysymyksiä isommiksi kokonaisuuksiksi tai jaoin niitä pienempiin osiin. Kysymyksiä tuli tällä tavalla yhteensä 29.

Kun olin käynyt läpi kaikki kahdeksan haastattelun litteraattia, kopioin ja liitin värikoodatut osiot erillisiin tiedostoihin. Se helpotti aineiston käsittelyä jatkossa. Loin myös tiedoston, johon kirjoitin

ranskalaisilla viivoilla yleisiä huomioita litteraateista. Tiedosto toimi myöhemmin ikään kuin muistilistana siitä, mitä asioita analyysiä kirjoittaessa ainakin tulee käsitellä.

Tämän jälkeen loin neljä taulukkoa haastatteluteemojen mukaan. Teemat ovat 1) omat kokemukset datajournalismista, 2) datajournalistisen jutun synty, 3) datajournalismin määritelmä ja 4) datajournalismin merkitys. Taulukoissa on vaakarivillä teemaan liittyvät kysymykset ja pystyrivillä kahdeksan haastateltavaa. Tiivistin taulukkoon haastateltavien vastaukset muutamaaan sanaan tai lauseeseen. Tein vielä toiset taulukot, joissa tiivistin vastaukset maksimissaan kymmenen sanan mittaisiksi. Tässä vaiheessa etsin jo vastauksista yhtäläisyyksiä ja eroavaisuuksia. Jos useampi haastateltava puhui eri sanoin samasta asiasta, pyrin toisinaan tiivistämään vastaukset samalla termillä. Jos yhteistä termiä ei löytynyt luontevasti, en väkisin keksinyt sellaista. Taulukot eivät siis ole tarkkarajaista luokittelua, vaan ennemmin tiettyjen asiakokonaisuuksien piirteiden kuvailua. Koin tällaisen työskentelytavan järkeväksi, koska haastateltavien ääni pääsi tällä tavalla hyvin kuuluviin.

En käyttänyt analyysissä puhtaasti mitään yksittäistä analyysitapaa. Hirsjärven ja Hurmeen mukaan analyysitapoja ovat laskeminen, asteikointi, teemoittelu, alaryhmien analysointi tai yhteyksien tarkastelu, kehityskulun tarkastelu ja metaforien käyttö (2011, 153). Näistä lähimmäksi omaa analyysiäni tulevat laskeminen ja teemoittelu, jotka kuvaavat erityisesti tapaa, jolla käsitteitin tekemiäni taulukoita.

Laskeminen tarkoittaa sen seikan selvittämistä, kuinka monta kertaa jokin ilmiö esiintyy aineistossa (emt., 172). Tein analyysissä paljon laskutoimituksia siitä, kuinka moni haastateltava mainitsi jonkin tietyn piirteen, vaikkapa uutisellisuuden hyvän datajournalismin piirteenä. Mitä enemmän piirre sai mainintoja, sitä selvemmin se määrittää asiakokonaisuutta, vaikkapa hyvää datajournalismia. Teemoittelu puolestaan tarkoittaa sellaisten aineistosta nousevien piirteiden tarkastelua, jotka ovat yhteisiä usealle haastateltavalle (emt., 173). Toisin sanoen se on askel laskemisesta eteenpäin. Taulukoista löytyi laskemalla ne asiakokonaisuutta määrittävät piirteet, joita käsitteitin kirjallisessa analyysissä tarkemmin. Teemoittelua ei pidä tässä sotkea isoihin haastatteluteemoihin.

Lopuksi kirjoitin taulukoiden ja litteraattien pohjalta tekemiäni tärkeimmät havainnot auki. Otin analyysin sekaan sitaatteja haastatteluista ja kaksi osataulukkoa, joissa on tiivistettynä haastateltavien ajatuksia.

## 4 Suomalaisen datajournalismin tila vuonna 2014

### 4.1 Datajournalismi Suomessa

Tutkivat toimittajat ovat käyttäneet Suomessa datajournalismin työtapoja jo paljon ennen kuin on alettu puhua datajournalismista. Esimerkiksi MTV:n tutkivan journalismin ryhmän tuottaja Hanna Takala kertoo hyödyntäneensä dataa journalismissa 1990-luvun lopulta lähtien. Samoihin aikoihin on alettu kehittää verkkokerrontaa. Helsingin Sanomat perusti 1990-luvun lopussa Verkkoliite-nimisen toimituksen, jossa tehtiin interaktiivista journalismia.

Datajournalismista on kuitenkin alettu puhua vasta 2010-luvun alussa. Se näyttää saaneen jalansijaa toimituksissa usein yksittäisten ihmisten innostuksen ansiosta. Esimerkiksi Helsingin Sanomien datadeskin tuottaja Esa Mäkinen kertoo, että sai kulttuuritoimittajana työskennellessään esimiehiltä luvan kehittää omia juttuprojekteja, jotka perustuivat dataan. Innoittajina toimivat ulkomaiset mediat, kuten The Guardianin Datablog. Työpariksi löytyi HS.fi:n kehitysryhmän graafikko Jarmo Lundgren, jonka kanssa Mäkinen sittemmin perusti datadeskin. Lokakuussa 2014 deskissä aloitti jo kuudes työntekijä.

Suomen Kuvalehdessä kehitys on mennyt toiseen suuntaan. Vuosina 2011 ja 2012 lehti teki ensimmäisten joukossa datan pohjalta verkkoon visualisointeja. Tekijöinä olivat Lauri Vanhala ja Juho Salminen. Sitten tekijät lähtivät muualle töihin, eikä verkkovisualisointeja ole enää näkynyt. Yle Uutisten tuottajana työskentelevä Salminen ei osannut vuoden 2014 loppupuolella arvioida, kuinka tietoisesti Suomen Kuvalehdessä päätettiin, että asiaan ei panosteta.

Tässä analyysin osassa kartoitan, mitä Suomessa on tehty datajournalismin saralla ja kenen toimesta lokakuuhun 2014 mennessä. Vastaan siis kolmanteen tutkimuskysymykseen (ks. luku 1.2). Osio on ennen kaikkea kuvaileva. Se perustuu tausta-aineistoon eli erityisesti puhelinkeskusteluihin, joita olen käynyt niiden toimitusten edustajien kanssa, jotka vastaavat datajournalismista toimituksissa. Tein puhelinhaastattelut lokakuussa 2014. Lisäksi käytän lähteenä joitakin asiasta kirjoitettuja lehtiartikkeleita ja tekemiäni teemahaastatteluja. Jos lähdeä ei ole mainittu, teksti perustuu tekemiini teema- ja taustahaastatteluihin.

Kuten yllä kertomani esimerkit osoittavat, datajournalismin määritelmästä ja merkityksestä on



Suomessa monia mielipiteitä. Se hankaloitti kartoituksen tekemistä. Kun kysyin toimitusten panoksesta datajournalismiin, sain usein vastaukseksi, että riippuu mitä sillä tarkoitetaan. Niinpä tässä osiossa korostuu se, miten toimitusten edustajat itse käsittävät datajournalismin. Heidän näkemyksensä eivät välttämättä kerro edes edustamiensa medioiden linjasta.

Pidän katsausta suomalaiseseen datajournalismiin tarpeellisena, koska se asettaa tekemäni teemahaastattelut kontekstiinsa ja pohjustaa tutkimukseni varsinaista analyysiä. Siksi lähden liikkeelle tausta-aineistosta ja pureudun vasta seuraavassa luvussa ensisijaiseen aineistoon. Katsauksella on myös itsenäinen arvonsa. Katsaus kertoo osaltaan siitä, minkälainen datajournalismin asema suomalaisessa mediakentässä on. Tietääkseni vastaavaa katsausta ei tässä laajuudessa ole aiemmin tehty.

## 4.2 Helsingin Sanomat

Helsingin Sanomien ensimmäinen varsinainen datajournalistinen juttu oli kirjakustantamoiden laatuvertailu. Esa Mäkisen ideoima juttu julkaistiin vuoden 2011 alussa. Toukokuussa 2012 Helsingin Sanomat ilmoitti perustavansa datajournalismin ryhmän. Tuottajaksi nimettiin Esa Mäkinen, ohjelmoija-graafikoksi Jarmo Lundgren. Helsingin Sanomien uutisen (2012) mukaan ryhmän tarkoitus on tuottaa aiempaa parempaa verkkojournalismia, vuorovaikutteisia grafiikoita ja joukkoistettuja artikkeleita.

Kesäkuussa 2013 Helsingin Sanomat vakinaisti aiemmin kokeiluna toimineen datadeskinsä, ja Esa Mäkisen titteli vaihtui uutispäälliköksi. Tuolloin Mäkinen kertoi Helsingin Sanomien uutisessa, että deskillä on kaksi päätehtävää: toimittajien auttaminen tiedon hankkimisessa, puhdistamisessa ja analysoinnissa sekä erilaisten interaktiivisten visualisointien tekeminen verkkosivuille. Esimerkkeinä interaktiivisista visualisoinneista Mäkinen mainitsee kartat, kyselyt, aikajanat ja laskurit. (Huuskonen 2013.)

Lokakuussa 2014 ryhmään kuului kahden alkuperäisjäsenen lisäksi toimitussihteeri Tuomas Peltomäki, toimittaja Jyri Hänninen sekä entisten Nelosen uutisten graafikot Uolevi Holmberg ja Boris Stefanov. Peltomäki ja Hänninen on nimitetty datadeskin työntekijöiksi sen jälkeen, kun haastattelin heitä keväällä 2014.

Mäkinen koordinoi deskin toimintaa ja yhteistyötä muiden osastojen kanssa, mutta myös hankkii ja käsittelee dataa sekä koodaa. Peltomäki ideoi ja suunnittelee juttuja sekä hallitsee datajournalistisen jutun kaikki tekovaiheet. Jyri Hänninen on toimittaja, joka tekee juttuja kaupunkitoimituksen aiheista. Hän ei osaa koodata, mutta pystyy käsittelemään dataa. Lundgren keskittyy usein isojen juttuprojektien ohjelmointiin ja ulkoasuun. Holmberg ja Stefanov tekevät etupäässä videografiikoita ja yksinkertaisempia visualisointeja, joita varten Helsingin Sanomat on kehittänyt oman Arkku-työkalun. Myös talon muita graafikoita on koulutettu käyttämään Arkkua, ja grafiikkatoimitus tekeekin osin samoja töitä kuin datadeski. Grafiikkatoimituksen Minna-Liisa Linjala oli aiemmin töissä datadeskissä.

Helsingin Sanomat laskee datajournalismiksi kaiken, mikä on interaktiivista tai missä on dataa taustalla. Käytännössä se tarkoittaa, että Helsingin Sanomat merkitsee datajournalismi-aihetunnisteella kaikki verkossa julkaistavat jutut, joiden tekemisessä datadeski on ollut mukana. Kaikki tuolla tunnisteella merkityt jutut kerätään samalle verkkosivulle osoitteeseen [hs.fi/aihe/datajournalismi](https://hs.fi/aihe/datajournalismi) (ks. kuvio 3). Datajournalismi-tunnisteen alla voi olla isoihin datoihin perustuvien juttujen ohella esimerkiksi erikoistaitot, testit ja artikkelit, joissa on vaikkapa interaktiivinen kartta tai aikajana. Tällaista sisältöä syntyy lähes päivittäin.

The screenshot shows the Helsingin Sanomat website interface. At the top, there's a search bar and navigation links like 'OSASTOT', 'UUTISET', 'HSTV', 'SÄÄ', 'PÄIVÄN LEHTI', and 'HS ILTA'. Below the navigation, there's a section for 'DATAJOURNALISMI' with a 'Suosittele' button and a count of 16. The main content area is a grid of articles and advertisements. The first article is 'Valitse ehdokas alle minuutissa HS:n uudenlaisella pikavaalikoneella'. The second is 'Testaa, minkälainen hankala tyyppi olet'. The third is 'Ukraina: Hyökkäykset maan itäosassa rauhoittuneet'. The fourth is 'Laske HS:n testillä, paljonko työmatkasi kuluttaa viikossa kaloreita'. To the right of the grid is a large yellow advertisement for 'MOTIVUS' with the text 'Jumpat alkaen 49€/kk' and 'EI VUOSI-SITOUTUMISTA! LUE LISÄÄ >>'. The bottom of the grid shows a partial article about 'Leikataanko palveluita vai korotetaanko veroja? Pitäisikö hakea Natoon? HS:n vaalikone auttaa löytämään ehdokkaan, joka'.

Kuvio 3: Ruutukaappaus osoitteesta [hs.fi/aihe/datajournalismi](https://hs.fi/aihe/datajournalismi) (1.4.2015).

## 4.3 Yleisradio

Yleisradio tekee aktiivisesti datajournalismia niin suomeksi kuin ruotsiksi. Ohjelmointitaitoinen Teemo Tebest aloitti vuoden 2012 alussa Svenska Ylen web designerina ja alkoi muiden töidensä ohessa tehdä yhteisiä projekteja toimittajien kanssa. Tammikuussa 2013 Tebest vaihtoi toimittajaksi suomenkieliselle puolelle, jonne perustettiin Plus-deski. Sen tarkoitus on auttaa uutis-, urheilu- ja ajankohtaistoimituksia verkonomaisten sisältöjen tekemisessä. Se ei siis keskity pelkästään datajournalismiin, vaan yleisesti verkonomaiseen kerrontaan. Vuonna 2013 Plus-deski osallistui kaikkiaan 350 jutun tekemiseen, ja Tebestin arvion mukaan niistä noin 50 oli datajournalismia. Tebest käsittää datajournalismin hyvin suppeasti verrattuna esimerkiksi Helsingin Sanomien määritelmään. Plus-deskin tekemät jutut kerätään verkkosivulle yle.fi/uutiset/plus (ks. kuvio 4).

Kuvio 4: Ruutukaappaus osoitteesta yle.fi/uutiset/plus (1.4.2015).

Plus-deskin tuottajan Juho Salmisen mukaan aloitteen deskin perustamisesta teki uutis- ja ajankohtaistoiminnan internet-päällikkö Mika Rahkonen. Tätä kirjoittaessa eli vuoden 2014 loppupuolella Yleisradiossa on menossa organisaatiouudistus ja yt-neuvottelut, joten Plus-deskin tulevaisuus on epäselvä. Uudistuksen yksi perimmäinen tavoite on kuitenkin verkon roolin kasvattaminen (Yleisradio 2014). Salmisen mukaan Plus-deskin kaltaista tekemistä halutaan lisätä, ja

sen halutaan olevan tiukemmin yhteydessä päivittäisen uutistuotannon kanssa.

Lokakuussa 2014 Plus-deskissä oli kuusi työntekijää: kaksi tuottajaa, kaksi graafikkoa ja kaksi koodaavaa toimittajaa. Heistä viisi on ollut mukana alusta alkaen. Lisäksi toimittajataustainen Ville Juutilainen palkattiin deskiin elokuussa 2014. Hän ja Teemo Tebest pääasiassa koodaavat, vaikka virallisesti he ovat toimittajia. Käytännössä Tebest tekee töitä ennen kaikkea datan parissa, kun taas Juutilainen tekee enemmän erikoistaittoja ja muita verkonomaisia sisältöjä. Tuottajat Juho Salminen ja Mika Pippuri ovat jakaneet työnsä siten, että Salminen vastaa deskin toiminnasta ja toimii yhteyshenkilönä uutistoimitukseen, kun taas Pippuri hoitaa yhteyksiä ajankohtaistoimitukseen ja toimii sen sparraajana verkkoasioissa. Graafikot ovat Juha Rissanen ja Stina Tuominen. Rissasen vastuulla on erityisesti käyttöliittymien suunnitteleminen ja datagrafiikkojen tekeminen. Tuominen puolestaan on erikoistunut kuvitusten ja infografiikkojen tekemiseen.

Myös Svenska Ylessä on panostettu datajournalismiin edelleen sen jälkeen, kun Tebest vaihtoi töihin suomenkieliselle puolelle. Alexander Ugglan aloitti syksyllä 2012 Svenska Yle Uutisten datajournalismista vastaavana toimittajana. Vuoden 2013 alkaessa hänen sijaisekseen tuli Peter Sjöholm, kun hän lähti virkavapaalle opettajaksi ja tutkijaksi. Elokuussa 2013 perustettiin kolmen hengen Grävnavet-ryhmä, joka keskittyy datajournalismin ja tutkivan journalismin tekemiseen eri välineisiin.

Siinä missä Plus-deski tekee juttuja lähinnä verkkoon, Grävnavet tekee niitä verkkoon, radioon ja televisioon. Sjöholmin mukaan melkein kaikissa ryhmän jutuissa käytetään jonkinlaista dataa, mutta kaikkiin juttuihin ei tehdä verkkovisualisointia.

Alun perin ryhmässä oli Sjöholm, virkavapaalta palannut Ugglan ja tuottaja Eva-Maria Koskinen. Sittemmin ryhmän kokoonpano on muuttunut. Koskinen jäi opintovapaalle, ja hänen sijaisekseen valittiin toimittaja Malin Ekholm. Ugglan muutti Ruotsiin, ja yt-neuvotteluiden takia hänen tilalleen ei saatu rekrytoida uutta henkilöä. Siksi syksyllä 2014 otettiin käyttöön kolmen viikon kiertosysteemi, jossa Svenska Yle Uutisten toimittajat työskentelevät kukin vuorollaan Grävnavet-ryhmässä. Ryhmän kolmantena jäsenenä on pysynyt Sjöholm, jolla on tällä hetkellä tuottajavastuu. Hän on myös tällä hetkellä ainoa ryhmän jäsenistä, joka osaa koodata. Aiemmin myös Ugglan koodasi. Ryhmässä ei ole varsinaista graafikkoa, mutta sen jäsenillä on jonkin verran graafista osaamista. Tarpeen mukaan ryhmä saa talon tv-graafikoilta apua.

Grävnavet-ryhmän lisäksi Svenska Ylellä työskentelee ohjelmointitaitoinen datatoimittaja Petri Kola. Hän toteuttaa datajournalistisia juttuja yhdessä Svenska Ylen toimittajien kanssa. Pääasiassa hän tekee yhteistyötä muiden kuin uutispuolen toimittajien kanssa. Kolan rooli on yleensä teknisen toteutuksen hoitaminen.

Sjöholm kuvaili lokakuussa 2014 datajournalismin asemaa Svenska Ylessä vakaaksi. Hän uskoo, että päällikkötasolla ollaan kiinnostuneita datajournalismista. Sjöholm mainitsi nimeltä Hans Sundquistin ja Kerstin Häggblomin, jotka ovat hänen mukaansa erityisesti vaikuttaneet siihen, että Svenska Ylellä on alettu tehdä datajournalismia.

Suurin ongelma päivittäisessä tekemisessä on koodaavien toimittajien puute. Sjöholm epäilee, että koodaajia on vaikea saada kiinnostumaan journalismista etenkin aikana, jolloin vähän jokaisessa mediatalossa käydään tai on käyty yt-neuvotteluja.

#### **4.4 STT-Lehtikuva**

STT-Lehtikuvassa datajournalismi oli elokuuhun 2014 saakka pitkälti ammattikorkeakoulu Metropolian harjoittelijoiden varassa. Ensimmäinen datajournalismiin keskittynyt harjoittelija oli Wolf Wikgren syksyllä 2013. Hän kartoitti ja loi uusia työkaluja sekä toteutti teknisesti toimittajien ideoita. Wikgrenin jälkeen talossa ovat käyneet Elina Kuivinen ja Anna Hurttu, ja välillä on ollut jaksoja, ettei talossa ole ollut harjoittelijoita.

Aloite datajournalismiin keskittyvistä harjoittelijoista tuli STT-Lehtikuvasta, joka haluaa panostaa asiaan. Harjoittelijoilta on haettu nimenomaan teknistä osaamista, jota toimituksessa ei muuten ole ollut. He ovat hankkineet ja käsitelleet dataa ja tehneet visualisointeja. Käytössä oli ja on edelleen loppuvuodesta 2014 pääosin valmiita työkaluja, joita käytetään työskentelyn pohjana. Hieman monimutkaisempien asioiden tekeminen vaatii koodaamista, ja osin käytetään Wikgrenin luomia työkaluja.

Elokuussa 2014 STT-Lehtikuva palkkasi Wikgrenin määräaikaiseksi datajournalismin tekijäksi ja kehittäjäksi. Uutistoimisto haluaa datajournalismin tekemisestä pysyvää ja säännöllistä. Tarkoitus on saada oma digiyksikkö, jossa työskentelisi teknisiä osaajia, toimittajia ja graafikoita. Tuottaja Harri Juutilaisen mukaan digiyksikkö aiotaan perustaa vuoden 2015 aikana.

Toimittajiensa kouluttamisen STT-Lehtikuva aloitti keväällä 2014. Helsingin Sanomien uutispäällikkö Esa Mäkinen opasti toimittajia datan hankkimisessa ja jalostamisessa sekä joidenkin visualisointityökalujen käytössä. Lisäksi Juutilainen piti excel-koulutusta. Kesällä 2014 datajournalismia ja visualisointeja harjoiteltiin ja tehtiin ilman lupauksia juttujen määrästä, ja saman vuoden elo-syyskuussa määriteltiin tuotantotavoitteet.

STT-Lehtikuva on luvannut asiakaslehdilleen tuottaa heidän käyttöönsä suunnilleen kaksi verkkovisualisointia viikossa. Utistointisto ei kuitenkaan ole luvannut niiden olevan välttämättä datajournalismia, vaan ne voivat olla esimerkiksi aikajanoja tai karttoja. Dataan pohjautuvaa ja kaivelua vaativaa journalismia STT-Lehtikuva tekee Juutilaisen mukaan harvemmin kuin viikoittain.

Asiakkaat eli ensisijaisesti maakuntalehdet ovat suhtautuneet STT-Lehtikuvan visualisointeihin vaihtelevasti. Juutilaisen mukaan yksittäiset asiakkaat aktivoituivat heti kesällä 2014 ja alkoivat nostaa visualisointeja näkyvästi esille verkkosivuillaan. Nämä asiakkaat ovat sellaisia, jotka tekevät itsekin aktiivisesti verkkojournalismia. Tällaisia ovat ainakin Aamulehti ja Keski-suomalainen. Kesälomien jälkeen uutistointiston visualisoinneille on tullut merkittävästi lisää käyttäjiä.

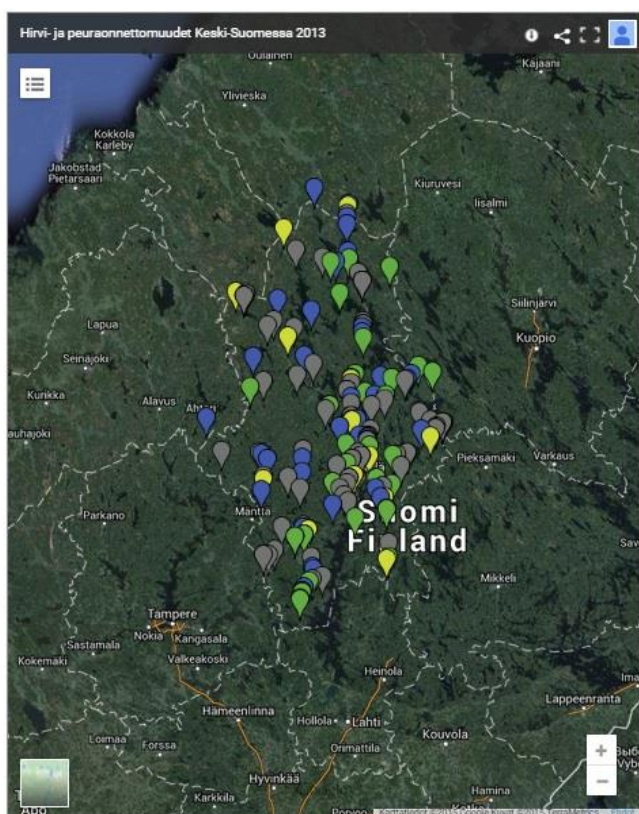
Juutilaisen mukaan datajournalismiin panostamista ovat ajaneet toimituksessa niin vastaava päätoimittaja ja toimitusjohtaja Mika Pettersson kuin uutispäätoimittaja Minna Holopainen. Myös monet uutispäälliköt ovat olleet kiinnostuneita asiasta. Pullonkaulana ovat olleet käytännön järjestelyt. Esimerkiksi excelin käyttö on ollut vaikeaa pitkän linjan toimittajille. Siksi asiasta on järjestetty koulutusta.

## **4.5 Maakuntalehdet**

Maakuntalehtien datajournalismi on ollut enimmäkseen yksittäisiä kokeiluja. Esimerkiksi Keski-suomalaisen verkkoliiketoiminnan päällikkö Ville Grahn kertoi lokakuussa 2014, että lehden datajournalistiset jutut ovat laskettavissa kahden käden sormilla. Hänen mukaansa lehden toimittajia koulutettiin tekemään datajournalismia puolivakavissaan puolitoista vuotta aiemmin, mutta juttuja ei ole juuri syntynyt. Asialle on Grahnin näkemyksen mukaan kaksi syytä: kiire ja sopivien tekijöiden puute. Koulutusta saaneet toimittajat eivät ole onnistuneet soveltamaan oppimaansa, ja ylipäätään toimittajien tietotekniset taidot ovat melko huonot. Grahnin mukaan datajournalismin tekeminen

vaatisi luontaista innostusta, jota ei juuri löydy toimituksesta. Rekrytointeihin puolestaan ei ole varaa.

Lokakuussa 2014 Keski-suomalaisessa oli yksi verkkotoimittaja, joka pystyy tekemään yksinkertaisia visualisointeja. Eriikka Hienonen on tehnyt esimerkiksi Googlen työkaluilla kartan hirvikolareista (ks. kuvio 5). Lisäksi Keski-suomalainen laittaa verkkosivuilleen kaikki STT-Lehtikuvan tuottamat datajournalistiset jutut.



Kuvio 5: Ruutukaappaus Keski-suomalaisen hirvikolarikartasta (1.4.2015).

Aamulehti on askeleen Keski-suomalaista edellä, vaikka sielläkin datajournalismin tekeminen on vähäistä. Muutaman ilmaistyökaluilla tehdyn jutun lisäksi Aamulehti kehittää uutta visualisointityökalua start up -yritys Dexmenin kanssa. Kehitystyötä tehdään Åkerlundin säätiön rahoituksen turvin. Tavoitteena on tehdä työkalu toimittajien ja graafikoiden käyttöön useisiin eri medioihin. Työkalua ideoineen Aamulehden toimittajan Taneli Koposen mukaan työkalun oli tarkoitus mennä pilottikäyttöön marraskuussa 2014.

Koponen ei tee päätoimisesti datajournalismia, mutta on tehnyt sitä Aamulehden toimittajista eniten. Hän kertoo, että talvella 2013–2014 Aamulehdessä tehtiin kokeiluja verkosta löytyvillä ilmaisilla

työkaluilla. Kokeiluissa syntyi muun muassa interaktiivista tilastografiikkaa ja aikajana. Lisäksi Aamulehdessä on tehty joitakin dataan pohjautuvia, mutta perinteisen muotoisia juttuja printtiin ja verkkoon. Mihinkään näistä ei ole tarvittu koodaustaitoja, mutta exceliä on pitänyt käyttää.

Aamulehden verkkopalvelujen kehityspäällikön Seppo Roth sanoi lokakuussa 2014, että Lännen media alkaa mahdollisesti tehdä datajournalismia tulevaisuudessa. Lännen media on 12 lehden muodostama yhtiö, joka tuottaa journalistista sisältöä ja jossa Aamulehti on mukana. Rothin mukaan syksy 2014 on mennyt toiminnan käynnistämiseen, ja digitaaliseen mediaan liittyvät kehityshankkeet ovat kilometrin mittaisella listalla. Datajournalismia Lännen media saattaisi alkaa tehdä vuonna 2015.

Keskisuomalainen, Karjalainen, Savon Sanomat ja Etelä-Suomen Sanomat puolestaan kertoivat alkavansa hyödyntää tutkivan ja datajournalismin menetelmiä, kun ne perustivat yhteisen virtuaalitoimituksen huhtikuussa 2014 (Kärkimedia 2014). Keskisuomalaisen Ville Grahnin mukaan tämä ei ole datajournalismin osalta toteutunut. Grahn sanoo, että virtuaalitoimitus on ehkä käyttänyt tutkivan journalismin menetelmiä ja onnistunut tekemään hyviä skuuppeja, mutta tekemisen tapa on perinteistä uutistystä.

Aamulehden ja Keskisuomalaisen esimerkit kuvastanevat sitä, mitä maakuntalehtien datajournalismi on vuoden 2014 lähestyessä loppuaan. Joissain maakuntalehdissä datajournalismia on tehty vielä vähemmän, tuskin missään juuri enemmän.

## 4.6 MTV

Kun MTV Uutiset perusti vuoden 2014 alussa tutkivaan journalismiin keskittyvän [tutkiva45.fi](http://www.tutkiva45.fi)-sivuston, vastaava päätoimittaja Merja Ylä-Anttila kuvaili asiaa näin:

*Journalismin kuluttaminen ja sen myötä myös tekemisen tavat muuttuvat nyt vauhdilla. Tutkivan ryhmämme uudelleenjärjestely tuo omat aiheet, selvitykset ja datajournalismin nopeammin yleisön ulottuville. Lopputuloksena on siis enemmän – ei vähemmän – omaa tiedonhankintaa. (MTV 2014.)*

Lausunto antaa ymmärtää, että datajournalismia on tehty ennenkin, esitystavat vain muuttuvat. Entistä enemmän tehdään verkkoon, eikä tutkiva journalismi ole enää television viikoittaisesta lähetyksistä kiinni. Lausunto liittyy datajournalismin vahvasti tutkivaan journalismiin.



Myös tuottaja Hanna Takala näkee datajournalismin ennen kaikkea osana tutkivaa journalismia. Hänen mukaansa MTV:llä on kerätty, analysoitu ja hyödynnetty dataa journalismissa 1990-luvun lopusta lähtien. Silloin sitä ei vain kutsuttu datajournalismiksi. Edelleenkin lokakuussa 2014 MTV:llä ei ollut datajournalismiin keskittyvää yksikköä, vaan tutkivan journalismin ryhmä tekee sitä osana työtään. Takalan vetämä ryhmä myös kouluttaa talon muita toimittajia esimerkiksi hyödyntämään julkisia tietokantoja ja sähköisiä rekistereitä sekä käyttämään exceliä. Ryhmän edeltäjiä olivat 45 minuuttia - ohjelma ja Sunnuntairaportti.

Tutkivan journalismin ryhmä koostuu toimittajista ja tuottajista. Sen jäsenillä on osaamista datan käsittelemisessä ja analysoimisessa, mutta ei niinkään visuaalisuudessa. Visuaalisuus ei nouse juttujen pääosaan, vaan keskiössä on uutinen. Ryhmä tekee kuitenkin harvakseltaan esimerkiksi interaktiivisia karttoja. Kukaan sen jäsenistä ei osaa koodata.

Datajournalismia tehdään jonkin verran myös muualla MTV Uutisissa. Takala mainitsee esimerkiksi Ulvilan murhamysteeriä käsittelevän aikajanan, jossa visuaalisuus nousee isoon rooliin. Syksyllä 2014 MTV:n multimediatoimitukseen otettiin dataharjoittelija, joka osaa ohjelmoida ja tehdä grafiikoita. Jani Savolainen tekee töitä ennen kaikkea uutispuolella, mutta Takalan mukaan tulossa on myös yhteistyöprojekteja tutkivan journalismin ryhmän kanssa.

## 4.7 Talousmediat

Talousmediat suhtautuvat datajournalismiin hieman samalla tavalla kuin MTV: niissä harvoin puhutaan datajournalismista, mutta jutut perustuvat usein dataan, ja grafiikka on kovassa käytössä. Se on luonnollista, koska mediat keskittyvät talouteen. Esimerkiksi Kauppalehden toimituspäällikön Jari Saarion mukaan toimituksessa ei lokakuussa 2014 ole varsinaista datajournalistia, eikä siellä puhuta datajournalismista. Kauppalehdellä on kuitenkin oma analyttikko, joka tekee numeroista kirjallisia analyyseja, ja toimittajat kirjoittavat niiden pohjalta juttuja.

Talouselämän toimituspäällikkö Olli Ainola mainitsee lehden kärkituotteiksi yritysanalyysit ja liikevaihdolla mitaten 500 Suomen suurimman yrityksen selvittämisen. Ne ovat isoja kokonaisuuksia, jotka tehdään vuosittain. Ainolan mielestä ne ovat siinä mielessä datajournalismia, että niihin liittyy valtava numerotiedon kerääminen. Se tehdään pääosin käsityönä.

Käytännössä prosessi menee niin, että Aalto-yliopiston kauppakorkeakoulusta palkataan vuosittain opiskelijoita avustajiksi keräämään dataa. Talouselämän omat informaattikot eli yritystutkijat analysoivat kerätyn datan. Lopulliset tuotokset tekevät lehden omat taloustoimittajat ja graafikot. Talouselämä myös myy keräämäänsä dataa saadakseen lisätuottoa.

500 Suomen suurinta yritystä -kokonaisuudesta tehtiin vuonna 2012 pelillinen visualisointi (ks. kuvio 6). Sen teknisestä toteutuksesta vastasi datajournalistina profiloitunut Jens Finnäs, joka teki työn freelancerina. Ainola näkee sen kokeiluna, jolle ei ainakaan vielä ole tehty jatkoa. Hänen mukaansa Talouselämä ei ole löytänyt sellaisia datajournalismin muotoja, joissa sijoitettu panos tuottaisi taloudellisesti tai journalistisesti sellaista palautetta, että niitä kannattaisi tehdä. Hän viittaa tällä verkkovisualisointeihin ja ohjelmointitaitoa vaativiin lopputuloksiin, ei perinteiseen grafiikkaan tai dataan pohjautuviin artikkeleihin.



*Kuvio 6: Ruutukaappaus 500 Suomen suurinta yritystä -visualisoinnin alusta (1.4.2015).*

Taloussanomien suhtautuminen datajournalismiin eroaa Kauppalehden ja Talouselämän suhtautumisesta. Datajournalismi kuuluu uutispäällikkö Aleksis Moision henkilökohtaisiin kiinnostuksen kohteisiin, ja Taloussanomien datajournalismissa korostuu visuaalisuus. Moision mukaan yleisölle näkyvät erityisesti erilaiset laskurit ja kartat. Hän huomauttaa, että Taloussanomien on tehnyt vanhimmat datajournalistiset projektinsa silloin, kun koko termiä ei oltu keksitty. Vuodesta 2011 lähtien asioita on vain tehty hieman modernimmalla tavalla, visuaalisuutta korostaen.

Moisio kertoi lokakuussa 2014, että Taloussanomien tekee suunnilleen yhden datajournalistisen jutun kuukaudessa. Pyrkimyksenä on kasvattaa juttujen määrää tekemällä valmiita malleja ja monistettavia tekotapoja sekä kouluttamalla toimittajia näkemään itse mahdollisuuksia ja käyttämään erilaisia työkaluja. Taloussanomien käyttää esimerkiksi Googlen tarjoamia työkaluja, ja vasta muutama toimittaja on koulutettu käyttämään niitä. Kouluttajana on toiminut Moisio, ja lisäksi toimittajia on käynyt Sanoma Akatemian kursseilla. Pääosin datajournalismia tekevät toimittajat Harri Pietarinen ja Johannes Niemeläinen. Toimituksessa ei ole ohjelmointitaitoisia henkilöitä, mutta nettikehittäjät osallistuvat tarvittaessa projekteihin. Lisäksi juttuja tilataan avustajilta, erityisesti Suomen Rahatiedolta.

## 4.8 Iltapäivälehdet

Iltapäivälehtien suhtautuminen datajournalismiin poikkeaa toisistaan. Iltalehden verkon esimiehen Timo Melarin mukaan lehti ei ole panostanut datajournalismiin, eikä sillä ole aikomustakaan panostaa siihen. Juttuja tehdään sisältö edellä, ei minkään tietyn esitystavan pakottamana. Lehdestä kyllä löytyisi tarvittaessa ulkoasu- ja tilasto-osaamista, mutta Melarin mukaan on vain vähän aiheita, joissa datajournalismi toisi lisäarvoa lukijalle. Hän ei saa mieleensä yhtään esimerkkiä Iltalehden tekemästä datajournalismista.

Ilta-Sanomien sosiaalisen median tuottaja Matti Markkola puolestaan sanoi lokakuussa 2014, että lehdessä tehdään datajournalismia joka päivä, jos termillä tarkoitetaan rikastettua journalismia kuten tietokantojen hyödyntämistä tai visualisointeja. Ilta-Sanomissa ei kuitenkaan juuri käytetä sanaa datajournalismi. Markkolan mielestä se on sisäinen termi, joka ei kiinnosta lukijaa.

Tyypillisimmin Ilta-Sanomien datajournalismi tarkoittaa karttoja tai grafiikkaa. Toimituksessa käytetään paljon esimerkiksi Googlen työkaluja, joiden käytön kuka tahansa toimittaja voi opetella. Tarvittaessa toimittaja saa apua kokeneemmilta, esimerkiksi verkkotoimituksen uutispäälliköiltä. Jos juttua varten tarvitsee koodata tai käyttää exceliä, apua voi pyytää kehitystiimin jäseniltä, jotka eivät kuulu varsinaisesti toimitukseen. Ilta-Sanomissa ei siis ole organisoitu datajournalismin tekemistä mihinkään erityiseen yksikköön, eikä sellaista ole suunnitelmissa. Ilta-Sanomissa tehdään paljon myös testejä ja pelejä, mutta niitä Markkola ei laske datajournalismiksi.

## 4.9 Ruotsinkieliset lehdet

Ruotsinkieliset lehdet ovat tehneet datajournalismia ennen kaikkea yhteistyönä. Hufvudstadsbladet, Vasabladet, Åbo Underrättelser ja Nya Åland toteuttivat joulukuussa 2013 ja kesällä 2014 isot projektit, joissa kerättiin lukijoilta dataa lehtien verkkosivuilla. Toisen projektin aiheena oli raha, toisen asuinpaikka. Lehdet laittoivat aiheisiin liittyvän kyselyn verkkosivuilleen, ja esimerkiksi asuinpaikkaa koskevaan kyselyyn tuli noin 5 000 vastausta. Saadun datan pohjalta tehtiin useita artikkeleita.

Projektien moottoreita olivat Hufvudstadsbladet ja Vasabladet eli nelikon isoimmat lehdet. Projekteja oli toteuttamassa useita toimittajia, graafikoita ja nettikehittäjiä sekä datajournalistina profiloitunut Jens Finnäs, joka teki oman osuutensa freelancerina. Hän osallistui ideointiin, hoiti teknisen toteutuksen ja teki myös sisältöä yhdessä lehtien toimittajien kanssa. Alkuperäisen idean takana on lehtien esimiehiä.

Hufvudstadsbladetin vastaavan päätoimittajan Lena Skogbergin mukaan suunnitelmissa on tehdä lisää yhteisiä projekteja Jens Finnäsin ja muiden ruotsinkielisten lehtien kanssa. Hän sanoo, että Hufvudstadsbladetilla on lisäksi joitain omia projekteja. Myös Vasabladetilla voi olla omia projekteja, mutta pienemmät lehdet eivät pysty yksin tuottamaan datajournalismia. Isommillekin se on helpompaa yhteistyönä.

Hufvudstadsbladetissa toimittaja Johan Gullmets teki jonkin verran datajournalismia, mutta on sittemmin siirtynyt Ylelle. Skogbergin mukaan syynä siirtymiseen oli se, että Yle tarjoaa paremmat mahdollisuudet datajournalismin tekemiseen. Hufvudstadsbladetissa on kyllä järjestetty koulutusta ja työkaluja, mutta silti datajournalismia on tehty hyvin vähän. Suunnitelmissa on kuitenkin panostaa siihen. Toimituksessa oli lokakuussa 2014 meneillään organisaatiomuutos, ja tarkoitus oli, että vuoden 2015 alusta lähtien toimittaja Staffan Småroos keskittyy datajournalismiin. Skogberg arvioi, että Småroos pystyy jatkossa käyttämään noin 40 prosenttia työajastaan datajournalismin tekemiseen. Småroos on itse osoittanut olevansa kiinnostunut aiheesta ja on esimerkiksi osallistunut koulutuksiin.

Hufvudstadsbladetissa on järjestetty kursseja esimerkiksi datan keräämisestä. Kouluttajana on ollut ruotsalainen toimittaja Helena Bengtsson, joka työskentelee nykyään brittiläisessä The Guardianissa. Kolmen vuoden aikana on järjestetty muutama kurssi, ja niihin on osallistunut kymmenisen toimittajaa. Lisäksi toimittajia on osallistunut Helsingin yliopiston ruotsinkielisen yksikön kursseille ja

ruotsalaisen tutkivan journalismin yhdistyksen seminaareihin. Skogbergin mukaan oppi ei vain juuri näy työskentelyssä.

#### 4.10 Yhteenveto

Katsaus osoittaa, että datajournalismin osalta medioissa liikutaan varsin monella tasolla. Helsingin Sanomien datadeskissä ja Ylen Plus-deskissä datajournalismia tehdään säännöllisesti, ja tuotokset kootaan omille sivuilleen. Osa jutuista on hyvinkin näyttäviä, ja niiden tekemiseen menee runsaasti aikaa. Toisaalta taas deskeissä tehdään paljon pieniä ja arkisia tuotoksia, joita on tehty jo niin paljon, että ne syntyvät nopeasti. Datajournalismista on siis tullut osa toimitusten kiireistä arkea.

Maakuntalehdissä sen sijaan pienikin datajournalistinen tuotos vaatii yksittäisen toimittajan harrastuneisutta ja esimiesten tukea. Harvalla toimittajalla on aikaa, innostusta tai taitoa tehdä datajournalistisia juttuja. Jos jotain tehdäänkin, tuotokset ovat yksinkertaisia visualisointeja.

STT-Lehtikuva pyrkii panostamaan verkonomaiseen kerrontaan maakuntalehtiä enemmän. Oman digiyksikön perustaminen on iso ponnistus, ja sille on asetettu konkreettisia tavoitteita esimerkiksi juttujen määristä. Juttujen ei kuitenkaan luvata olevan datajournalismia, vaan ne voivat olla myös esimerkiksi aikajanoja tai karttoja, joita STT-Lehtikuva ei laske datajournalismiksi. Yksi haaste on, miten STT-Lehtikuva saa sen asiakaslehdet opetettua julkaisemaan verkkovisualisointeja.

Ruotsinkielisissä medioissa, Yle mukaan lukien, tuntuu olevan runsaasti kiinnostusta datajournalismia kohtaan. Niissä tehdään isojaakin projekteja, ja etenkin Ylessä datajournalismin asema on melko vakaa. Ongelmana on kuitenkin hyvien tekijöiden puute.

Loput katsaukseen valikoituneet mediat – MTV, iltapäivälehdet ja talousmediat – suhtautuvat datajournalismiin varsin kirjavasti. Yhteistä niille on, että datajournalismi-termiä ei juuri käytetä, eikä datajournalismin tekeminen ole säännöllistä. Ennemmin tunnutaan tekevän jotain sen suuntaista: grafiikkaa, visualisointeja tai tutkivaa journalismia. Ainoastaan Taloussanomissa on pyrkimys tehdä datajournalismia vakiintuneemmin, sillä datajournalismi kuuluu uutispäällikkö Aleksin Moision henkilökohtaisiin kiinnostuksen kohteisiin.

Katsauksen pohjalta voi sanoa, että suomalaisissa toimituksissa ei ole selvää, mitä datajournalismilla oikein tarkoitetaan. Esimerkiksi Helsingin Sanomat on ratkaissut ongelman laskemalla

datajournalismiksi kaiken, mikä on interaktiivista tai missä on dataa taustalla. Kaikki eivät kuitenkaan niele näin laajaa määritelmää. Joidenkin mielestä koko termi on turha.

Katsauksessa nousi esiin monta muutakin datajournalismiin liittyvää ongelmaa. Mistä löytää päteviä tekijöitä? Entä aikaa? Miten datajournalismia voi tehdä kannattavasti? Monet auki olevat kysymykset ovat hyvin konkreettisia ja arkisia. Samalla ne kertovat siitä, että datajournalismi ei ole vakiinnuttanut asemaansa suomalaisessa mediakentässä.

## 5 Datajournalismin ilmiön selittämistä

### 5.1 Haastateltavat ilmiön arvioijina

*Minä: No osaatko sä arvioida sitten datajournalismin laajempaa merkitystä? Mitä merkitystä sillä on yleisölle tai yhteiskunnalle?*

*Wolf Wikgren: Suureellista. Ja mun pitäis vastaa, ei vit... huh huh. Mä pelkään että mut vielä otetaan tosissaan näissä mun vastauksissa.*

*Minä: Totta kai otetaan.*

*Wolf Wikgren: Oi voi, sillee kesän ollu harjottelijana ja... öö, hmm, öööö... yhteiskunnallinen merkitys. Kai se liittyy vähän samoihin asioihin sinänsä mikä se lisäarvo on journalismille, eli pystytään esittämään asioita mitä ei oo aiemmin pystytty esittämään. Ainakin tietyissä määrin näin. Et et pystytään käsittelemään semmosia asioita mitä on ehkä ollu vaikeampi käsittää ja esittää aiemmin. En tiedä.*

*(Wolf Wikgren, STT-Lehtikuva, datajournalismiharjoittelija.)<sup>3</sup>*

Tämä luku pohjautuu kahdeksaan teemahaastatteluun, jotka tein maaliskuussa 2014. Asetin haastateltavat hieman epäkiitolliseen asemaan ja kysyin kaikilta samat datajournalismiin liittyvät kysymykset, vaikka heidän taustansa ovat hyvin erilaiset. Puolet haastateltavista kertoi tehneensä alle kymmenen juttua, jotka he laskevat datajournalismiksi. Loput olivat tehneet reilu kymmenen, kymmeniä tai jopa satoja datajournalistisiksi katsomiaan juttuja. Kuten yllä oleva ote Wolf Wikgrenin haastattelusta osoittaa, haastateltavat kykenivät kuitenkin vastamaan kysymyksiin ja arvioimaan datajournalismia myös omaa kokemusta laajempaan ilmiöönä.

Haastateltavat ovat STT-Lehtikuvan Esa Koivuranta ja Wolf Wikgren, Yle Uutisten Teemo Tebest, Tuomo Björkstén ja Johan Jaakkola sekä Helsingin Sanomien Esa Mäkinen, Jyri Hänninen ja Tuomas Peltomäki. Esittelin kaikki haastateltavat tarkemmin luvussa 3.3.

Tässä analyysin osassa pureudun datajournalismin määritelmään ja merkitykseen. Vastaan siis ensimmäiseen ja toiseen tutkimuskysymykseen (ks. luku 1.2). Luvun tavoite on tehdä tulkintoja siitä, mitä ajatuksia haastateltavilla on datajournalismista. Aloitan termin määrittelystä ja etenen datajournalismin merkitykseen journalismille, yhteiskunnalle ja yleisölle. Siihen liittyen käsittelen myös datajournalismin asemaa Suomessa nyt ja tulevaisuudessa.

---

<sup>3</sup> Kerron jokaisen sitaatin yhteydessä haastateltavan nimen ja sen, missä taustaorganisaatiossa ja työtehtävässä henkilö on tehnyt datajournalismia, jotta sitaattia voi peilata haastateltavan kokemukseen.

## 5.2 Kuka on datajournalisti?

Kysymys siitä, pitääkö datajournalismin tekijän olla journalisti, on pitkälti sen määrittelyä, kuka voi olla journalisti. Perinteisesti journalismin pariin on hakeutunut paljon ihmisiä, joilla ei ole journalistista koulutusta. Moni on myös jättänyt journalismin opinnot kesken. Journalismi ei ole professio, jonka tekemiseen vaadittaisiin akateeminen loppututkinto (Pietilä 2012).

On selvää, että haastateltavien omat kokemukset vaikuttavat tulkintaan datajournalismista. Esimerkiksi Wikgren opiskeli mediatekniikan insinööriksi, eikä hänellä ollut journalistista kokemusta ennen STT-Lehtikuvan datajournalismiharjoittelua. Sielläkin hänen tehtävänsä liittyivät pitkälti tekniikkaan. Wikgren toi monta kertaa haastattelun aikana esiin, että hän vastaa kysymyksiin ennen kaikkea teknisestä näkökulmasta.

Myös Ylen Plus-deskissä työskentelevä Tebest on ennen kaikkea koodaaja. Hän on valmistunut tietotekniikan diplomi-insinööriksi pääaineenaan hypermedia ja työskennellyt aiemmin tutkijana. Haastattelujen perusteella molemmilla on kuitenkin hyvä taju journalismista ja sen tekemisestä. Heidän kanssaan työskennelleet toimittajat tuntuvat myös arvostavan heidän osaamistaan. Esimerkiksi Wikgrenin kanssa yhteisiä projekteja tehnyt Koivuranta sanoi, että kokee Wikgrenin olevan toimittaja. Mäkinen huomautti, että koodaajan on helpompi opetella journalismia kuin journalistin koodaamista, joskin hän mielestään “skene on vähintäänkin epäselvä” sen takia.

Muiden haastateltavien kuin Wikgrenin ja Tebestin tausta on ennen kaikkea journalismissa. Heille on kertynyt eri verran vuosia journalismin parissa, enimmillään 12 vuotta. Peltomäki myös koodaa ja tekee grafiikkaa, ja uutispäällikkönä työskentelevä Mäkinen koordinoi, koodaa ja tekee vähän kaikkea. Muiden rooli datajournalistisissa projekteissa on ennen kaikkea toimittajan rooli. He ideoivat, haastattelevat, kirjoittavat, ehkä myös hankkivat ja käsittelevät dataa.

Useampi haastateltava sanoi, ettei koe itse olevansa datajournalisti, vaikka en kysynyt sitä suoraan kaikilta.



*No mä ajattelen itseäni toimittajana, en mä, ei mulla oo koskaan käynyt mielessäkään että mä olisin datajournalisti. Mä tykkään kaivaa, siis penkoa tällöisiä isoja asiakokonaisuuksia, jotka vaatii sitä että siinä on teknistä apua, mut lähtökohtaisesti mä en ole datajournalisti ja mä katson, että ennen kaikkea sen takia että mä en esimerkiksi pystyisi itse toteuttamaan tällöistä, tällöisiä hankkeita yksin. Mun on pakko kävellä nyt vaan Esa Mäkisen luo tai jonkun muun, joka pystyy ennen kaikkea teknisissä asioissa jeesaamaan mua ja viemään hommia eteenpäin.*

(Jyri Hänninen, Helsingin Sanomat, taloustoimittaja.)

*On toimittajia ja koodareita ja graafikoita, ja sit on lopputuloksia, jotka on datajournalismia.*  
(Johan Jaakkola, Yle Uutiset, taloustoimittaja.)

Haastatteluissa tuli siis esiin yhteistyön merkitys. Ei ole datajournalisteja, vaan on datajournalismia tekeviä toimittajia, koodareita ja graafikoita. Kun kysyin, kuka on datajournalisti tai mitä datajournalistin pitää osata, vastauksissa korostui kuitenkin tekninen osaaminen. Puolet mainitsi vähimmäisvaatimuksena excel-aidot ja loputkin jossain muodossa tietotekniset taidot. Koodaaminen nousi usein esiin, mutta ei välttämättömyytenä vaan keinona syventää osaamista. Suurin osa mainitsi vaatimuksena myös toimittajuuden tai joitakin toimittajalle tyypillisiä ominaisuuksia, kuten kiinnostuksen maailmaa kohtaan.

Tebest nosti haastattelussa esiin, että Suomessa ei voi kouluttautua datajournalistiksi. Esimerkiksi visuaalisten esitystapojen osaajaksi tai toimittajaksi voi kouluttautua, mutta ei suoraan datajournalistiksi. Jos – etenkin teknisiä – taitoja haluaa opetella lisää, pitää olla omatoiminen. Usein se tarkoittaa haastateltavien mukaan istumista koneen ääressä, googlaamista ja kokeilemistä. Verkko on pullollaan englanninkielistä materiaalia itseopiskeluun. Datajournalismin tekeminen ei ole vielä kovin vakiintunutta, ja jopa sen konkari Mäkinen kertoo, että lähes joka päivä joutuu opettelemaan jotain uutta. Opettely kulkee usein käsi kädessä kulloisenkin projektin kanssa, eli taitoja omaksutaan sitä mukaa, kun niitä tarvitsee. Verkon lisäksi ideoita ja apua haetaan kollegoilta.

### **5.3 Kohti datajournalismin määritelmää**

Haastatteluissa nousi monta kertaa esiin datajournalismi-käsitteen epämääräisyys. On vaikea kertoa, monta datajournalistista juttua on tehnyt, jos määritelmä ei ole selvä. En antanut haastateltaville edes suuntaa antavaa määritelmää, vaan jokainen sai itse kertoa, mitä laskee datajournalismiksi.

Osa haastateltavista ei pidä datajournalismin määrittelemistä kovin kiinnostavana. Mäkinen kertoi,

ettei jaksakaan enää väitellä ihmisten kanssa aiheesta.

*Täähän on sellanen keskustelu mitä aina välillä käydään, että onko joku juttu datajournalismia vaiko ei, ja mä oon niinku tavallaan lopettanu sen asian pohtimisen. Että niinku jos se on meidän yksikön tekemää, ni silloin sitä nimitetään Hesarissa datajournalismiksi, joskin ei nyt ihan kaikkea, jos mä kirjoitan ite jonkun kolumnin tai jonkun ihan tavallisen uutisen ni ei sitä nyt nimitetä -- et varmaankin se että siinä on joko käytetty dataa jollain tavalla tai sitten siinä niinku on jonkunnäkönen interaktiivinen visualisointi, ja ne on niinku ehkä ne kaks olennaista tekijää.*

(Esa Mäkinen, Helsingin Sanomat, datajournalismista vastaava uutispäällikkö.)

Helsingin Sanomat on siis tavallaan ratkaissut määrittelyongelman perustamalla datadeskin. Sen tuotoksia nimitetään datajournalismiksi, jos ne joko pohjautuvat dataan tai näyttävät datajournalismilta eli niissä on interaktiivinen visualisointi. Ylen Plus-deski ei ole lähtenyt samalle linjalle, vaan se kertoo tekevänsä verkonomaista journalismia, josta osa on datajournalismia. Tebest arvioi, että Plus-deskin vuonna 2013 tekemästä 350:stä jutusta vain 50 olisi ollut datajournalismia.

Tiivistin kuvioon 7 haastateltavien vastauksia kysymyksiin, jotka auttavat datajournalismin määrittelyssä:

	<b>Hyvä datajournalismi</b>	<b>Huono datajournalismi</b>	<b>Ero muuhun journalismiin</b>	<b>Datajournalismin tehtävä</b>
<b>Esa Koivuranta</b>	sisältö edellä, tutkivaa journalismia	keinotekoinen visualisointi, tiedotepohjaisuus, merkityksettömyys	datapohjaisuus, usein koneella tekeminen, usein visuaalisuus	sama kuin muulla journalismilla
<b>Wolf Wikgren</b>	kaivaa datasta ilmiön, selkeä esitystapa, interaktiivisuus ei itseisarvona	huono tai turha grafiikka, virhepäätelmät, interaktiivisuus itseisarvona	datan käsittely ja analyysi, usein interaktiivinen grafiikka	sama kuin muulla journalismilla, löytää datasta tarinan tai yhdistää dataa järkevästi
<b>Teemo Tebest</b>	vaikuttavuus, merkittävyys, kaivaa datasta uutisen	sama kuin muu huono journalismi, sokaistuu datalle ja kadottaa kontekstin	verkonomaisuus, datapohjaisuus	sama kuin muulla journalismilla, tehdä datasta ymmärrettävää
<b>Johan Jaakkola</b>	hyvä esitystapa, interaktiivisuus, visuaalisuus, kaivaa datasta uutisen	vaikeakäyttöinen visualisointi, huonosti rajattu aihe	interaktiivisuus, datan rooli ja tulkinta	sama kuin muulla journalismilla, laajentaa sitä mitä pystyy tekemään
<b>Tuomo Björkstén</b>	sisältö edellä, hyvä käyttökokemus, helppo omaksua	aikajanan hehuttaminen käänteentekevänä	datapohjaisuus, oma analyysi	sama kuin muulla journalismilla
<b>Esa Mäkinen</b>	uutisellisuus, hyvä analyysi, mielenkiintoisuus, tärkeys, selkeys, teknisesti toimiva	ei toimi teknisesti, data ilman analyysiä, huono analyysi	datan käyttäminen tai interaktiivinen visualisointi	sama kuin muulla journalismilla, datan penkominen
<b>Jyri Hänninen</b>	oleellisuus, merkityksellisyys, uutisellisuus, tärkeys	esim. verkkografiikassa visualisoidaan itsestäänselvyksiä	työtavoissa eroja, uuden yleisön tavoittelu	sama kuin muulla journalismilla, hahmottaa monimutkaisia asioita
<b>Tuomas Peltomäki</b>	selkeys, tärkeys, ymmärrettävyys, merkittävyys, kiinnostavuus	turha interaktiivisuus, epäselkeys	datapohjaisuus, usein visuaalisuus	sama kuin muulla journalismilla, datan selittäminen

*Kuvio 7: Tiivistettyjä poimintoja haastatteluista datajournalismin määrittelyyn liittyen.*

Kysyin siis haastateltavilta, millaista on heidän mielestään hyvä ja huono datajournalismi, miten se eroaa muusta journalismista ja mikä on sen tehtävä. Käyn seuraavaksi hieman tarkemmin läpi haastateltavien vastauksia.

Kun pyysin haastateltavia kertomaan, millaista heidän mielestään on hyvä datajournalismi, esiin nousi monesti esimerkiksi yhteiskunnallinen merkittävyys, uutisellisuus ja kiinnostavuus. Ne ovat adjektiiveja, jotka liitetään usein hyvään journalismiin, ennen kaikkea uutisiin ja tutkivaan journalismiin. Vastauksista huomasin, että kaikki haastateltavat työskentelevät uutistoimituksissa. Jos olisin haastatellut esimerkiksi naistenlehtien tekijöitä, vastaukset olisivat voineet olla hyvin erilaisia.

*Mun mielestä hyvä datajournalismi on semmosta, että joku saa potkut. Ei, se on ehkä vitsi, mutta tarkoitan sitä, että niitten pitää olla vaikuttavia, ja niil pitää olla joku merkitys, vaikuttavuus.*

(Teemo Tebest, Yle Uutiset, Plus-deskin toimittaja.)

Moni mainitsi myös erityisesti datajournalismiin liittyviä piirteitä: uutinen kaivetaan datasta, ja esitystavan tulee olla selkeä ja teknisesti toimiva. Jutussa voi olla esimerkiksi interaktiivinen visualisointi, mutta vain, jos se sopii sisältöön.

Käänteisesti huonossa datajournalismissa on interaktiivisuutta interaktiivisuuden takia, data ilman analyysiä ja toimimatonta tekniikkaa. Haastateltavien mainitsemat huonon datajournalismin piirteet liittyivät usein nimenomaan visuaalisuuteen, interaktiivisuuteen ja dataan, vaikka joukossa oli myös mainintoja yleisesti huonosta journalismista. Moni pohti, onko esimerkiksi keinotekoinen visualisointi lainkaan datajournalismia. Pohdinnasta paistoi läpi datajournalismin määrittelyn vaikeus: haastateltavalle voi tulla interaktiivisesta verkkokartasta mieleen datajournalismi, vaikka se ei hänen mielestään datajournalismia olisikaan. Se kertonee osaltaan, että datajournalismin määritelmä ei ole vakiintunut. Termiä tunnutaan käytettävän kuvaamaan uudenlaisia asioita, joille ei oikein ole keksitty muuta nimitystä.

Ero muuhun journalismiin on häilyvä. Löytyi kuitenkin yksi erottava tekijä, jonka kaikki haastateltavat mainitsivat: datapohjaisuus. Myös datan analyysi on tärkeä. Datan esittäminen ilman omaa analyysiä ei ole hyvää datajournalismia. Haastateltavien mielestä ei siis riitä, että jutun taustalla on data, vaan myös käsittelytapa erottaa jutun tavallisesta journalismista. Lisäksi datajournalismiin liittyvät usein hienot visualisoinnit tai interaktiivinen kerronta, mutta esitystapa ei ole yhtä oleellinen kuin data ja sen käsittelytapa, siis prosessi. Datajournalistinen juttu voi olla myös pelkästään tekstiä.

*Siis niinkö grafiikkatoimituksiaahan on ollu iät ja ajat. Se että ne käppyrät on netissä eikä paperilla, ja se että siinä on joku nappula mitä painaa, ni sehän on grafiikkaa, ei se oo mitään data-, niinku... sitä on tehty aina grafiikkatoimituksissa, joka ikisessä jutussa on aina joku käppyrä, ei siinä oo mitään erikoista. Se että ne liikkuu ja vilkkuu ni sehän monesti vaan niinku huonontaa sitä juttua. Et kyl datajournalismi sinänsä on mun mielestä vähän niinku eri asia, et siinä pitää niinku jollain tavalla kaivaa sitä dataa tai sitten jollain tavalla prosessoida ja analysoida koneellisesti sitä dataa. Sit se menee niinku sen verran askeleen, muutaman askeleen eteenpäin, että sitä voi nimittää joksikin muuksi kuin ihan perusgraffaksi.*  
(Tuomas Peltomäki, Helsingin Sanomat, hs.fin toimitussihteeri.)

Kaikki haastateltavat sanoivat, että datajournalismilla on pääasiallisesti sama tehtävä kuin muulla journalismilla. Useimmille se tarkoittaa uuden tiedon tuottamista tai asioiden selittämistä ihmisille. Yksi sanoi, että niin perusuutisen kuin datajournalismin tehtävä on keksiä journalismi uudestaan, siis keksiä tapoja houkutella lisää yleisöä ja hankkia rahaa. Moni kuitenkin näki datajournalismilla erityistehtäviä tai erityisiä keinoja toteuttaa journalismin tehtävää. Siinä missä muu journalismi selittää muita asioita, datajournalismi selittää dataa. Datajournalismin tekijä etsii uutisia datasta tai datoja yhdistelemällä.

Luvussa 2 esittelin kaksi kontekstia datajournalismille: verkkojournalismi ja osallistava journalismi osana sitä sekä tutkiva ja tietokoneavusteinen journalismi. Seuraavissa kahdessa alaluvussa käsitelen datajournalismia näiden kontekstien kautta. Ensimmäinen niistä liittyy ennen kaikkea esitystapaan ja jälkimmäinen prosessiin.

### 5.3.1 Verkonomaisuus, interaktiivisuus ja visuaalisuus

*Pahimmillaanhan se on sitä, että laitetaan aikajana nettiin ja sit sanotaan, että jeejee tää on todella sensaatiomaista, meillä on tämmönen aikajana. Koska se aikajana on ehkä niinku 1900-luvun alkupuolelta, että mä en oo ihan varma, että onks se edes kunnolla datajournalismia. Se on visualisointia, mut ei siinä oo mitään ihmeellistä, sitä ei kannattais mun mielestä hehkuttaa eikä mainostaa minään käännteentekevänä asiana.*  
(Tuomo Björksten, Yle Uutiset, kotimaan toimittaja.)

Björksten puhui haastattelussa siitä, miten paljon datajournalismia hehkutettiin, kun siitä ylipäättään alettiin puhua Suomessa. Puheissa se oli käännteentekevää ja fantastista, mutta ensimmäiset jutut olivat kömpelöitä interaktiivisia karttoja tai aikajanoja. Kuten luvussa 4 totesin, esimerkiksi maakuntalehtien sovellukset ovat yhä samalla tasolla. Nykyään etenkin Yle ja Helsingin Sanomat tekevät niin sisällöltään kuin visuaaliselta ilmeeltään paljon hienostuneempia juttuja, mutta datajournalismin

rantautuminen Suomeen tuntuu jättäneen lajiin leimansa. Esimerkiksi Björkstenillä tulee aikajanasta miellelyhtymä datajournalismiin, vaikka hän lisääkin perään, ettei tiedä, onko se datajournalismia.

Kun pyysin kahdeksaa haastateltavaani määrittelemään datajournalismin omin sanoin, kuusi heistä liitti sen ennen kaikkea prosessiin, jonka osana data on jollain tavalla. Mäkinen ja Jaakkola puolestaan korostivat kerrontatapaa. Mäkinen kuvaili datajournalismia lyhyesti verkkokerronnaksi, joskin totesi, että se on myös työkaluvalikoima. Jaakkola sanoi, että hänen mielestään datajournalismiin liittyy datan lisäksi selkeästi visuaalisuus ja käytettävyys.

Visuaalisuus, interaktiivisuus ja verkonomaisuus nousivat hieman vahvemmin esiin, kun kysyin datajournalismin eroja muuhun journalismiin. Vain kaksi kahdeksasta ei maininnut näistä mitään. Kolme sanoi, että visuaalisuus tai interaktiivisuus liittyy datajournalismiin usein, mutta ei välttämättä. Jaakkolan mielestä interaktiivisuus on selkeästi erottava tekijä, kun taas Tebest mainitsi verkonomaisuuden. Mäkisen mukaan datajournalismin erottaa muusta journalismista joko datan käyttäminen tai interaktiivinen visualisointi. Hänen mielestään on siis olemassa kahdenlaista datajournalismia.

*Mut tosiaan ku tässä on tää niinkun suomalainen erikoisuus. Siis jos sä kysyisit Simon Rogersilta mitä datajournalismi on ni se varmaan antais sulle suppeemman vastauksen, joka liittyy nimenomaan siihen tiedon hankkimiseen, analysointiin ja puhdistamiseen. Mutta koska meillä on tää niinkun anglosaksisesta perinteestä poikkeava, ainakin Hesarilla on tää anglosaksisesta perinteestä poikkeava tapa määritellä datajournalismi, niin tavallaan sen takia tää mun vastaus on laajempi, se on niinku kaksiosainen kenttä meillä.*  
(Esa Mäkinen, Helsingin Sanomat, datajournalismista vastaava uutispäällikkö.)

Mäkisen mukaan datajournalismi ja interaktiivinen journalismi “löysivät toisensa”, kun asiasta alettiin puhua vähän ennen 2010-lukua. Se johtuu hänen mukaansa siitä, että Suomessa ei ole vahvaa interaktiivisen journalismin perinnettä. Oli luonnollista ja myös käytännön sanelemaa, että se sulautui osaksi datajournalismia. Toisin kuin esimerkiksi The New York Timesilla, Helsingin Sanomilla ei ole Mäkisen mukaan varaa pitää erikseen datadeskiä ja interaktiivista deskiä.

Vaikka interaktiivisuus ei olisi välttämätöntä datajournalismissa, ne näyttävät tosiaan kulkevan usein rinta rinnan. Pyysin haastateltavia valitsemaan mielestään parhaan datajournalistisen jutun, jota on ollut tekemässä. Vain Hänninen valitsi jutun, jossa oli pelkästään tekstiä eikä mitään interaktiivista elementtiä. Kun kysyin häneltä, eroako juttu jotenkin normaalista datajournalistisesta jutusta, hän sanoi sen olevan ilmaisuanniltaan normaalia köyhempi. Lähes kaikki muut sanoivat oman

esimerkkijuttunsa olevan normaalia datajournalismia tai eroavan ainoastaan siinä, että ne ovat normaalia laajempia ja isotoisempia. Jaakkola ei osannut vastata kysymykseen, koska ei osannut määrittellä, millainen on normaali datajournalistinen juttu.

Kaikki haastateltavat olivat yhtä mieltä siitä, että datajournalismia tehdään ennen kaikkea verkkoon. Sitä voi tehdä muihinkin välineisiin, mutta parhaiten se taipuu verkkoon. Datajournalismi ja verkonomainen kerronta tukevat toisiaan. Ottamatta kantaa siihen, onko vaikkapa interaktiivinen kartta datajournalismia, se on usein luonnollinen tapa esittää maantieteellisesti jaettuja tilastoja. Datan hankkimis- ja analysointitekniikat kehittyvät samaan aikaan kuin visualisointitekniikat.

### 5.3.2 Tutkiva journalismi

*-- on hyvin helppo puhua datajournalismista vaan tommosina järjestettävänä taulukoina tai verkkokarttoina, jotka ei sinänsä, mä ainakin koen että se datajournalismi vaatis vähän enemmän viel. Tai jotenkin mä haluaisin kokee, että datajournalismi on semmosta tutkivampaa, tutkitaan jotain, yhdistetään aineistoja ja analysoidaan niiden perusteella jotain eikä vaan esitetä jo olemassa olevaa tietoa.*

(Teemo Tebest, Yle Uutiset, Plus-deskin toimittaja.)

Tutkiva journalismi nousi haastatteluissa esiin terminä melko harvoin. Tebest mainitsi useamman kerran kaipaavansa datajournalismiin enemmän tutkivaa otetta, ja Koivuranta sanoi näkevänsä datajournalismin osana tutkivaa journalismia. Muut puhuivat siitä enintään ohimennen, osa ei ollenkaan.

Kuitenkin haastatteluista paistoi läpi tutkivan journalismin ihanne. Luvussa 2.2.3 lainasin Heikki Kuutin tutkivan journalismin kriteerejä, jotka ovat 1) omaan tutkimukseen perustuminen, 2) aiheen yhteiskunnallinen merkittävyys ja 3) salatun tai kiusallisen tiedon käsitteleminen. Etenkin kaksi ensimmäistä kriteeriä nousivat hyvin usein esiin, kun kysyin haastateltavilta hyvän datajournalismin piirteitä. Samoin ne nousivat esille, kun kysyin, miksi haastateltava oli halunnut valita esimerkkitutuksi juuri valitsemansa jutun tai miksi se oli hänen mielestään hyvä.

Tutkiva journalismi ja datajournalismi siis perustuvat omaan tutkimukseen. Haastatteluista käy ilmi, että datajournalismissa oman tutkimuksen tekeminen eli datan kerääminen ja analysointi on osa jutun syntyprosessia. Kuten luvussa 5.2.1 kerroin, kaikki haastateltavat mainitsivat datapohjaisuuden tekijänä, joka erottaa datajournalismin muusta journalismista. Myös datan analysoiminen on tärkeää.

Epäselväksi jäi kuitenkin se, minkälainen, miten hankittu ja kuinka iso datan pitäisi olla. Joidenkin mielestä datan pitää olla niin iso, että sen käsitteleminen ja analysoiminen manuaalisesti ei ole järkevää, vaan tarvitaan tekniikka apuun. Osan mielestä datan pitää olla jollain tavalla itse koostettua. Osa sen sijaan laskee datajournalismiksi esimerkiksi jutun, joka pohjautuu tilastokeskukselta saatuihin tilastoihin.

*Et tosi paljon se on, no, siks mä tykkään siitä atk-toimittajanimikkeestä, koska datajournalismi, se on automaattista tietojenkäsittelyä tosi usein myös, et sä vaan automatisoit niitä asioita mitä sä oot aikaisemmin tehny journalismissa. Nyt tietokone tekee ne sun puolesta.*  
(Teemo Tebest, Yle Uutiset, Plus-deskin toimittaja.)

## 5.4 Kohti datajournalismin merkitystä

Ennen kuin menen datajournalismin merkitykseen, on syytä käydä lyhyesti läpi, millaisena haastateltavat näkevät datajournalismin aseman Suomessa. Nämä kaksi asiaa liittyvät kiinteästi toisiinsa: datajournalismin asema määrittää sitä, kuinka paljon sillä voi olla merkitystä.

Kun kysyin haastateltavilta, tehdäänkö Suomessa tarpeeksi datajournalismia, kaikki haastateltavat antoivat ymmärtää, että enemmänkin voisi tehdä. Osa höysti vastausta toteamalla, että datajournalismi ei ole itseisarvo. Enemmän datajournalismi on keino tehdä hyviä juttuja, ja niille on haastateltavien mielestä tilausta.

Datajournalismin tämänhetkistä asemaa kuvailtiin monin tavoin: se on lapsenkengissä, siitä puhutaan enemmän kuin tehdään, se on nouseva trendi. Muutama puhui Yleisradiosta ja Helsingin Sanomista edelläkävijöinä. Pari mainitsi, että koulutusta järjestetään entistä enemmän. Esiin nousi myös kehitystä jarruttavia tekijöistä: kolme mainitsi pulan hyvistä tekijöistä, kaksi yt-neuvottelut ja kaksi pomot, jotka eivät anna resursseja tai ymmärrä datajournalismista mitään.

*Viime kesänä [2013] sanoin uutispäällikölle, että tulee interaktiivinen verkkokartta tähän nettiversioon yhteyteen ja siitä pitää ohjeistaa toimituksia, ni sit uutispäällikkö ei tajunnut mikä on interaktiivinen verkkokartta, et hän luuli että se on joku printtikartta. Että puhutaan tälläkin tasolla asioista vielä.*

(Esa Koivuranta, STT-Lehtikuva, kotimaan toimittaja.)

Haastatteluista kävi kuitenkin ilmi, että haastateltavat uskovat datajournalismin kehitykseen ja aseman vahvistumiseen. Haastateltavat liittyvät datajournalismin kehityksen muun muassa (avoimen) datan lisääntymiseen ja tekniikan kehittymiseen. Tulevasta tosin puhuttiin varovaisin sanankääntein. Moni



käytti toivoo-verbiä, kun puhui datajournalismin positiivisesta tulevaisuudesta. Varmistin monelta haastateltavalta erikseen, että he uskovat siihen mitä toivovat.

Selvitin haastateltavien näkemyksiä datajournalismin merkityksestä kahdelta kannalta: mitä annettavaa datajournalismilla on ensinnäkin journalismille ja toisekseen yleisölle ja yhteiskunnalle. Tiivistin vastaukset kuvioon 8:

	<b>Datajournalismin anti journalismille</b>	<b>Datajournalismin anti yleisölle ja yhteiskunnalle</b>
<b>Esa Koivuranta</b>	itsenäinen tiedonhankinta, rahan lähde	uutiset, yhteiskunnallisten epäkohtien paljastaminen, tutkiva journalismi
<b>Wolf Wikgren</b>	datan ja uuden tiedon esittäminen, verkkonatiivius	datan ja uuden tiedon esittäminen
<b>Teemo Tebest</b>	tutkiva asenne, omat uutiset, haastaminen, selittäminen, taustoitus	asioiden muuttaminen, uudet näkökulmat, asioiden suhteuttaminen
<b>Johan Jaakkola</b>	erilainen kerronta, uudenlaista tietoa, datan hyödyntäminen	asioiden muuttaminen, uudet näkökulmat, datan ja uuden tiedon esittäminen
<b>Tuomo Björksten</b>	elämyksellisyys, viihdyttävyys, yleisön houkuttelu	uusien ihmisten tavoittaminen, oivalluksia ja tietoa yhteiskunnasta
<b>Esa Mäkinen</b>	skuupit, visualisoinnit, ymmärrettävyys, uudet keinot uutisten löytämiseen	skuupit yhteiskunnallisesti vaikuttavia; pelit, visat ja koneet yleisölle kiinnostavia
<b>Jyri Hänninen</b>	uudet näkemykset, aiempaa monipuolisempi ilmaisutapa, verkkonatiivius	yhteiskunnallinen keskustelu, asioiden muuttaminen
<b>Tuomas Peltomäki</b>	uutisaiheita, selittävyys, uudenlaista tietoa	laajojen trendien paljastaminen ja todistaminen

*Kuvio 8: Tiivistettyjä poimintoja haastatteluista datajournalismin merkitykseen liittyen.*

Käyn seuraavassa kahdessa alaluvussa läpi haastateltavien pohdintoja datajournalismin merkitykseen liittyen. Lisäksi pureudun heidän näkemyksiinsä siitä, miten datajournalismi muuttuu tulevaisuudessa. Muutokset liittyvät erityisesti siihen, millainen merkitys datajournalismilla on alan sisällä. Siksi käsittelen muutoksia samassa yhteydessä, kun käsittelen datajournalismin merkitystä journalismille. Raja ei kuitenkaan ole tarkka, vaan muutoksilla voi olla merkitystä myös yleisölle ja yhteiskunnalle.

#### 5.4.1 Merkitys journalismille

*Pystytään hyväksikäyttämään semmosia aineistoja, mitä ei oo välttämättä aiemmin pystytyt niin hyvin. Ja pystytään tosiaan luomaan semmosta verkkonatiivia journalismia.*  
(Wolf Wikgren, STT-Lehtikuva, datajournalismiharjoittelija.)

Wikgren kiteyttää aika hyvin ne kaksi asiaa, jotka nousivat vahvimmin esiin, kun kysyin datajournalismin merkityksestä journalismille. Ensinnäkin aineistosta eli datasta löytyy uutisia, skuppeja tai uutta tietoa – haastateltavat ilmaisivat asian hieman eri tavoin – jota ei aiemmin ole saatu kaivettua esiin. Toisekseen esitystapa voi olla verkossa muutakin kuin näköispainos printtilehdestä tai tv-insertti. Senkin haastateltavat ilmaisevat eri tavoin, esimerkiksi sanoilla verkkonatiivius, elämyksellisyys ja visualisoinnit.

Osa kuvaili tarkemmin, millaisia datasta kaivetut uutiset ovat. Ne voivat olla maailmaa selittäviä, ja analyysiä tehdessä voi löytyä uusia näkemyksiä. Tebest arvelee, että datajournalismi voi tuoda tutkivaa asennetta journalismiin: Uutisia ei vain kerrota, vaan niitä tehdään itse. Uutisissa haastetaan, selitetään ja taustoitetaan.

Moni haastateltava näkee verkkonatiiviuden yleisön palveluna. Visualisoinnit ja elämyksellisyys voivat tehdä asioista helposti ymmärrettäviä. Björksten sanoi, että datajournalismi voi tehdä haudanvakavastakin uutisesta elämyksellisen, mikä voi houkutella lukijoita. Parinkin haastateltavan vastauksiin sisältyi ajatus siitä, että datajournalismi tarjoaa sisältöjä, joista yleisö haluaa maksaa. Datajournalismin nähdään siis tuovan rahaa aikana, jolloin journalismin ansaintalogiikka on hukassa ja ihmiset ovat tottuneet ilmaiseen verkkojournalismiin.

Oletettavasti datajournalismin merkitys kasvaa sen myötä, kun se kehittyy. Kaksi haastateltavaa sanoi, että etenkin nuoret toimittajat edistävät datajournalismin yleistymistä.

*Ja nyt koko ajan tulee sitte nuorempia ja nuorempia tyyppejä alalle, jotka on tottunut tähän käsitteeseen ja heidän tämmösessä toimittajan taitopaketissa on datajournalismia, nää käsitteet jollain tavalla sisäänrakennettuna jo, niin totta kai se varmasti lisää datajournalismia ja se tekee siitä yhä yleisempää.*

(Johan Jaakkola, Yle Uutiset, taloustoimittaja.)

On huomionarvoista, että kaikki haastateltavat ovat melko nuoria. Vanhin oli haastatteluhetkellä 36-vuotias, nuorin vasta 24-vuotias. Ehkä nuorten on helpompi omaksua uusia tekemisen tapoja, tai ehkä heillä on enemmän tilaisuuksia tai intoa opetella uusia taitoja. Kaikki haastateltavat ovat myös miehiä. Tietotekniikka onkin perinteisesti miesvaltainen ala.

Sekä tekijöiden ikä että sukupuoli vaikuttanee osaltaan siihen, millaista datajournalismia tehdään ja millaiseksi se kehittyy. Ikä ja sukupuoli voivat vaikuttaa esimerkiksi aihevalintoihin, näkökulmiin ja siihen, missä datajournalismia tehdään.

Haastateltavat arvelevat datajournalismin muuttuvan monella tavalla tulevaisuudessa. Lähes kaikki sanoivat, että toimittajien taidot kehittyvät tai tekijöitä tulee lisää. Moni ajattelee tekniikan, prosessien ja ilmaisutapojen kehittyvän. Ylipäättään he tuntuvat uskovan siihen, että datajournalismi yleistyy ja arkipäiväistyy. Samalla juttuaiheiden kirjo voi laajentua tai ideat jalostua.

Pisimmälle vietyjä näkemyksiä datajournalismin muuttumisesta esitti Mäkinen. Se on luontevaa, sillä hänellä on haastateltavista eniten kokemusta datajournalismista, hän toimii uutispäällikön tehtävissä, ja hän on ollut alusta saakka vastuussa Helsingin Sanomien datadeskin johtamisesta ja kehittamisestä. Mäkinen puhui työkaluistumisesta, joka on hänen mukaansa parhaillaan käynnissä. Hän ennustaa, että muutaman vuoden aikana työprosessit hahmottuvat selkeämmiksi, ja jokaiselle prosessin osalle hahmottuu selkeämmin omat työkalut. Sen jälkeen mediatalojen johtajien on helpompi käsittää, mitä datajournalismin tekemiseen vaaditaan.

*-- ku työprosessit on hahmottunu, ni jokaiselle työprosessille tulee selkeemmin omia työkaluja, tulee tavallaan, se sen logiikka mikä siel toimii taustalla on se niinku mediapomon näkemys maailmasta, että mitä ohjelmia mun pitää ostaa, mitä koulutuksia mun pitää järjestää, jotta mä saan noi tyypit tekemään datajournalismia.*

(Esa Mäkinen, Helsingin Sanomat, datajournalismista vastaava uutispäällikkö.)

Toinen Mäkisen mainitsema seikka on se, että yhä useampi mediatalo alkaa tehdä datajournalismia. Tällä hetkellä Suomen ykkönen datajournalismissa on Mäkisen näkemyksen mukaan Helsingin Sanomat. Jotta se pystyy säilyttämään asemansa, sen täytyy tulevaisuudessa tehdä entistä enemmän ja paremmin.

Kolmanneksi Mäkinen puhui datajournalismin kaupallistamisesta, joka tuo rahan myötä uskottavuutta ja lisäksi muuttaa sisältöjä. Mäkinen kertoi, että Helsingin Sanomat on alkanut myydä yksinoikeudella mainospaikkoja helposti ennakoitaviin datajournalistisiin tuotoksiin. Tällainen voisi olla interaktiivinen verkkototeutus esimerkiksi jalkapallon MM-kilpailuihin tai pääsiäisen resepteihin liittyen. Koska uutisia ei voi kovin hyvin ennakoida, sisällöt ottavat askeleen aikakauslehtien suuntaan.

#### 5.4.2 Merkitys yleisölle ja yhteiskunnalle

Kiinnostavasti ainoastaan Hänninen nosti esiin, että datajournalismin merkitys yhteiskunnalle on vielä vähäinen ja että se voisi ottaa journalismin kentällä suuremman roolin. Muut tuntuivat ottavan annettuna, että datajournalismilla on merkitystä. Osa totesi ääneen, että merkitys on suuri. En tosin pyytänyt haastateltavia suoraan arvioimaan, kuinka suuri datajournalismin merkitys on, vaan millainen sen merkitys on.

*No jos aatellaan, että tää tämmönen vanha sivistystehtävä ei oo medialta vielä ihan kokonaan kadonnut, niin tota kyllähän se siis parhaimmillaan antaa niille ihmisille oivalluksia ja tietoa yhteiskunnasta, ja semmonen nykyaikaa vaivaava välinpitämättömyys sitten parhaimmillaan ehkä häviäis tai vähenis tai sillä tavalla.*

(Tuomo Björkstén, Yle Uutiset, kotimaan toimittaja.)

Kun kysyin datajournalismin merkitystä yleisölle ja yhteiskunnalle, esiin nousi jälleen vahvasti tutkivan journalismin ihanne. Se tuli esiin jollain tavalla jokaisessa haastattelussa. Haastateltavat puhuivat esimerkiksi omien uutisten tekemisestä ja asioiden muuttamisesta journalismin avulla. Voi olla, että myös yhteiskunta-sanana mainitseminen kysymyksessä ohjasi haastateltavien ajatuksia yhteiskunnallisten epäkohtien paljastamiseen ja yhteiskunnalliseen keskusteluun. Toisaalta tutkiva journalismi oli muutenkin vahvasti läsnä monessa haastattelussa.

Tuomas Peltomäki esitti konkreettisen ajatuksen siitä, mihin datajournalismi pystyy. Hänen mukaansa datajournalismi voi paljastaa ja todistaa laajoja trendejä, joista ei ole vastuussa kukaan yksittäinen ihminen.

*Ja sen se voi antaa yhteiskunnalle, se pystyy näyttämään semmosia isoja aiheita, joista kukaan yksittäinen tyyppi ei oo vastuussa. Koska jos se ois joku yksittäinen tyyppi vastuussa, ni kaikista järkevin ja tehokkain journalistinen tapa ois mennä kysymään siltä, että näinkö. Mut tosi moni asia maailmassa, tosi moni tärkeä asia maailmassa on semmonen, että ei siitä kukaan yksittäinen tyyppi vastaa millään tavalla. Niin ni datajournalismi pystyy näyttään tän. (Tuomas Peltomäki, Helsingin Sanomat, hs.fin toimitussihteeri.)*

Mäkinen näkee datajournalismilla kahdenlaista merkitystä: skuupit ovat yhteiskunnallisesti vaikuttavia, ja pelit, visat ja koneet ovat yleisölle kiinnostavia ja tärkeitä, koska niiden kautta yhä useampi kuluttaa journalismia. Myös Björkstén puhui uusien ihmisten tavoittamisesta uusien kerrontatapojen avulla.

## 5.5 Yhteenveto

Analyysistä nousi vahvasti esiin ne kaksi asiaa, joiden kautta kontekstoin datajournalismia luvussa 2: tutkiva journalismi ja verkkojournalismi. Haastattelujen perusteella nämä kaksi asiaa liittyvät niin datajournalismin määritelmään kuin merkitykseen. Muita, yllättäviä konteksteja ei tullut ilmi.

Tutkiva journalismi kytkeytyy datajournalististen juttujen tekoprosessiin. Kaikki haastateltavat ovat sitä mieltä, että datajournalismia määrittää datapohjaisuus, siis uuden tiedon etsiminen datasta. Tärkeää on myös datan analysoiminen. Se tarkoittaa, että datajournalismin tekijät tuottavat omia uutisia. Jäi kuitenkin epäselväksi, minkälainen, miten hankittu ja kuinka iso datan pitäisi olla. En ehkä tajunnut kysyä siitä tarpeeksi spesifisti.

Tutkiva journalismi tuli esiin myös siinä, että haastateltavat painottivat datajournalismin yhteiskunnallista merkitystä. Useimmille heistä hyvä datajournalismi tarkoittaa merkittävää uutista. Siihen vaikuttanee osaltaan se, että kaikki haastateltava työskentelevät uutisorganisaatioissa. Heillä tuntuu olevan vahva halu herättää yhteiskunnallista keskustelua ja jopa muuttaa maailmaa. Se ilmeni esimerkiksi, kun pyysin haastateltavia pohtimaan datajournalismin merkitystä yleisölle ja yhteiskunnalle sekä omien parhaiden juttujen valinnassa.

Siinä missä tutkiva journalismi kytkeytyy prosessiin, verkkojournalismi liittyy esitystapaan. Kuusi kahdeksasta haastateltavasta sanoi, että visuaalisuus, interaktiivisuus tai verkonomaisuus liittyy datajournalismiin usein tai aina. Näyttää siltä, että datajournalismi ja verkonomainen kerronta kehittyvät käsi kädessä. Joidenkin mielestä verkonomainen kerronta on datajournalismia. Esimerkiksi

Esa Mäkinen näkee, että on olemassa kahdenlaista datajournalismia: datajournalismin erottaa muusta journalismista joko datan käyttäminen tai interaktiivinen visualisointi. Toisinaan nämä kaksi asiaa yhdistyvät.

Kaikki haastateltavat ajattelevat, että datajournalismia tehdään ensisijaisesti verkkoon. Verkkoon liittyy myös se, mitä annettavaa datajournalismilla on – uutisten lisäksi – journalismille. Datajournalismi tuo journalismiin elämyksellisyyttä, visuaalisuutta, interaktiivisuutta. Tällaiset elementit puolestaan houkuttelevat yleisöä ja siten voivat tuoda rahaa mediataloille.

On huomionarvoista, että kuusi haastateltavaa kahdeksasta liitti datajournalismin määritelmän ensisijaisesti prosessiin ja vain kaksi esitystapaan, siis verkonomaiseen kerrontaan. Verkonomaisuus nousi hieman vahvemmin esiin, kun kysyin datajournalismin eroja muuhun journalismiin. Harva piti sitä kuitenkaan välttämättömytenä. Enemmän moni sanoi, että datajournalismiin voi liittyä interaktiivisia visualisointeja.

Analyysin pohjalta vaikuttaa siltä, että datajournalismi on nouseva trendi. Datajournalismi on vasta kehittymässä, eivätkä sen määritelmä ja merkitys ole täysin selviä edes sen tekijöille. Haastateltavat ovat yksimielisiä siitä, että datajournalismia voisi tehdä enemmänkin. He myös uskovat, että se yleistyy ja muuttuu yleistyessään. Työkalut kehittyvät, työprosessit selkiytyvät, tekijöitä tulee ehkä lisää.

On kiinnostavaa, että harva haastateltava koki itse olevansa datajournalisti. Ehkä ei olekaan olemassa datajournalisteja, vaan on olemassa toimittajia, koodaajia ja graafikoita, jotka tekevät datajournalismia. Usein datajournalismia tehdään joka tapauksessa yhteistyönä.

## 6 Pohdinta

### 6.1 Kohti johtopäätöksiä

Tutkimukseni tarkoitus oli selvittää, miten suomalaiset datajournalismin tekijät määrittelevät datajournalismin ja mitä merkitystä he näkevät sillä olevan journalismille, yleisölle ja yhteiskunnalle. Esittelin keväällä 2014 tekemiäni teemahaastattelujen pohjalta syntyneitä havaintoja luvussa 5. Lisäksi halusin kartoittaa, mitä Suomessa on tehty datajournalismin saralla. Kirjoitin syksyllä 2014 tekemiäni kartoituksen tuloksista luvun 4.

Tässä luvussa syvennän tekemiäni havaintoja. Aluksi esittelen tärkeimmät tutkimustulokset ja vertaan niitä aiempaan tutkimukseen, jota esittelin luvussa 2. Lisäksi laajennan tuloksia vielä hieman ja pyrin tekemään oman määritelmän datajournalismille. Sen jälkeen tarkastelen tutkimusta ja sen merkitystä kriittisellä otteella. Tähän sisältyy myös tutkimusmenetelmän arviointi. Lopuksi esittelen muutaman jatkotutkimusidean.

### 6.2 Tulokset ja aiempi tutkimus

Ehkä tärkein havaintoni on se, että kahdeksasta haastateltavasta kuusi liitti datajournalismin määritelmän ensisijaisesti prosessiin ja vain kaksi esitystapaan. Toisin sanoen datajournalismia määrittää ennen kaikkea datan käyttö prosessissa eikä esitystapa. Prosessi ja esitystapa kytkeytyvät vahvasti myös niihin kahteen kontekstiin, jotka ennakko-oletukseni mukaan nousivat analyysissä esiin: tutkivaan journalismiin ja verkkojournalismiin. Tutkiva journalismi yhdistyy dataan prosessin osana. Esitystapa puolestaan liittyy erityisesti verkonomaiseen kerrontaan.

Toinen tärkeä havainto on se, että datajournalismin merkitys kytkeytyy ennen kaikkea datasta löytyvään uuteen tietoon, monesti siis tutkivaan journalismiin. Toisaalta merkitys kytkeytyy myös verkonomaisen kerronnan kehittämiseen ja sitä kautta uuden yleisön ja rahan houkutteluun. Näyttää siltä, että merkitys kumpuaa datajournalismia eniten määrittävistä tekijöistä.

Analyysi vahvisti myös käsitystä siitä, että datajournalismi on vasta hakemassa muotoaan Suomessa. Se on nouseva trendi, mutta sen tekeminen on tällä hetkellä ennen kaikkea Yleisradion ja Helsingin

Sanomien harteilla. Tämän osoittivat niin tekemäni kartoitus datajournalismin tilasta syksyllä 2014 kuin teemahaastattelut. Muissa mediataloissa tehdään melko vähän ja yksinkertaista, jos ollenkaan – joskin STT-Lehtikuva panostaa tällä hetkellä datajournalismiin tai ainakin visuaalisiin elementteihin.

Suomessa ei tietääkseni ole aiemmin tehty samanlaista kartoitusta datajournalismin tilasta. Myöskään datajournalismin määritelmää ja merkitystä ei ole selvitetty teemahaastatteluilla. Niinpä peilaan tuloksia ulkomaisiin tutkimuksiin.

Ainakin Belgiassa, Norjassa, Ruotsissa ja Yhdysvalloissa on tehty samankaltaiset, mutta laajemmat tutkimukset (De Maeyer et al. 2014; Karlsen & Stavelin 2014; Appelgren & Nygren 2014; Fink & Anderson 2014). Esittelin näiden tutkimusten tuloksia luvussa 2.2.5. Niissä nousi esiin samankaltaisia seikkoja kuin omassa analyysissäni: Datajournalismista on monenlaisia käsityksiä, ja sen kehitys on eri paikoissa eri vaiheissa. Ajan ja taitojen puute on ongelma. Kaikki eivät näe datajournalismin eroavan journalismista, onhan aiemminkin tehty journalismia tilastojen perusteella. Esimerkiksi Ruotsissa datajournalismi on vielä lapsenkengissä.

On kiinnostavaa, että kahdessa maassa datajournalismin nähtiin jakautuvan kahtia. Belgiassa tehdyssä tutkimuksessa erottui tavallinen ja perinpohjainen datajournalismi. Ruotsissa puolestaan datajournalismin tekijöistä erottui kaksi ryhmää: leikkisät uusien työtapojen etsijät ja tietokoneavusteisen journalismin perinteen jatkajat, jotka tekevät yksinkertaisesti datan pohjalta journalismia. Omassa tutkimuksessani erityisesti Esa Mäkinen – kaikkein kokenein datajournalismin tekijä – jakoi datajournalismin interaktiivisiin visualisointeihin ja datapohjaisiin uutisiin. Eri tavoissa jakaa datajournalismia voi nähdä yhtäläisyyksiä.

Kun perehdyin datajournalismin käsitteeseen luvussa 2.2.5, totesin kolmen asian tekevän datajournalismista monimutkaista: 1) tuoreen käsitteen alle lasketaan sekalainen joukko asioita, 2) myös muut kuin journalistit tekevät datajournalismia ja 3) niin data kuin journalismi ovat problemaattisia käsitteitä. Pohdin seuraavaksi hieman näitä väitteitä tutkimukseni valossa.

Ensinnäkin tutkimus vahvisti käsitystä siitä, että datajournalismiksi luetaan sekalainen joukko asioita. Tämä nousi erityisesti esiin tausta-aineistosta (ks. luku 4). Haastattelin siihen mediatalojen edustajia, joilla ei välttämättä ollut kovin järeää kokemusta datajournalismista. Datajournalismin määrittelemisen tuotti heille vaikeuksia, mutta ei se ollut helppoa myöskään ensisijaisen aineiston haastateltaville. Useampi haastateltava pohti ääneen, onko esimerkiksi aikajana tai interaktiivinen



kartta datajournalismia. Vaikka he olisivat päätyneet siihen, että ei ole, yhtä kaikki heille tuli datajournalismista puhuttaessa mieleen aikajana tai interaktiivinen kartta. Visualisoinnit ovat siis jättäneet leimansa käsitteeseen.

Kuitenkin datapohjaisuus ja datan analyysi ovat ainoat asiat, jotka nousivat jokaisessa teemahaastattelussa esiin datajournalismia määritellessä. Muuten käsitykset eroavat toisistaan, ja niihin liittyy lieventäviä sanoja kuten “usein” tai “yleensä”.

Toiseksi on totta, että myös muut kuin journalistit tekevät datajournalismia. Etenkin tekniset osaajat ovat välttämättömiä datajournalismin tekemisessä. Analyysin perusteella näyttää kuitenkin siltä, että siitä ei kannata tehdä ongelmaa. Kuten Esa Mäkinen huomautti luvussa 5.2, koodaajan on helpompi opetella journalismia kuin journalistin koodaamista. Moni muukin on opetellut tekemään journalismia ilman akateemista loppututkintoa.

On myös olennaista huomata, että datajournalismia tehdään usein yhteistyönä. Tällöin journalististen päätösten tekeminen ei jää datajournalismin tekemistä aloittelevan koodaajan harteille. Analyysistä nousikin vahvasti esiin se, että datajournalismin tekemiseen ei tarvita erillisiä datajournalisteja. Datajournalismia syntyy toimittajien, koodaajien ja graafikoiden yhteistyön tuloksena. Tämä on tärkeää pitää mielessä, jotta kynnyks datajournalismin tekemiseen ei nouse valtavaksi. Julkisuudessa puhutaan toisinaan kaiken osaavista datavelhoista (ks. luku 1.1), mikä on omiaan tappamaan orastavan kiinnostuksen datajournalismiin. Aiemmin esimerkiksi Royal (2010) on kirjoittanut yhteistyön merkityksestä (ks. luku 2.2.5).

Kolmanneksi analyysi todensi sen, että erityisesti data on problemaattinen käsite. Minkälainen, kuinka suuri ja miten hankittu datan pitää olla, jotta sen pohjalta tehdystä journalismista voidaan puhua datajournalismina? Tähän kysymykseen en pysty vastaamaan aineistoni perusteella. En ehkä kysynyt tarpeeksi tarkkoja kysymyksiä asiaan liittyen. Toisaalta asiaa sivunneista keskusteluista kävi ilmi, että haastateltavilla ei myöskään ole asiasta selkeää näkemystä. Se kävi kuitenkin selväksi niin haastatteluista kuin aiemmasta tutkimuksesta (ks. luku 2.2.4), että datan määrä ja hankkimiskeinot lisääntyvät koko ajan. Kaikki Poikolan mainitsemat datalähteet (ks. luku 2.2.4) eli avoin data, tietopyynnöt, ruudun raavinta ja joukkoistaminen nousivat jollain tavalla esiin haastatteluissa. Vahvimmin esillä oli avoin data, heikoiten joukkoistaminen.

Myös journalismi on problemaattinen käsite, ja sen tarkka määrittelemine on internet-aikana yhä

vaikeampaa (ks. luku 2.2.1). Siihen kuitenkin tavataan liittää joukko määritteitä, jotka ovat yleisesti tiedossa. Kaikilla on jonkinlainen käsitys journalismista ja erityisesti uutisista. Kaikki haastateltavat työskentelevät uutistoimituksissa, ja heidän puheistaan nousi esiin kirkkaana mielessä olevat uutiskriteerit ja tutkivan journalismin ihanne. Se vaikuttanee osaltaan heidän näkemyksiinsä.

Kuten jo aiemmin totesin, haastatteluissa nousi esiin verkkojournalismin ja tutkivan journalismin kontekstit. Bradshawn ja Rohumaan (2011) luettelemista kahdeksasta verkkojournalismin piirteestä erityisesti interaktiivisuus ja internet juttujen alustana painottuivat haastatteluissa. Datajournalismia tehdään ensisijaisesti verkkoon, ja siihen liittyy usein interaktiivisia visualisointeja. Kukaan ei maininnut termiä osallistava journalismi (ks. luku 2.2.2), mutta interaktiivisuus on osallistavaa. Tutkimuksen perusteella datajournalismin merkitys perustuu osittain ihmisten haluun osallistua: uudenlainen verkkokerronta houkuttelee yleisöä ja siten rahaa.

Ennen kaikkea merkitys perustuu kuitenkin datasta löytyvään tietoon, monesti siis tutkivaan journalismiin. Haastatteluissa nousi vahvasti esiin kaksi Kuutin (2002) luettelemista tutkivan journalismin kriteereistä: perustuminen omaan tutkimukseen ja aiheen yhteiskunnallinen merkittävyys. Kolmas kriteeri on salatun tai kiusallisen tiedon käsitteleminen, joka sekin oli luettavissa joistain vastauksista.

### **6.2.1 Oma määritelmä datajournalismille**

Kerroin johdannossa, että ymmärrän datajournalismin alustavasti isojen datamassojen pohjalta tehdyksi journalismiksi, jota tehdään yleensä verkkoon. Huomautin myös, että datajournalismiin liittyy usein – mutta ei välttämättä – hienoja visualisointeja ja interaktiivisuutta. Lisäksi totesin, että parhaimmillaan datajournalismi on tutkivaa journalismia, sillä erilaisia dataja yhdistämällä ja analysoimalla voidaan synnyttää uutta tietoa.

Tutkimukseni muutti tätä alustavaa määritelmää monin tavoin. Ensinnäkin, data voi näkyä prosessin eri vaiheissa eikä ole pelkästään jutun lähde. Toiseksi, en uskalla tutkimukseni perusteella ottaa kantaa datan kokoon. Kolmanneksi, verkko on oleellinen osa datajournalismia. Datajournalismia tehdään verkkoon tai vähintäänkin verkko edellä. Neljänneksi, en koe tarpeelliseksi tuoda määritelmässä esiin tutkivan journalismin termiä. Tutkiva journalismi tulee esiin siinä, että datajournalismi tuottaa uutta tietoa. En kuitenkaan näe estettä sille, etteikö datajournalismia voisi tehdä myös kevyistä, ei-

yhteiskunnallisista aiheista. Tässä teen tietoisesti valinnan tehdä hieman eroa haastateltaviin, joiden näkemyksissä korostui tutkivan journalismin ja yhteiskunnallisesti tärkeiden aiheiden ihanne. On huomioitava, että kaikki haastateltavat työskentelevät uutistoimituksissa, mikä ohjanee heidän ajatteluaan.

Ehdotukseni datajournalismin määritelmäksi on seuraava: *Datajournalismi on journalistinen prosessi, jossa tuotetaan datan avulla uutta tietoa verkkoon tai verkko edellä ja johon liittyy usein visuaalisia ja interaktiivisia elementtejä.*

Määritelmä jää väistämättä väljäksi. Tarvittaisiin lisätutkimusta, jotta esimerkiksi datan roolista tai koosta pystyisi sanomaan mitään täsmällisempää. Samalla määritelmä kuitenkin rajaa ulkopuolelle asioita, joita on tavattu nimittää datajournalismiksi. Tämän määritelmän mukaan esimerkiksi aikajana tai yksinkertainen testi ei sellaisenaan ole datajournalismia. Toisin kuin Esa Mäkinen, teen siis eron interaktiivisen kerronnan ja datan avulla tuotetun uuden tiedon välille, vaikka ne voivat myös yhdistyä. Voi tietysti kysyä, onko pelkkä aikajana lainkaan journalismia.

Kuitenkin datajournalismi on selkeästi 2000-luvun journalismia. Se on tietokoneavusteisen journalismin päivitetty versio, jota tehdään kehittyneemmillä välineillä ja ensisijaisesti verkkoon. Howardin (2014) mukaan tietokoneavusteisessa journalismissa on perinteisesti keskitytty datan analysoimiseen, kun taas datajournalismi sisältää myös datan julkaisemisen, kierrättämisen ja käytettävyyden. Mediatilat julkaisevatkin toisinaan jutussa käytetyn datan sellaisenaan, ja käytettävyys voi näkyä esimerkiksi näyttävänä interaktiivisena visualisointina. Toisinaan datajournalistinen juttu toimii parhaiten pelkkänä tekstinä. Esitystavat ovat siis moninaiset.

### **6.3 Tutkimuksen kriittinen tarkastelu ja merkitys**

Tutkimukseni tarkoitus oli tuottaa uutta tietoa eikä testata aiempaa teoriaa (ks. luku 3.5). Mielestäni onnistuin kertomaan uutta suomalaisen datajournalismin tilasta ja merkityksestä. Yksi tärkeä tulos oli oman määritelmän luominen datajournalismille.

Vaikka en testannut teoriaa, teoria ohjasi analyysiä. Esittelin tutkimukseni konteksteina verkkojournalismin ja tutkivan journalismin (ks. luvut 2.2.2 ja 2.2.3). Valinta oli mielestäni onnistunut, sillä nämä kontekstit nousivat esiin myös teemahaastatteluissa. En usko ohjanneeni haastateltavia

ainakaan vahvasti näiden kontekstien suuntaan, vaan haastattelurungon kysymykset olivat tarkoituksellisesti laajoja ja avoimia (ks. liite).

Avasin omaa taustaani ja tutkimuksen lähtökohtia johdannossa (ks. luku 1.1). Pysin myös perustelevaan ja selventämään valintojani läpi tutkimuksen sekä kirjoittamaan auki ennakkokäsityksiäni datajournalismista. Se on tärkeää, sillä käyttämäni käsitteet, tutkimusasetelma ja menetelmät vaikuttavat tuloksiin (Sarajärvi & Tuomi 2009, 96). Virheiden välttämiseksi annoin niin ensisijaisen aineiston kuin tausta-aineiston haastateltaville mahdollisuuden lukea tutkimuksen ennen sen julkaisua.

Luvussa 3 kerroin tutkimusmenetelmistä ja aineistosta. Keräsin ensisijaisen aineiston teemahaastatteluilla (ks. luvut 3.2 ja 3.3). Valitsin oman perehtyneisyyteni perusteella kolme avainhaastateltavaa, joita pyysin ehdottamaan lisähaastateltavia. Tämä menetelmä osoittautui toimivaksi. En usko, että olisin itsenäisesti osannut valita yhtä sopivia haastateltavia. Toisaalta avainhaastateltavat olivat niin ilmiselviä, että heidän valitseminen ei voinut epäonnistua.

Tausta-aineistoni koostui eri toimitusten edustajien kanssa käydyistä puhelinkeskusteluista sekä muutamista lehtiartikkeleista. Tämän aineiston keräämisessä etenin toisinaan melko sattumanvaraisesti. Saatoin soittaa toimitukseen ja kysyä, kuka voisi olla oikea henkilö kertomaan minulle heidän tekemästään datajournalismista. Luotan siihen, että toimitusten sisällä osataan arvioida tällaiset asiat. Sain mielestäni kasaan melko kattavan aineiston, joskin maakuntalehtiä olisi voinut ajan salliessa ottaa mukaan enemmänkin kuin kaksi.

Jaoin analyysin kuvailevaan osioon ja teemahaastatteluille tehtyyn sisällönanalyysiin. Kuvaileva osio eli suomalaisen datajournalismin tilan kartoitus toimi hyvänä johdatteluna varsinaiseen analyysiin. Kartoitusta varten käymäni puhelinkeskustelut olivat melko vapaamuotoisia. Keskusteluissa korostui se, miten toimitusten edustajat itse käsittävät datajournalismin. Heidän näkemyksensä eivät välttämättä kerro edes edustamiensa medioiden linjasta. Siksi kartoitus ei varmasti ole aukoton kuvaus toimitusten tekemästä datajournalismista.

Myös sisällönanalyysissä korostuvat teemahaastateltavien subjektiiviset näkemykset. Pysin pitämään analyysiä tehdessäni mielessä, että haastateltavat tulevat eri lähtökohdista ja ovat tehneet eri verran datajournalismia. Analyysiin nostamieni sitaattien yhteydessä kerroin haastateltavan nimen ja sen, missä taustaorganisaatiossa ja työtehtävässä tämä on tehnyt datajournalismia, jotta sitaattia voi peilata

haastateltavan kokemukseen. Kirjoitin myös jokaisesta haastateltavasta lyhyen esittelyn (ks. luku 3.3).

Monet haastateltavien vastaukset olivat melko jäsentymättömiä ja höystetty erilaisilla lieventävillä ilmauksilla. Syynä saattoi olla niin vähäinen kokemus kuin se, että haastateltava ei välttämättä ollut pohtinut datajournalismia teoreettisesti. Kuitenkin haastateltavat kykenivät käsittelemään datajournalismia myös omaa kokemusta laajempaan ilmiönä.

Tekemäni haastattelurunko palveli tarkoitustaan enimmäkseen hyvin. Haastateltavan kokemukset datajournalismista ja parhaan datajournalistisen jutun ruotiminen toimivat hyvänä taustana etenkin datajournalismin määritelmälle, mutta niitä olisi voinut tiivistää hieman. Sen sijaan datajournalismin merkitykselle olisi voinut antaa hieman enemmän tilaa. Lisäksi huomasin analyysivaiheessa, että olisin voinut kysyä tarkemmin datan roolista juttujen tekoprosessissa. Samoin olisin voinut kysyä kaikilta suoraan, kokevatko he itse olevansa datajournalisteja.

Jätin paljon haastateltavien harkinnan varaan. En esimerkiksi antanut erityisiä kriteerejä parhaan oman jutun valintaan, vaan päin vastoin pyysin heitä kertomaan valintaperusteensa. Moni haastateltava koki valinnan vaikeaksi. Halusin kuitenkin antaa haastateltaville vapaat kädet, koska esimerkiksi parhaan jutun valintaperusteet kertovat osaltaan siitä, millaista datajournalismi heidän mielestään on tai millaista sen pitäisi olla.

Kaikkiaan uskon tutkimukseni tuovan uusia näkökulmia keskusteluun, jossa pohditaan datajournalismin olemusta ja merkitystä. Se tarjoaa yhden mahdollisen määritelmän datajournalismille. Samalla se antaa ymmärtää, että datajournalismi voi kehittyä ja vakiinnuttaa asemansa. Siihen kuitenkin tarvitaan lisää tekijöitä ja resursseja.

## **6.4 Jatkoa varten**

Pureduin tutkimuksessani perustavanlaatuisiin kysymyksiin lähtien siitä, mitä datajournalismi oikeastaan on. Uskon tällaisten perusasioiden käsittelyn tarjoavan paljon mahdollisuuksia syventäviin jatkotutkimuksiin. Jatkotutkimus on tarpeen, sillä datan määrä lisääntyy, journalismi siirtyy yhä enemmän verkkoon ja esitystavat kehittyvät.

Haastattelin tutkimukseen teknisiä osaajia ja toimittajia, mutta graafikot jäivät tutkimuksen ulkopuolelle. Helsingin Sanomien Esa Mäkistä lukuun ottamatta haastattelin esimiehiä vain taustaksi.

Yksi kiinnostava tutkimusasetelma olisi jonkin datajournalismia tekevän yksikön kaikkien työntekijöiden haastattelemine. Suomessa kyseeseen tulevat Ylen Plus-deski ja Helsingin Sanomien datadeski.

Laajojen aiheiden syleilyn lisäksi olisi tarpeen keskittyä myös kapeampiin osa-alueisiin. Tutkimuksessani nousi esiin, että datajournalismia tehdään ensisijaisesti verkkoon. Olisi paikallaan tutkia, millaisen roolin printti, televisio ja radio voivat datajournalismissa saada.

Ehkä vahvimmin nousi esiin tarve tutkimukselle, jossa selvitetään kattavasti datan roolia journalismissa. Sellainen voisi toimia myös rajanvetona sille, milloin on kyse datajournalismista.

## Lähdeluettelo

Aitamurto, Tanja & Sirkkunen, Esa & Lehtonen, Pauliina (2011) *Trends in Data Journalism*. Next Media. <[http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2011/D3.2.1.2.B\\_Hyperlocal\\_Trends\\_In%20Data\\_Journalism.pdf](http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2011/D3.2.1.2.B_Hyperlocal_Trends_In%20Data_Journalism.pdf)> Viitattu 24.3.2014.

Ala-Kyyny, Minna & Patrikainen, Susanna (2013) *Kuluttajat mukaan jutun tekoon – joukkoistaminen ja datajournalismi Kuningaskuluttajan juttuprosessissa*. Haaga-Helia-ammattikorkeakoulu: journalismin opinnäytetyö. <<http://publications.theseus.fi/bitstream/handle/10024/61380/Kuluttajat%20mukaan%20jutun%20teko.on.pdf?sequence=1>> Viitattu 24.3.2014.

Appelgren, Ester & Nygren, Gunnar (2014) Data Journalism in Sweden. Introducing new methods and genres of journalism into “old” organizations. *Digital Journalism* 2(3): 394–405. DOI: 10.1080/21670811.2014.884344. Viitattu 21.12.2014.

Bradshaw, Paul & Rohumaa, Liisa (2011) *The Online Journalism Handbook. Skills to survive and thrive in the digital age*. Englanti: Pearson Education Limited.

Bradshaw, Paul (2012) What is Data Journalism? Teoksessa Gray, Jonathan & Chambers, Lucy & Bounegru, Liliana (toim.) *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. O'Reilly Media. Saatavilla sähköisesti: <<http://datajournalismhandbook.org/>> Viitattu 23.4.2015.

Bounegru, Liliana (2012) Data Journalism in Perspective. Teoksessa Gray, Jonathan & Chambers, Lucy & Bounegru, Liliana (toim.) *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. O'Reilly Media. Saatavilla sähköisesti: <<http://datajournalismhandbook.org/>> Viitattu 23.4.2015.

Coddington, Mark (2014) Clarifying Journalism's Quantitative Turn. *Digital Journalism* 3(3): 331–348. DOI:10.1080/21670811.2014.976400. Viitattu 21.12.2014.

De Maeyer, Juliette & Libert, Manon & Domingo, David & Heinderyckx, François & Le Cam, Florence (2014) Waiting for Data Journalism: A Qualitative Assessment of the Anecdotal Take-up of Data Journalism in French-speaking Belgium. *Digital Journalism* 3(3): 432–446. DOI: 10.1080/21670811.2014.976415. Viitattu 21.12.2014.

Eskola, Jari & Vastamäki, Jaana (2010) Teemahaastattelu: opit ja opetukset. Teoksessa Aaltola, Juhani & Valli, Raine (toim.) *Ikkunoita tutkimusmetodeihin I. Metodien valinta ja aineistonkeruu: virikkeitä aloittelevalle tutkijalle*. 3. painos. Jyväskylä: PS-kustannus.

Fink, Katherine & Anderson, C. W. (2014) Data Journalism in the United States. *Journalism Studies*. DOI:10.1080/1461670X.2014.939852. Viitattu 21.12.2014.

Gray, Jonathan & Chambers, Lucy & Bounegru, Liliana toim. (2012) *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. O'Reilly Media. Saatavilla sähköisesti: <<http://datajournalismhandbook.org/>> Viitattu 23.4.2015.

Hirsjärvi, Sirkka & Hurme, Helena (2011) *Tutkimushaastattelu. Teemahaastattelun teoria ja käytäntö*. Helsinki: Gaudeamus.

Heinonen, Ari (1999) *Journalism in the Age of Net. Changing Society, Changing Profession*. Väitöskirja. Tampere: Tampere University Press.

<<https://tampub.uta.fi/bitstream/handle/10024/67684/951-44-5349-2.pdf?sequence=1>> Viitattu 25.3.2014.

Helsingin Sanomain Säätiö vuonna 2013 myöntämät apurahat ja stipendit (i.v.).

<[http://www.hssaatio.fi/wp-content/uploads/2011/09/hss\\_apurahat\\_2013.pdf](http://www.hssaatio.fi/wp-content/uploads/2011/09/hss_apurahat_2013.pdf)> Viitattu 24.3.2014.

Helsingin Sanomat (2012) *HS perustaa datajournalismin ryhmän*. Helsingin Sanomien uutinen.

<<http://www.hs.fi/kotimaa/a1305567392362>> Viitattu 27.10.2014.

Hermida, Alfred (2011) Mechanisms of Participation. How audience options shape the conversations. Teoksessa Singer, Jane B. & Hermida, Alfred & Domingo, David & Heinonen, Ari & Paulussen, Steve & Quandt, Thorsten & Reich, Zvi & Vujnovic, Marina *Participatory Journalism. Guarding Open Gates at Online Newspapers*. Chichester: Wiley-Blackwell, 13–33.

Holovaty, Adrian (2006) *A fundamental way newspaper sites need to change*. Blogikirjoitus.

<[www.holovaty.com/writing/fundamental-change/](http://www.holovaty.com/writing/fundamental-change/)> Viitattu 21.12.2014.

Howard, Alexander B. (2014) *The Art and Science of Data-driven Journalism*. Tow Center for Digital Journalism at Columbia University. <<http://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Data-Driven-Journalism.pdf>> Viitattu 12.11.2014.

Huuskonen, Matti (2013) *HS vakinaistaa datadeskinsä*. Helsingin Sanomien uutinen.

<<http://www.hs.fi/kotimaa/a1372389389078>> Viitattu 27.10.2014.

Jones, Janet & Salter, Lee (2012) *Digital journalism*. Lontoo: Sage.

Joukkouttaminen (i.v.). Wikipedia. <<http://fi.wikipedia.org/wiki/Joukkouttaminen>> Viitattu 28.3.2014.

Karlsen, Joakim & Stavelin, Eirik (2014) Computational Journalism in Norwegian Newsrooms. *Journalism Practice* 8(1): 34–48. DOI: DOI:

10.1080/17512786.2013.813190.10.1080/17512786.2013.813190. Viitattu 21.12.2014.

Kinnari, Tomi & Lindman, Juho & Rossi, Matti (2012) *New Value Network Profiles for Open Data*. Next Media.

<[http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2012/D3.3.2.2\\_HYPERLOCAL\\_New\\_value\\_network\\_profiles\\_for\\_open\\_data.pdf](http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2012/D3.3.2.2_HYPERLOCAL_New_value_network_profiles_for_open_data.pdf)> Viitattu 24.3.2014.

Kinnari, Tomi & Lindman, Juho & Rossi, Matti (2012(2)) *Detailed sample profiles of Finnish open data companies*. <[http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2012/D3.3.2.3\\_HYPERLOCAL\\_Detailed%20sample%20profiles%20of%20Finnish%20open%20data%20companies.pdf](http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2012/D3.3.2.3_HYPERLOCAL_Detailed%20sample%20profiles%20of%20Finnish%20open%20data%20companies.pdf)> Viitattu 28.3.2014.

Kinnari, Tomi (2013) *Open data business models for media industry – Finnish case study*. Aalto-yliopiston kauppakorkeakoulu: tietojärjestelmätieteen pro gradu -tutkielma.

<[https://aaltodoc.aalto.fi/bitstream/handle/123456789/10140/hse\\_thesis\\_13166.pdf?sequence=1](https://aaltodoc.aalto.fi/bitstream/handle/123456789/10140/hse_thesis_13166.pdf?sequence=1)>



Viitattu 24.3.2014.

Kiviniemi, Kari (2010) Laadullinen tutkimus prosessina. Teoksessa Aaltola, Juhani & Valli, Raine (toim.) *Ikkunoita tutkimusmetodeihin II. Näkökulmia aloittelevalle tutkijalle tutkimuksen teoreettisiin lähtökohtiin ja analyysimenetelmiin*. 3. painos. Jyväskylä: PS-kustannus.

Koskenranta, Riikka (2013) *Toimittaja, data tarinaksi. Datajournalismi uutistoimituksen työkalupakkiin*. Metropolia-ammattikorkeakoulu: mediatuottamisen opinnäytetyö.  
<[https://www.theseus.fi/bitstream/handle/10024/59105/130516\\_Koskenranta\\_R\\_opinnaytetyo\\_final.pdf?sequence=1](https://www.theseus.fi/bitstream/handle/10024/59105/130516_Koskenranta_R_opinnaytetyo_final.pdf?sequence=1)> Viitattu 24.3.2014.

Kunelius, Risto (2003) *Viestinnän vallassa*. Porvoo: WS Bookwell Oy.

Kuutti, Heikki (2002) *Tutkittu juttu. Johdatus tutkivaan journalismiin*. Jyväskylä: Atena Kustannus Oy.

Kähkönen, Heidi (2013) *Dataa, ei dadaa – Näin syntyi datajournalismin kurssi*. Haaga-Helia-ammattikorkeakoulu: journalismin opinnäytetyö.  
<[http://publications.theseus.fi/bitstream/handle/10024/62050/Kahkonen\\_Heidi.pdf?sequence=1](http://publications.theseus.fi/bitstream/handle/10024/62050/Kahkonen_Heidi.pdf?sequence=1)> Viitattu 24.3.2014.

Kärkimedia (2014) *Keskisuomalainen, Karjalainen, Savon Sanomat ja ESS yhteistyöhön*. Suomen lehdistön uutislaina. <<https://www.karkimedia.fi/karkimedia/mita-mediassamme-tapahtuu/10517/keskisuomalainen-karjalainen-savon-sanomat-ja-ess-yhteistyoehoen.html>> Viitattu 6.11.2014.

Laakso, Ville & Pirhonen, Kalle (2013) *Vuoden luetuimmat jutut: Pisa-testi, uusi painoindeksi, kuvat liikkuvan poliisin arjesta...* Yle Uutisten uutinen. <[http://yle.fi/uutiset/vuoden\\_luetuimmat\\_jutut\\_pisa-testi\\_uusi\\_painoindeksi\\_kuvat\\_liikkuvan\\_poliisin\\_arjesta/7003416](http://yle.fi/uutiset/vuoden_luetuimmat_jutut_pisa-testi_uusi_painoindeksi_kuvat_liikkuvan_poliisin_arjesta/7003416)> Viitattu 25.3.2014.

Laki viranomaisten toiminnan julkisuudesta (i.v.).  
<<http://www.finlex.fi/fi/laki/ajantasa/1999/19990621>> Viitattu 24.3.2014.

Lehtonen, Pauliina (2011) *Open data in Finland – Public sector perspectives on open data*. Next Media. <[http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2011/D3.2.1.2.\\_Hyperlocal\\_Open%20data%20in%20Finland.pdf](http://virtual.vtt.fi/virtual/nextmedia/Deliverables-2011/D3.2.1.2._Hyperlocal_Open%20data%20in%20Finland.pdf)> Viitattu 25.3.2014.

Leivonniemi, Samuli (2012) *Datajournalismi sanomalehden uutistuotannossa*. Aalto-yliopiston perustieteiden korkeakoulu: mediatekniikan kandidaatin tutkielma. <<https://docs.google.com/file/d/0B-wmyyQmG0gxbW1NdWp2d2VtWm8/edit?pli=1>> Viitattu 24.3.2014.

Lindenberg, Friedrich (2012) *Getting Data from the Web*. Teoksessa Gray, Jonathan & Chambers, Lucy & Bounegru, Liliana (toim.) *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. O'Reilly Media. Saatavilla sähköisesti: <<http://datajournalismhandbook.org/>> Viitattu 23.4.2015.

Lorenz, Mirko (2010) Status and Outlook for data-driven journalism. Julkaisussa *European Journalism Center: Data-driven journalism: What is there to learn?* A paper on the data-driven journalism roundtable held in Amsterdam on 24 August 2010, 8–17.  
<[http://mediapusher.eu/datadrivenjournalism/pdf/ddj\\_paper\\_final.pdf](http://mediapusher.eu/datadrivenjournalism/pdf/ddj_paper_final.pdf)> Viitattu 24.3.2014.

Lorenz, Mirko (2012) Why Journalists Should Use Data. Teoksessa Gray, Jonathan & Chambers, Lucy & Bounegru, Liliana (toim.) *The Data Journalism Handbook. How Journalists Can Use Data to Improve the News*. O'Reilly Media. Saatavilla sähköisesti: <<http://datajournalismhandbook.org/>> Viitattu 23.4.2015.

Meyer, Philip (1991) *The New Precision Journalism*. Bloomington & Indianapolis: Indiana University Press.

Mikä on Plus? (i.v.) Yle Uutisten määritelmä Plus-deskille. <<http://yle.fi/uutiset/plus/>> Viitattu 24.3.2014.

Minkkinen, Mika (2013) *Infografiikan uudet muodot. Graafista suunnittelua, dataa ja journalismia verkkojulkaisemisen aikakaudella*. Lahden ammattikorkeakoulun muotoilu- ja taideinstituutti: viestinnän opinnäytetyö.  
<[http://www.theseus.fi/bitstream/handle/10024/57119/minkkinen\\_mika.pdf?sequence=2](http://www.theseus.fi/bitstream/handle/10024/57119/minkkinen_mika.pdf?sequence=2)> Viitattu 24.3.2014.

MTV (2014) *MTV Uutisten tutkivan journalismin ryhmä aloittaa*. MTV:n uutinen.  
<<http://www.mtv.fi/uutiset/kotimaa/artikkeli/mtv-uutisten-tutkivan-journalismin-ryhma-aloittaa/2756272>> Viitattu 29.10.2014.

Nikander, Pirjo (2010) Laadullisten aineistojen litterointi, kääntäminen ja validiteetti. Teoksessa Ruusuvuori, Johanna & Nikander, Pirjo & Hyvärinen, Matti (toim.) *Haastattelun analyysi*. Tampere: Vastapaino.

Open data (i.v.). Wikipedia. <[http://en.wikipedia.org/wiki/Open\\_data](http://en.wikipedia.org/wiki/Open_data)> Viitattu 28.3.2014.

Pietilä, Kauko (2012) *Journalismi ammattina: journalistiprofession teoria*. Helsinki: Gaudeamus.

Poikola, Antti & Kola, Petri & Hintikka, Kari A. (2010) *Julkinen data – johdatus tietovarantojen avaamiseen*. <<http://www.scribd.com/doc/28845102/Julkinen-data>> Viitattu 3.3.2014.

Poikola, Antti (2013) *Mistä dataa saa*. Luentokalvot.  
<[https://dl.dropbox.com/u/2949803/datajournalismi\\_tutoriaalit/LONG%20B\\_Mista\\_dataa\\_saa.pdf](https://dl.dropbox.com/u/2949803/datajournalismi_tutoriaalit/LONG%20B_Mista_dataa_saa.pdf)> Viitattu 3.3.2014.

Rapeli, Minna (2013) *Data Journalism: An Outlook for the Future Processes*. Aalto-yliopiston kauppakorkeakoulu: tietojärjestelmätieteen pro gradu -tutkielma.  
<[http://epub.lib.aalto.fi/en/ethesis/pdf/13341/hse\\_ethesis\\_13341.pdf](http://epub.lib.aalto.fi/en/ethesis/pdf/13341/hse_ethesis_13341.pdf)> Viitattu 24.3.2014.

Royal, Cindy (2010) *The Journalist as Programmer: A Case Study of The New York Times Interactive News Technology Department*. International Symposium in Online Journalism. The University of Texas at Austin, April 2010. <<https://online.journalism.utexas.edu/2010/papers/Royal10.pdf>> Viitattu 30.3.2014.

Ruokoski, Vilma (2013) *Klikkailtava kartta, kysely ja aikajana*. Tampereen yliopisto: journalistiikan ja viestinnän kandidaatin tutkielma.  
<<https://docs.google.com/file/d/0B1kOnxI1kRPuZUHGOHUyalBtSVk/edit?pli=1>> Viitattu 24.3.2014.

Ruusuvuori, Johanna (2010) Litteroijan muistilista. Teoksessa Ruusuvuori, Johanna & Nikander, Pirjo

& Hyvärinen, Matti (toim.) *Haastattelun analyysi*. Tampere: Vastapaino.

Sarajärvi, Anneli & Tuomi, Jouni (2009) *Laadullinen tutkimus ja sisällönanalyysi*. Helsinki: Tammi.

Singer, Jane B. & Hermida, Alfred & Domingo, David & Heinonen, Ari & Paulussen, Steve & Quandt, Thorsten & Reich, Zvi & Vujnovic, Marina (2011) *Participatory Journalism. Guarding Open Gates at Online Newspapers*. Chichester: Wiley-Blackwell.

Tebest, Teemo (2015) *NODA15 – Päällimmäiset tuntemukset*. Blogikirjoitus.  
<<http://datajournalismi.blogspot.fi/2015/02/noda15-paallimmaiset-tuntemukset.html>> Viitattu 21.4.2015.

Terveisiä maailmalta (2011). Helsingin Sanomat. <<http://terveysiamailmalta.hs.fi/>> Viitattu 28.3.2014.

Toikkanen, Ilkka (2014) *Pinnallisista klikkikartoista yhteiskunnalliseen merkittävyyteen. Datajournalismin työprosessi Helsingin Sanomissa ja Ylessä*. Jyväskylän yliopisto: journalistiikan pro gradu -tutkielma.  
<<https://jyx.jyu.fi/dspace/bitstream/handle/123456789/43189/URN%3ANBN%3Afi%3Aju-201404081484.pdf?sequence=1>> Viitattu 21.12.2014.

Valkama, Meri (2012) *Tulevaisuuden tekijät*. Journalisti 15/2012.  
<<http://www.journalistiliitto.fi/journalisti/lehti/2012/15/artikkelit/tulevaisuuden-tekijat/>> Viitattu 24.3.2014.

Vehkoo, Johanna (2011) *Painokoneet seis! Kertomuksia uuden journalismin ajasta*. Teos, Helsinki.

Väliverronen, Esa (2009) Journalismi kriisissä? Teoksessa Väliverronen, Esa (toim.) *Journalismi murroksessa*. Helsinki: Gaudeamus.  
Yleisradio (2014) *Yle siirtää painopistettä verkkoon ja aloittaa yt-neuvottelut*. Yleisradion uutinen.  
<<http://yle.fi/yleisradio/ajankohtaista/yle-siirtaa-painopistetta-verkkoon-ja-aloittaa-yt-neuvottelut>> Viitattu 27.10.2014.

# Liite: Teemahaastattelun runko

## 1. Taustatiedot

- Nimi ja ikä?
- Koulutus, myös datajournalismikoulutus?
- Media, jossa työskentelet?
- Nykyinen työtehtävä ja sen kesto?
- Relevantti työhistoria?

## 2. Omat kokemukset datajournalismista

- Miten kiinnostuit datajournalismista?
- Miten päädyit tekemään datajournalismia?
- Kuvaile, millaisia tehtäviä työsi datajournalistina sisältää tai sisälsi. Missä roolissa yleensä osallistut datajournalististen juttujen tekemiseen?
- Ketä muita (henkilö/rooli) juttuprosesseihin osallistuu?
- Montaa datajournalistista juttua olet ollut tekemässä (suunnilleen, jos määrä on suuri)?
- Tuleeko datajournalistisista jutuista paljon palautetta yleisöltä? Minkälaista palautetta?
- Miten pidät yllä ammattitaitoa datajournalistina tai opettelet uusia taitoja datajournalistina?
- Aiotko tulevaisuudessa opetella lisää datajournalistisia taitoja? Mitä, miten, miksi?
- Miksi teet datajournalismia?

## 3. Datajournalistisen jutun synty

- Esittele lyhyesti paras datajournalistinen juttu, jota olet ollut tekemässä.
- Miksi valitsit juuri tämän jutun?
- Kerro jutun syntyprosessista.
  - Ketkä juttua olivat toteuttamassa?
  - Mitä itse teit jutussa?
  - Kuka keksi idean juttuun?
  - Kuka keksi jutun toteutustavan?
  - Eroaako juttu jotenkin “normaalista” datajournalistisesta jutusta?

- Mikä tekee jutusta datajournalismia?
- Miksi juttu on mielestäsi hyvä?

#### **4. Datajournalismin määritelmä**

- Millaista on mielestäsi hyvä datajournalismi? Kerro esimerkkejä.
- Millaista on mielestäsi huono datajournalismi? Kerro esimerkkejä.
- Miten datajournalismi eroaa muusta journalismista?
- Miten määrittelet datajournalismin (omin sanoin)?
  - Onko datajournalismin sisällä eri tyyppisiä?
  - Mikä on yleisin datajournalismin julkaisukanava? Miten eri julkaisukanaviin tehty datajournalismi eroaa toisistaan?
- Kuka on datajournalisti? Mitä datajournalistin pitää mielestäsi osata?
- Mikä on mielestäsi datajournalistin tehtävä?
- Onko datajournalismilla eri/sama tehtävä kuin muulla journalismilla?

#### **5. Datajournalismin merkitys**

- Millaisena näet datajournalismin aseman suomalaisessa mediakentässä nyt? Entä tulevaisuudessa?
  - Tehdäänkö Suomessa riittävästi datajournalismia?
- Mitä annettavaa datajournalismilla on tai voi olla journalismille?
- Millainen on datajournalismin yhteiskunnallinen merkitys/merkitys yleisölle?
- Miten uskot datajournalismin muuttuvan tulevaisuudessa?