

Riikka Koskenranta

Toimittaja, data tarinaksi

Datajournalismi uutistoimituksen työkalupakkiin

Metropolia Ammattikorkeakoulu

Medianomi (ylempi AMK)

Mediatuottamisen koulutusohjelma

Opinnäytetyö

3.5.2013

Tekijä(t) Otsikko Sivumäärä Aika	Riikka Koskenranta Toimittaja, data tarinaksi – Datajournalismi uutistoimituksen työkalupakkiin 34 sivua + 5 liitettä 03.05.2013
Tutkinto	Medianomi (ylempi AMK)
Koulutusohjelma	Mediatuottaminen
Suuntautumisvaihtoehto	-
Ohjaaja(t)	FT, yliopettaja Robert Arpo
<p>Tämän opinnäytetyön tavoitteena on tutustuttaa pieniä uutistoimituksia datajournalismin tekemiseen. Datajournalismilla tarkoitetaan journalistista prosessia, jossa laajoja tietoaaineistoja yhdistelemällä, analysoimalla ja visualisoimalla tuotetaan uutisia. Teen opinnäytetyön työnantajani Suomen Yrittäjien Sypoint Oy:n toimeksiannosta. Opinnäytteen viitekehyksenä on toteutettu kehittämishanke Suomen Yrittäjien toimituksessa. Tutkimuksen lähestymistapa on laadullinen. Tutkimuksessa on syvennetty tutkimuskirjallisuudesta esiin nousseita asioita asiantuntijahaastatteluilta. Kehittämishankkeen menetelmänä on toimintatutkimus. Opinnäytetyössä analysoidaan kehittämishankkeessa toteutetun datajournalismijutun onnistumiset ja epäonnistumiset. Kehittämishanketta on arvioitu oppimisen ja tulosten näkökulmasta.</p> <p>Datajournalismilla on annettavaa tänä päivänä, kun uutismediat etsivät uusia tapoja esittää juttuja verkossa ja hyödyntää digitaalisen alustan tuomat mahdollisuudet. Aiempien datajournalismia koskevien selvitysten ja asiantuntijahaastatteluiden perusteella toimituskunta tarvitsee monipuolisesti oppia datajournalismin tekemiseen. Tähän tarpeeseen opinnäytetyö osaltaan vastaa. Mediatiloille datajournalismi kannattaa myös siksi, että datavisualisoinnit keräävät huomiota pitkän aikaa, ns. pitkä häntä. Tutkimuksessa kävi ilmi, että päätoimittajien kannattaa motivoida toimituskuntaansa korostamalla sitä, että datajournalismissakin juttu on se pääasia. Taidot ja työkalut ovat opetettavissa ja uutisvainu on kaiken a ja o.</p> <p>On todennäköistä, että toimittaja, jolla on ammattitaito tehdä datajournalismia, on edullisemmassa asemassa työmarkkinoilla. Tuloksista ilmeni, että suunnitelmallisuus alusta alkaen on tärkeää, sitä vähemmän juttuprosessi vie työaikaa. On niin ikään pohdittava printin ja verkon funktiot, niin että lisäarvoa syntyy kummassakin kanavassa. On selvää, että toimituksen juttupalavereissa pitää totutun lähestymistavan lisäksi kysyä myös, millaista dataa kulloisenkin juttuidean taustalla on. Jos data ei ole mukana juttupalavereissa, on vaarana, että datajournalismin tekemisestä tulee luonteeltaan satunnaista.</p>	
Avainsanat	datajournalismi, verkkojournalismi, avoin data

Author(s) Title	Riikka Koskenranta Integration of Data Journalism into the News Media Tool Kit
Number of Pages Date	34 pages + 5 appendices 3 May 2013
Degree	Master of Arts
Degree Programme	Media Production and Management
Specialisation option	-
Instructor(s)	Robert Arpo, PhD
<p>The objective of the present thesis is to provide an introductory guide for small newsrooms to data journalism. The term "data journalism" stands for a journalistic process, which produces news by combining, analyzing and visualizing data. My employer, the Federation of Finnish Enterprises, commissioned the thesis. A developing project, which has been carried out in the newsroom of the Federation of Finnish Enterprises, forms a frame of reference for the present thesis.</p> <p>The approach of the research is qualitative. The study has employed specialist interviews to further examine the issues which emerged from the research literature. The used method of the project is action research. The thesis analyses the successes and failures of a data journalism story produced during the developing project. The developing project has been assessed with respect to learning and the gained results.</p> <p>Data journalism is topical now when the news media are looking for new ways to show stories on the web and make use of opportunities in the digital platform. On the grounds of previous research on the subject and specialist interviews, it is clear that journalists need thorough education on data journalism skills. The present thesis will thus meet the educational gap and need. Data journalism is worthwhile for the media houses also because data visualization stories gather attention over a long period of time. This is referred to by the term "long tail". The research results indicated that managing editors should motivate their staff by emphasizing journalistic instinct being the most important skill whereas data management skills can be taught and learned.</p> <p>It is likely that a journalist with data journalism skills will have an advantage in the recruitment process. The results indicate that with a better planning, the shorter the journalistic process will be. It is likewise necessary to contemplate the function for print versus web, so that added value will be created in both the channels. It is clear that stories which involve data will need to be present in editorial meetings, and the question to be asked is 'what data is behind the story'.</p>	
Keywords	data journalism, online journalism, open data

Sisällys

1	Johdanto	1
1.1	Tutkimuksen tavoite ja työn rakenne	3
1.2	Tutkimusmenetelmät	4
2	Datajournalismin kehitys	5
2.1	Datajournalismin tavoitteet	7
2.2	Vuorovaikutteinen journalismi	10
2.3	Avoin data ja journalismi	11
3	Datajournalismi käytännössä	12
3.1	Mistä dataa	15
3.2	Datajournalistin työkalut	16
3.3	Datajournalistinen työprosessi ja arviointi	20
4	Analyysi datajournalismijutusta, case: kaupunkien sotepalveluostot	25
5	Pohdinta ja arvio tutkimuksen hyödyntämismahdollisuuksista	28
	Lähteet	32
	Liitteet	
	Liite 1. Datajournalismin työnkulun kehittämisriihi	
	Liite 2. Ohjeet datan hankintaan	
	Liite 3. Excel-kuvaajan valitseminen	
	Liite 4. Datajournalismin toimitusprosessin malli (lähde: Antti Poikola)	

Copyright © 2013 Riikka Koskenranta

Tämän teoksen käyttöoikeutta koskee Creative Commons Nimeä 1.0 Suomi -lisenssi.

1 Johdanto

Olen työskennellyt vuodesta 2007 Suomen Yrittäjien mediajulkaisuissa toimittajana. Teemme kuukausittain ilmestyvää Yrittäjäsanomien tabloidlehteä, Yrittäjä-aikakauslehteä ja yrittajat.fi – sivuille uutisointia. Olen mielenkiinnolla seurannut vuoden 2012 aikana Suomessa nousuun lähtenyttä datajournalismia, jota vasta muutama media Suomessa tekee säännöllisesti. Halusin lähteä kehittämään toimituksemme osaamista datajournalismin tekemisessä.

Tämän opinnäytetyön tavoitteena on tutustuttaa pieniä uutistoimituksia datajournalismin tekemiseen. Datajournalismilla tarkoitetaan journalistista prosessia, jossa laajoja tietoaaineistoja yhdistelemällä, analysoimalla ja visualisoimalla tuotetaan uutisia. Avoin data eli viranomaisten ja yritysten julkaisemat tietoaaineistot toimii merkittävänä uutislähteenä datajournalismissa. Tutkimus vastaa kysymykseen, miten uutistoimitus voi ottaa datajournalismin keinovalikoimaansa luomalla katsauksen datan käsittely- ja visualisointityökaluihin ja ohjaamalla verkosta saataviin oppaisiin toimitusprosessin eri vaiheissa. Opinnäytetyön viitekehyksenä on toimintatutkimus Suomen Yrittäjien Sypoint Oy:n (jäljempänä SY) toimituksessa. Toivon, että opinnäytetyössä kuvatus toimintatutkimuksen ja kehittämishankkeen keskeiset havainnot ja dokumentointi auttavat lukijaa hänen kehittäessään datajournalismin osaamistaan.

Työn tavoitteena on kehittää datajournalismin toimitusprosessin malli työnantajani SY:n viiden hengen toimituksen käyttöön. Suomen Yrittäjien Sypoint Oy muun muassa kustantaa ja julkaisee kahta lehteä ja yrittajat.fi -verkkopalvelua. Toimintatutkimus sopii kehittämishankkeeni menetelmäksi, sillä sen avulla tavoitellaan käytännön hyötyä.

Tämä työ kuvaa samalla datajournalismin oppimisprosessiani toimittajana.

Toimintatutkimuksessa määritellään luonnollisesti toiminnan kehittämisen tavoite, josta lisää jäljempänä (Kuula 1999, 218).

Syvensin kehittämishankkeen kanssa samanaikaisesti datajournalismin työkalujen osaamistani Otavan opiston järjestämällä datajournalismin kurssilla helmi-huhtikuussa 2013. Oppimisprosessi toki jatkuu itsenäisesti kurssin jälkeen. Toteutin datajournalismikurssilla yhden SY:n ensimmäisistä datajournalismiuutisista printtiin ja

verkkoon. Opinnäytetyön 5. luvussa on analyysi maakuntien keskuskaupunkien sosiaali- ja terveystalvueluostoja koskevasta toteuttamastani datajournalismijutusta.

Opinnäytetyön rakenne muodostuu kirjallisuuskatsauksesta ja toimintatutkimusprosessista. Tutkimuksessa käyn vuoropuhelua aineiston ja kehittämishankkeen tavoitteiden siivittämänä aineistosta esiin nousseiden teemojen kautta. Kirjallisuuden koodaamisessa käytetty aineiston teemoittelu käydään läpi luvussa 1.2. Asiantuntijahaastattelujen tarkoituksena on syventää ja tuoda mahdollisimman tuoretta näkemystä tutkimukseni tueksi. Arja Kuulan kuvaus toimintatutkimuksesta auttoi minua kehittämishankkeeni suunnittelussa. Hänen mukaansa toimintatutkimus voidaan ymmärtää tutkimuksellisenä lähestymistapana, jossa hyödynnetään tutkimukselle olennaisia erilaisia kvalitatiivisen ja/tai kvantitatiivisen tutkimuksen menetelmiä (1999, 218).

Tutkimus jakautuu viiteen osaan. Datajournalismin kehitys - kappaleessa luodaan katsaus datajournalismin kehitykseen, tavoitteisiin ja ominaispiirteisiin aiemman kirjallisuuden pohjalta. Tutkimusta varten tehdyt asiantuntijahaastattelut tuovat lisävalaisua datajournalismin merkityksestä eurooppalaiselle journalismille ja kertovat datajournalismin tilanteesta tänä päivänä. Kolmannessa kappaleessa yhdistyvät toimintatutkimuksen vaiheet suunnittelusta tulosten arviointiin sekä tutkimuskysymykseen vastaaminen datajournalistisen työprosessin kautta. Tässä luvussa käydään läpi datajournalistin työkalupakkia. Seuraavassa, Analyysi datajournalismijutusta - luvussa, arvioin työyhteisöni erään dataa hyödyntävän uutisen työvaiheiden ja -prosessin onnistumisen. Lopuksi tutkimus tarjoaa suosituksia jatkoa ajatellen, ja tuo esiin näkemyksiä datajournalismiin panostamisesta kustannustehokkaalla tavalla. Kehittämishankkeen kohderyhmiä ovat pienet uutis- ja mediatoimitukset sekä yksittäiset toimittajat, jotka eivät vielä ole perehtyneet datajournalismiin.

Mitä datajournalismi sitten tarkasti ottaen tarkoittaa? Wikipedian mukaan datajournalismilla tarkoitetaan journalismin suuntausta, joka hyödyntää laajasti numeerista dataa uutistyössä ja myös uutisten esittämistavassa (Wikipedia 2012). Tämän journalismin muodon yhtenä ulottuvuutena on vuorovaikutus yleisön kanssa internetissä. Suomessa tehty datajournalismi on vielä melko nuorta. Datajournalismi ei ole ilmestynyt tyhjästä, vaan tietokantajournalismia ja infografiikoita on tehty jo vuosikymmeniä. Datajournalismi on tutkivaa journalismia; tärkeintä on datasta löytyvä

tarina. Datajournalismille tyypillinen jutun visuaalinen esitystapa tukee tarinaa. Ison-Britannian yleisradioyhtiö BBC määrittelee datajournalismin muun muassa todeten, että se mahdollistaa lukijalle sellaisen tiedon löytämisen, joka on hänelle relevantti. Datajournalismin tavoitteena on paljastaa asia tai juttu, joka on merkittävä ja aiemmin tuntematon (Hurrell & Leimdorfer 2012, Grayn, Bounegrün ja Chambersin mukaan 2012, 28).

1.1 Tutkimuksen tavoite ja työn rakenne

Toimintatutkimuksen tavoitteena on kehittää datajournalismin toimitusprosessin konsepti/malli YrittäjäSanomien, yrittajat.fin ja Yrittäjä-lehden toimitukselle. Teen tämän kehittämishankkeen työnantajani Suomen Yrittäjien Sypoint Oy:n käyttöön. Yhtiö on Suomen Yrittäjät ry:n omistama.

Datajournalismi käytännössä -luvussa esittelen datajournalistisen prosessin vaiheet ideasta valmiiseen tuotokseen ja keskeiset datajournalistin työkalut. Tässä luvussa tarkastellaan lisäksi SY:n toimituksen datajournalismiin vihkiytymisen tavoitteita ja päämääriä sekä muun muassa toimituksen näkemyksiä siitä, miten SY:ssä valitaan dataa, jota lähdetään penkomaan. Seuraava luku ”Analyysi datajournalismijutusta” tuo erehdyksen ja oppimisen kautta kokemuksellista sisältöä aloittelevalle datatoimittajalle. Näin syntyvän datajournalistin työkalupakin arvioidaan tarjoavan valmiuksia toimittajille lähteä toteuttamaan omia, ensimmäisiä datajournalismijuttujaan. Suomenruotsalainen toimittaja Jens Finnäs (2013) määritteli datatoimittajan näppärästi: ”Datajournalisti on kuin tavallinen toimittaja, mutta muistiinpanot ovat Excelissä.”

Aiempien selvitysten, asiantuntijahaastatteluiden sekä subjektiivisten ja jaettujen mediakentän havaintojen perusteella toimituskunta tarvitsee monipuolisesti oppia datajournalismin tekemiseen (Poikola 26.2.2013). Tähän tarpeeseen opinnäytetyö osaltaan vastaa. Tämä tutkimus vastaa myös siihen, miksi toimittajien pitäisi käyttää dataa yhtenä lähdemuotona. Kehittämishankkeen lähtökohtana on myös tavoite nostaa entisestään journalismin laatua, datajournalismijutut ovat keino profiloitua ja herättää lukijoiden kiinnostus SY:n kaltaiselle talouspainotteiselle verkkouutisten tarjoajalle.

1.2 Tutkimusmenetelmät

Tutkimuskysymykseni on: miten uutistoimitus voi ottaa datajournalismin keinovalikoimaansa. Tutkimuksen päätavoitteen tukena tutkimus vastaa myös kysymykseen: miten datajournalismin toimituksellinen prosessi etenee. Tämän opinnäytetyön laadullinen tutkimus rakentuu aineistoanalyysistä sisältäen myös seminaariaineistoja sekä kolmesta datatoimittajan ja yhdestä avoin data -aktiivin asiantuntijahaastatteluista. Tutkimuksen kehittämishankkeen tavoitteiksi SY:n toimitus määritteli marraskuussa 2012: 1. nostaa journalistista laatua, 2. tuoda lukijoille datajournalismilla tietoa, jota muut eivät tuota, 3. datajournalismilla tuotetaan uutisia, joita muut eivät voi verkossa lainata linkittämättä ja 4. datatyökalujen osaamisen levittäminen, ettemme ole datajournalismin teossa riippuvaisia vain yhdestä henkilöstä. Näistä etenkin 1, 2 ja 4 ovat toimintatutkimuksen päämääriä, joiden suuntaan toimituksen yhteisen oppimisprosessin aikana kuljetaan. 3. tavoite on luonteeltaan sivutuotteena saatava, eli kun teemme datajournalismijuttuja, niitä ei voi ”varastaa” muiden medioiden sivuille. Haemme datajournalismista lisäarvoa niin yrittäjät.fi:hin kuin Yrittäjäsanomiiin. Olennainen huomio palaverissa oli myös datatyökalujen osaamisen levittäminen, ettemme ole datajournalismin teossa riippuvaisia vain yhdestä henkilöstä. Kehittämishankkeen lähestymistapa on toimintatutkimuksellinen. Toimintatutkimuksen rakennetta voi kuvata perinteisenä spiraalimaisena syklinä, joka muodostuu suunnittelusta, toiminnasta, havainnoinnista ja reflektoinnista (Anttila 2007, 47).

Kehittämishanketta arvioidaan oppimisen ja tulosten näkökulmasta. Tässä työssä dokumentoidaan SY:n toimituksen arviointi sen huhtikuussa 2013 toteutetun datajournalismijutun työprosessin onnistumisista ja epäonnistumisista. On kuitenkin huomioitava, että kehittämishankkeen verraten lyhyt ajanjakso marraskuusta 2012 huhtikuun 2013 loppuun tarkoitti tämän työn kannalta sitä, että kehittämistyö jatkuu SY:n toimituksessa.

Tutkimuksen kirjallisuusanalyysia varten koko aineisto teemoitettiin seuraavanlaisesti: 1) datajournalismin tausta ja kehitys, 2) journalismin vuorovaikutteisuus, 3) avoin data vs. journalismi, 4) datajournalismin toimitusprosessi ja työkalut, 5) datajournalismin tavoitteet ja 6) datajournalismin seuraava taso.

Tämä työ saa lisää painoarvoa asiantuntijahaastatteluista, sillä verkkoympäristössä kehitys on nopeaa, ja haastattelut antavat tuoreen kuvan siitä, miten datatoimittajat

kokevat verkkotyökalujen laajan kirjon datan työstämisessä ja visualisoinnissa. Tutkimusta varten on kerätty aineistoa niin ikään datajournalismiin ja datan visualisointiin keskittyneissä seminaareissa, kuten Open Knowledge Festival - tapahtuman datajournalismisessiossa syyskuussa 2012. Opinnäytetyöprosessin aikana keräämäni linkkilista on nähtävissä Delicious-palvelussa (<https://delicious.com/#riikkis1>).

Kehittämishankkeessa roolini on ollut kaksijakoinen, olen ollut yhtäältä osallinen ja toisaalta olen ollut ulkopuolinen tutkija. Olen pyrkinyt toimimaan käytäntöhakuisesti ja objektiivisesti (Anttila 2007, 36).

2 Datajournalismin kehitys

”Datan analysoinnissa on journalistien tulevaisuus”, Tim Berners-Lee, world wide webin keksijä sanoi vuonna 2010 (Charles 2010). Hänen viestinsä toimittajille oli, että toimittajien pitää olla datalukutaitoisia. Eikä hän ole yksin mielipiteinen: datajournalismi on päivän sana journalismissa. Datajournalismin vaikuttimet tulevat ainakin kahta kautta: toisaalta journalismia ravisteleva uudistuminen on sopeutumista tilanteeseen, jossa toimittajan rooli tiedon portinvartijana on murentunut ja toisaalta kyseessä on digitalisoitumisen tarjoama mahdollisuus, jossa verkkoteknologioiden kehitys mahdollistaa toimittajalle uudenlaisten uutisten esitystapojen haltuunoton. Tampereen yliopiston *Viestintäalan ja - ammattien tulevaisuuden osaamistarpeita* käsittelevässä verkkodokumentissa Paula Haara kirjoittaa toimittajien roolin muuttumisesta maailmassa, jossa avoimen tietomäärän kasvu on vienyt toimittajilta tiedon portinvartijan aseman. Haara nostaa yhtenä mahdollisuutena esille datajournalismin. (2012, 16) Muun muassa The BBC:n Steve Herrmann sanoi journalism.co.uk:n artikkelissa, että uutissivustot etsivät nyt tapoja käyttää dataa kertomaan tarinaa (Herrmann 2013, Marshallin mukaan 2013). Myös Suomessa mediatilat ovat tehneet ”löytöretkiä” datajournalismiin, mutta säännöllisesti datajournalismia tuottavat mediat voi vielä laskea yhden käden sormilla.

Datajournalismi ei ole syntynyt tyhjästä, vaan sen edeltäjiä ovat muun muassa tietokantajournalismi, tietokoneavusteinen journalismi ja infografiikat. Bradshaw’n ja Rohumaan (2011, 48) mukaan tietokantojen potentiaali uutishankinnassa löydettiin jo 1960-luvulla, kun amerikkalainen toimittaja Philip Meyer hyödynsi niitä rotumellakoiden tutkinnassa. Silloin puhuttiin tietokoneavusteisesta journalismista. Varsinaisesti

datajournalismin, sellaisena kuin datajournalismi tänä päivänä tunnustetaan, ensimmäisiä muotoja nähtiin vuonna 2006. Adrian Holovaty'n perustama, nyt jo lopetettu, EveryBlock –tietopalvelu auttoi kansalaisia näkemään, mitä tapahtuu heidän asuinalueellaan. (Gray ym. 2012, 18).

Datajournalismin kenttä kattaa muun muassa infografiikoiden käyttöä, tietoverkkojournalismin (eli tiedonhallintajärjestelmien, jossa tieto organisoidaan tietokannassa) ja tiedon visualisoinnin (Wikipedia 2012). Siitä, onko infografiikka datajournalismia, on erilaisia näkemyksiä. Ylen PlusDeskin datajournalisti Teemo Tebest (2012) kirjoittaa verkkoblogissaan, että: ”datajournalismi-käsitteessä oleellista on, että taustalla on tietoaaineisto, jota on käsitelty sekä analysoitu, ja joka esitetään journalistisesta näkökulmasta esimerkiksi visualisoinnin keinoin”.

Tietokannoista avoimeen dataan

2000-luvulla tietokantajournalismi on muuttunut datajournalismiksi, sillä dataa on saatavilla yhä enemmän (Bradshaw ja Rohumaa 2011, 48). Datajournalismi lähti isommassa mittakaavassa liikkeelle monen tapahtuman seurauksena. Guardianin pitkän linjan datatuottaja Simon Rogers pohtii datajournalismin nousun syitä.

Nostaisin tärkeimmiksi syiksi, että hallitukset alkoivat julkistaa dataansa, toisaalta Tim Berners-Leen esiintulo ja vaatimus avoimesta datasta oli iso asia. Mielestäni uutisorganisaatiot heräsivät tilanteeseen vähän myöhässä. Guardian aloitti Data Blogin vuonna 2009. (10.1.2013)

Helsingin Sanomien datatuottaja Esa Mäkinen nostaa merkittävimmiksi asioiksi datajournalismin yleistymisessä valmiiden datatyökalujen syntymisen sekä julkisen ja avoimen datan määrän kasvun (Mäkinen 15.3.2013). Myös Suomen hallitus teki maaliskuussa 2011 periaatepäätöksen julkisen tietoaaineiston saatavuudesta niin, että data on avoimesti saatavilla ja uudelleenkäytettävissä yhtenäisin ehdoin. Tämä on antanut aivan uudenlaisen mahdollisuuden datajournalismille. Nyt osa valtion virastoista, kuten Maanmittauslaitos ja Ilmatieteenlaitos, ovat avanneet dataansa kiinnostuneiden käyttöön. Verohallinto avasi joulukuussa 2012 ensimmäisen kerran yhteisöjen tuloverotuksen dataa verkossa. Datajournalismin voi perustellusti sanoa olevan nousukiidossa suomalaisissa medioissa.

Julian Assangen perustama WikiLeaks nousee usein esiin, kun kerrataan datajournalismin historiaa. WikiLeaks on voittoa tavoittelematon organisaatio. Muun

muassa Guardian ja New York Times käsittelivät WikiLeaksin avaamia valtavia tietoaaineistoja esimerkiksi Afganistanin ja Irakin sotien salaisista dokumenteista. (Gray, Bounegru ja Chambers 2012). WikiLeaksin vuonna 2010 vuotamat diplomaattisähkeet ja niitä seurannut uutisointi toivat muun muassa Guardianin datajournalismia näyttävästi esille.

Amerikkalaisen Burt Hermanin perustamalla Hacks/hackers –yhteenliittymällä on alajaosto Suomessakin. H/H ilmoittaa missiokseen luoda toimittajien (hacks) ja teknologian tuntijoiden (hacker) verkosto, joka pohtii uutisten ja tiedon tulevaisuutta. H/H:lla on ”sormensa pelissä” digitaalisen journalismin kentässä. Johanna Vehkoon (2011, 93) mukaan Hacks/Hackers oli ensimmäinen merkki journalismin uudesta trendistä, jolle ei näy loppua.

2.1 Datajournalismin tavoitteet

Uutismedioiden verkkosivut eivät vielä hyödynnä digitaalisuuden tarjoamia mahdollisuuksia siinä mittakaavassa kuin olisi mahdollista. Ruotsissa asuva datatoimittaja Jens Finnäs toivoo, että mediat löytävät datajournalismimaailmasta uusia tapoja esittää juttuja ja hyödyntää digitaalisen alustan tuomat mahdollisuudet. ”Vaikka lehdillä ja toimituksilla on ollut verkkosivuja jo lähes 20 vuotta, ne näyttävät vielä suurin piirtein samalta kuin ovat aina näyttäneet: samat 3000 merkkiä ja kuva” (Finnäs 7.3.2013).

Viime vuosina Suomen valtakunnallisiin medioihin on rekrytoitu ns. uuden ajan journalisteja ja datatoimittajia. Helsingin Sanomissa aloitti kesällä 2012 kaksihenkinen datatoimitustiimi. Tammikuussa 2013 aloittaneen Yle Uutisten PlusDesk-toimituksen tehtävä on auttaa muun muassa uutis- ja urheilutoimituksia internetin mahdollisuuksia hyödyntävien sisältöjen toteuttamisessa. PlusDeskissä on kaksi graafikkoa, kaksi tuottajaa ja datajournalisti. Suomessa on vajaa 10 datatoimittajaa.

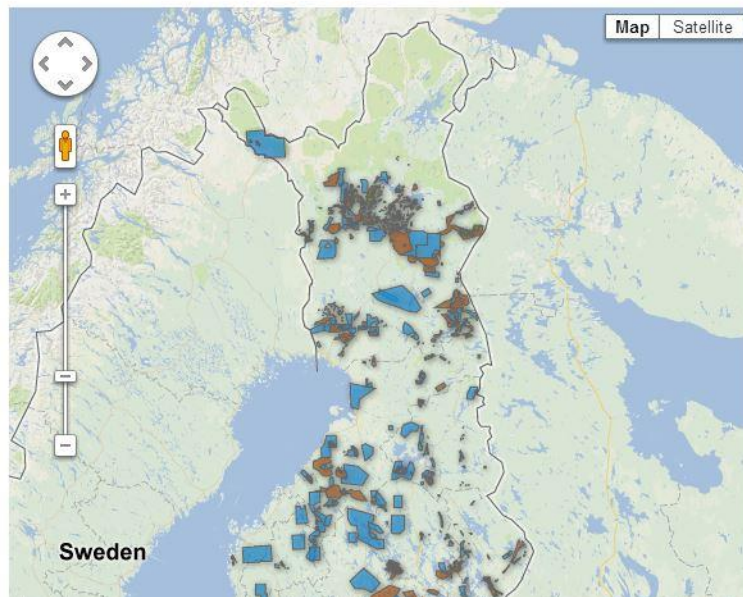
Datajournalismi on vielä terminä monelle toimittajallekin uusi, joten puheena olevaa journalismin suuntausta on tarpeen määritellä tarkemmin. Simon Rogersin sanoin: ”journalisti tutkii, ottaa selvää, kirjoittaa, raportoi, kiinnittää lukijoiden huomion ja paljastaa. Mitä sitten tekee datajournalisti? Samoja asioita, mutta työkalut ovat kokonaan muuttuneet”, Rogers jatkaa. (Rogers 2012) Finnäs nostaa esiin aiemmin kuulemansa määritelmän, että datajournalisti on kuin tavallinen toimittaja, mutta

muistiinpanot ovat Excelissä (7.3.2013). Datajournalismissa niin kuin muussakin journalismissa uutinen on pääosassa, lopputuotteen ei aina tarvitse olla visualisointi.

Tässä luvussa käydään läpi datajournalismin ja vuorovaikutteisen journalismin edelläkävijä-tiedotusvälineiden edustajien näkemyksiä siitä, mitä he ovat tehneet uudistuakseen ja houkutelakseen lukijoita. Nämä mediat ovat myös onnistuneet keräämään verkossa ns. pitkää häntää datajournalismillaan. New York Times on yksi arvostettu edelläkävijä datajournalismissa. Brittiläisessä Guardianissa on kolmen hengen datatoimitustiimi (tammikuu 2013), ja lisäksi toimituksessa on muun muassa vuorovaikutteinen tiimi sekä tutkivan journalismin tiimi. Guardian tuottaa paljon datajournalismia: sekä parissa tunnissa tehtäviä datavisualisointeja että pitkäjänteisempiä visualisointiprojekteja. Eniten aikaa vievät vaiheet datajournalistisessa prosessissa ovat datan hankkiminen, käsittely ja analysointi. Suuntaa antavana arviona journalistien kannattaa varautua siihen, että jutun kirjoittaminen ja julkaisu vievät vain noin viidesosan koko datajuttuun käytetystä työajasta (Poikola 26.2.2013 ja Tebest 2013).

Guardian julkisti vuonna 2012 linjauksensa avoimesta journalismista (open journalism policy), joka tarkoittaa asioita avoimesta käyttöliittymästä lisääntyneeseen vuorovaikutukseen lukijoiden kanssa. Rogers mainitsee ihmisten sitouttamisen ja osallistamisen yhdeksi digitaalisen tai avoimen journalismin tavoitteeksi (2012). Tästä erinomainen esimerkki on Guardianin ja London School of Economicsin yhteistyössä tekemä *Reading the riots* – juttukokonaisuus. Englannin kesän 2011 mellakoinnista uutisoidessaan Guardian keräsi tietokantaan muun muassa 2,5 miljoonaa Twitter- viestiä. Guardianin ja muun muassa tutkija Farida Vis'in tekemä visualisointi Twitter- viesteistä kertoi, miten mellakointia koskevat juurut levisivät. Tällaista vuorovaikutteista tunti tunnilta tarkasteltavissa olevaa visualisointia olisi ollut mahdotonta tehdä printtilehteen. Guardian onnistui mellakkauutisoinnissaan, sillä visualisoinnit olivat todella suosittuja. Farida Vis painottaa, että tekstuaalisen datan visualisointi ei ole ongelmaton; mainintojen esiintymistiheys ja asiayhteys ovat tärkeitä (2012). Rogers jatkaa, että Guardianissa yritetään saada ihmisiä osallistumaan jo siinä vaiheessa, kun tilanne on päällä. Muutoin riskinä on, että jutun valmistuessa ihmiset ovat jo menneet. Hänen mukaansa mellakoiden uutisointi eteni nopeasta monimutkaisempiin kerronnan muotoihin. (Rogers 2012)

Datajournalismilla voidaan näyttää monimutkainen ja laaja asia havainnollisesti. Tämä juuri on tiedon visualisoinnin vahvuus. ”Hyvin suunniteltu datagrafiikka on yleensä yksinkertaisin ja samalla vaikuttavin tilastollisen tiedon kaikista analysointi- ja viestintätavoista” (Tufte, 2001, introduction). “- - datajournalismi on yksi keskeinen keino, jonka avulla toimittajat ja kansalaiset voivat tuoda ilmi tietämättömyyttä ja haastaa ennakkoluuloja, saattaa monimutkainen tieto saavutettavaksi ja vaatia vallanpitäjiä vastuuseen (Bradshaw ja Rohumaa 2011, 49). Ns. tietokantajournalismin, ja erityisesti datajournalismin avulla toimittaja voi ikään kuin raportoida yksittäisiä uutisia ”pidemmälle” esittämällä uutisen laajemmassa asiayhteydessä antamalla näin yksittäiselle uutiselle merkityksen (Gray ym. 2012, 19). Datajournalistisen visualisoinnin avulla toimittaja voi esimerkiksi tuoda ison ilmiön lukijan omaan ympäristöön. Näin lukija voi itse sovelluksen tai visualisoinnin kautta tehdä omia johtopäätöksiä (Poikola 26.2.2013). Oheisessa Yle Uutisten *Tuoko ryntäys kaivoksen naapuriisi* -verkkokartassa/kartassa lukija voi tarkastella kaivoksiin liittyvä toimintaa kotiseudullaan (2013).



Kuva 1: Yle Uutisten *Tuoko ryntäys kaivoksen naapuriisi* -verkkokartta

Mediataloille datajournalismi kannattaa myös siksi, että datavisualisoinnit keräävät huomiota pitkän aikaa, ns. pitkä häntä (Gray ym. 2012, 52). Tätä näkemystä tukee myös toimintatutkimuksessa tehdyt havainnot yrittajat.fi:n ensimmäisestä datajournalismijutusta lokakuussa 2012 kunnallisvaalien aikaan. Yrittäjäkansanedustajat kunnittain esittelevä kartta oli suosittu Suomen Yrittäjien

Facebook-sivuilla (www.yrittajat.fi 2012). SY:n Facebook-sivulla on lähes 8900 tykkääjää, joten uutisjutut tai datajutut leviävät hyvin verkossa Facebook-sivun kautta.

Datajournalismissa vuorovaikutus toimittajien, graafisten osaajien ja koodareiden välillä korostuu. Yhdyn muun muassa Jonesin ja Salterin (2012, 172) näkemykseen, että vaatimukset toimittajan datajournalismiosaamiselle voivat jossakin aikaikkunassa tulla myös ylhäältä päin; toimittajilta aletaan pikku hiljaa edellyttää uusien osaamisalueiden haltuunottoa, muun muassa valmiutta opetella moninaisten tietotyökalujen käyttöä. Australian Yleisradioyhtiön (ABC) Wendy Carlislen mukaan suorituskykyyn on kiinnitettävä huomiota, jotta mediatalojen datajournalismivalmiudet paranevat. Carlisle jatkaa, että ”tarvitsemme uutisjohtajia, jotka ymmärtävät digitaalisuus/journalismi yhteyden ja alkavat investoida siihen”. (Gray ym. 2012, 28)

Datajournalismi on luonteeltaan tutkivaa journalismia, jossa pengotaan aihetta syvällisemmin. Suomen tutkivan journalismin yhdistys onkin kouluttanut jäseniään muun muassa Excel- ja datajournalismitaidoissa.

2.2 Vuorovaikutteinen journalismi

Päivälehti Helsingin Sanomat määritteli datajournalismin verkkosivuillaan seuraavasti käynnistäessään datajournalismitiimin toiminnan toukokuussa 2012. ”Datajournalismilla tarkoitetaan viime vuosina syntynyttä uudenlaista journalismin kenttää, jossa isoista tietoaaineistoista etsitään uutisia ja tehdään niistä uudenlaisia, verkossa toimivia visualisointeja.” (2012a) Helsingin Sanomat toi samassa yhteydessä esille joukkoistamisen, eli lukijoiden ottamisen mukaan jutuntekoon, olennaisena osana datajournalismia.

Vuorovaikutus eri muodoissaan on keskeinen osa tämän päivän journalismia. Bradshaw ja Rohumaa (2011, 123) sanovat asian ytimekkäästi: [Journalismin näkökulmasta] ”se tarkoittaa, että sisällöt laitetaan sinne, missä käyttäjä on, ei toisin päin.” Verkkjournalismi mahdollistaa lukuisia vuorovaikutteisuuden elementtejä, kuten Facebook- ja Twitter-nappuloiden lisääminen verkkouutisen loppuun jutun levittämiseksi. Myös datajournalismijuttu voi olla vuorovaikutteinen, jossa käyttäjä klikkaa auki visuaalisen esityksen tietoja. Britanniassa toimiva Help me investigate – sivusto on joukkoistamisalusta, jossa toimittajat ja kansalaiset voivat joukkoistaa

tutkittavaksi julkisen edun mukaisia kysymyksiä. Sivuston kautta tarjottuun apuun ovat turvautuneet useat brittiläiset mediatalot.

Mediatalot ovat Suomessakin toteuttaneet uutisjuttuja joukkoistamisen avulla. Mielestäni hyvä esimerkki on Helsingin Sanomien kiinteistönvälittäjien palkkioita kartoittanut juttuprojekti, jossa lehti kysyi lukijoiltaan välittäjien palkkioista verkkokyselyllä (2012b). Juttu näyttää todistavan, että välittäjät neuvottelevat palkkioitaan listahinnoista alaspäin. Joukkoistaminen tuntuu olevan myös lukijoiden keskuudessa hyvin vastaanotettu keino hankkia uniikkia dataa, pyytää apua datan siivoamisessa lukijoilta, tai selvittää ruohonjuuritasolta ponnahtavia juttuaiheita. Datajournalismi käytännössä -luku kertoo lisää vuorovaikutteisuudesta journalismissa. SY:n toimitus on muun muassa tehnyt lähemmän tarkastelun mahdollistavia verkkokarttoja yrittajat.fi-sivulle, mihin palataan tuonnempana.

Suomalaisten medioiden datajournalismia tiiviisti seuranneen Poikolan mukaan täkäläisissä datajournalismijutuissa vuorovaikutteisuuutta ei ole vielä paljon hyödynnetty. ”Usein vuorovaikutteisuus vaatii sitä, että tehdään sovellus verkkosivuille, vielä ei sellaista osaamista välttämättä ole. Esimerkkejä, joissa yleisö otettaisiin pitkälti mukaan, on aika vähän”, Poikola (26.2.2013) pohtii.

Bradshaw ja Rohumaa (2011) kiinnittävät huomion siihen, että tutkimusten mukaan käyttäjät viettävät enemmän aikaa sivuillasi tarkastellen vuorovaikutteista uutisjuttua. Kun vuorovaikutteinen uutisjuttu, kuten datajournalistinen juttu, on hyvin punnittu ja laadukas, se kerää käyttäjiä ja kävijöitä pidemmän aikaa verkossa kuin tavallinen tekstipainotteinen juttu (Bradshaw ja Rohumaa 2011, 41).

2.3 Avoin data ja journalismi

Suomessa datajournalismikenttään kuuluu niin toimittajia, koodareita kuin avoin demokratia - toimijoita, jotka pyrkivät edistämään julkishallinnon tietovarantojen avoimutta. Teemo Tebest (2013) nostaakin Datajournalismia valottavassa esityksessään esille, että datajournalismissa uutta on koodaajien ja tekniikan osaajien mukaantulo journalistiseen prosessiin. Avointa dataa vaativalla kansainvälisellä liikkeellä on ollut iso merkitys siinä, että hallitukset ympäri maailman ovat avanneet julkisten tahojen keräämää dataa kansalaisten saataville (Bradshaw ja Rohumaa, 2011, 49). Yksi avoimen datan ja datan saatavuuden parantamisen eteen

työskentelevä järjestö on Open Data Foundation. Apps4Finland järjestää vuosittain kilpailun, joka kannustaa kehittämään avointa dataa hyödyntäviä palveluja. Pysyvän Datavaalit-hankkeen tavoitteena on tukea vaaleihin kytkeytyvien tietovarantojen avautumista ja laajaa hyödyntämistä. Hanke oli aktiivinen syksyllä 2012 kunnallisvaalien alla. Sen nimissä järjestettiin muun muassa koulutussessio datajournalismityökaluista.

Helsingin Sanomat järjestää aika ajoin HS Open -hackathoneja (intensiivinen muun muassa uutissovellusten kehittämistäpahtuma). HS järjesti ensimmäisen HS Openin maaliskuussa 2011. HS Open -konseptissa tapahtuman alussa muodostettavissa tiimeissä on jokaisessa vähintään yksi toimittaja, graafikko ja koodari. Tampereen yliopiston tutkijan Auli Harjun mukaan HS on ostanut parhaimpia HS Openin tuotoksia ja julkaissut ne. Osallistujien motiiveja on monenlaisia: datajournalismin ja avoimen datan ymmärtäminen, verkostoituminen, työkalujen oppiminen, kiinnostus yhdessä tekemisessä saatavista tuloksista tai työnantajan pyyntö osallistua (Harju 2012). HSOpen konseptia on mahdollista soveltaa muissakin yhteyksissä; datajournalismin kehittämistäpahtuma edellä mainitunkaltaisella kokoonpanolla olisi toimiva malli esimerkiksi SY:n yrittajat.fin ympärille vietynä.

3 Datajournalismi käytännössä

Verkkoteknologioiden kehitys 2000-luvulla ja avoimen datan valtaisa kasvu mahdollistavat toimittajalle monipuolisen keinovalikoiman verkkojournalismissa. Johanna Vehkoon kahden vuoden takaisen teoksen (2011, 93) mukaan verkolle ominaisia kerronnan muotoja aletaan vasta kehittää. Nyt niitä nähdään jo runsaasti enemmän kuin tuolloin vuonna 2011. Tämä kertoo kehityksen nopeudesta verkkomaailmassa. Verkkotoimittaja voi käyttää raportoinnissaan muun muassa videota, animaatiota, vuorovaikutteisuutta, dataa, aikajanoja ja yhteisöllisiä elementtejä, kuten kyselyitä, foorumeita sekä blogeja (Bradshaw ja Rohumaa 2011, 41).

Millainen on suomalaisten toimittajien datajournalismin tietotaito alkuvuodesta 2013? Avoimen datan parissa toimiva konsultti Antti Poikola on kouluttanut toimituksia datajournalismin tekoon. Hän on yksi harvoista kouluttajista Suomessa.

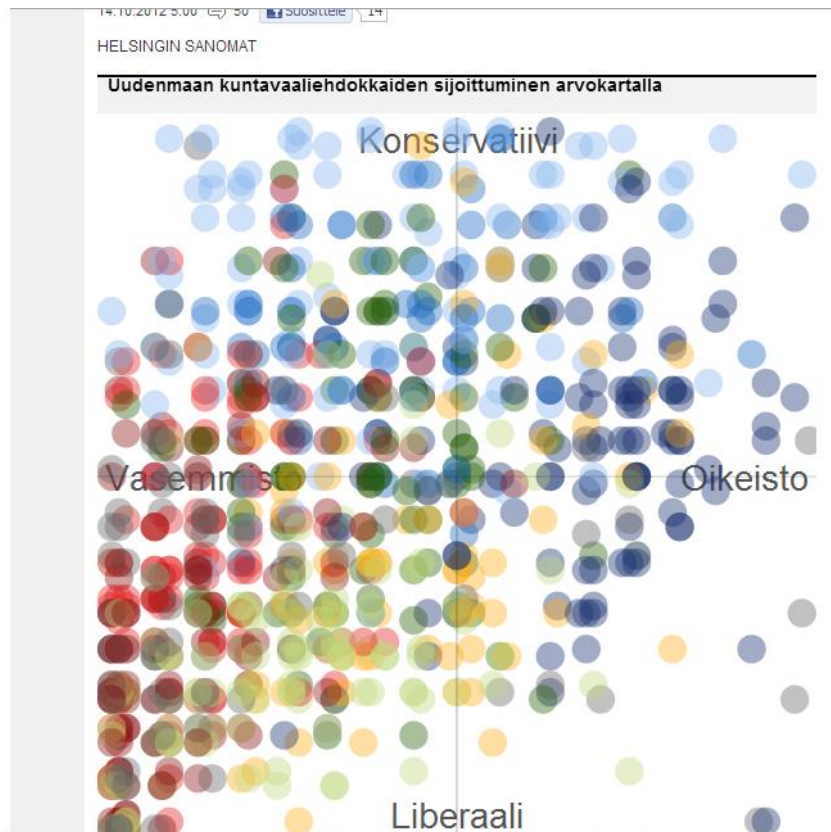
Poikolan mukaan koulutuksissa näkyy, että toimittajien perustyökulttuuriin ei kuulu datan käsittely ollenkaan.

Monet toimittajat ovat täysin alussa ja koskevat Exceliin ensimmäistä kertaa kymmeneen vuoteen. Silloin huolenaiheet pyörivät pitkälti siinä, miten datajuttu pystyy toteuttamaan, kun omat taidot ja kokemus eivät riitä. Osa on hiukan pidemmällä, silloin keskitytään siihen, mistä löytyy dataa jostakin tietystä aiheesta. (Poikola 26.2.2013)

Seuraava vaihe henkilön oppimisprosessissa on juttunäkökulman ja datan yhdistäminen.

Halusin kuulla myös Suomen eturivin tekijöiden näkemystä siitä, millaisissa taidoissa mediatalojen tulisi toimittajia kouluttaa. Riittääkö se, että toimittajille järjestetään koulutusta datan käsittelyssä ja visualisoinnissa? Vai pitäisikö kouluttautua myös muissa tarvittavissa taidoissa, kuten tilastotiede ja koodaus? Helsingin Sanomien datatuottajan Esa Mäkisen (15.3.2013) mukaan tilastolliset ja koodaamiseen liittyvät asiat kuuluvat olennaisesti datajournalismiin. Mutta alkuun pärjää työkaluilla, joita pystyy käyttämään ilman ohjelmointitaitoja, ja työkaluja tulee koko ajan enemmän. ”Edistyneemmän datajournalismin tekemisessä on hyötyä vähintäänkin ymmärtää, mitä voi tehdä ja mitä ei. Toimittajien pitäisi ymmärtää datajournalismin mahdollisuudet ja rajoitukset paremmin ” (Mäkinen 15.3.2013). Helsingin Sanomissa perinteinen grafiikkaosasto on aika syrjässä datajournalismin tekemisestä. Mäkisen mukaan datajournalismi lähti liikkeelle toimituksesta ja verkon kehitysrühmästä. (Mäkinen 15.3.2013) Guardianin datatuottaja Simon Rogers kannustaa toimittajia ajattelemaan dataa samalla tavalla kuin he ajattelevat muita uutisjuttuja. ”Toimittajia on hyvä rohkaista korostamalla sitä, että se on vain yksi numero ja juttu on pääasia. Taidot ja työkalut ovat opetettavissa. Uutisvainu on kaiken a ja o.” (Rogers 10.1.2013). Olen samaa mieltä Rogersin kanssa, numeraalinen tai muu data on uutislähde muiden joukossa; sen käsittelemiseen tosin tarvitaan muun muassa taulukkolaskennan taitoja. Uutisvainu yhdistyy taitoon löytää ja ”haistaa” uutisia, ja se on toimittajan avaintyökalu myös tänä päivänä. Korostan kuitenkin motivaation merkitystä uuden opiskelussa.

Helsingin Sanomien tuottaja Esa Mäkinen (15.3.2013) nosti HS:n parhaimmaksi datajuttuksi kunnallisvaaliehdokkaiden arvokartan, jonka lehti toteutti 14. lokakuuta 2012. *HS sijoitti kuntavaalien ehdokkaat arvokartalle mielipiteiden mukaan* –jutun visualisoinnissa näkyy kuntavaaliehdokkaiden sijoittuminen oikeisto-vasemmisto ja konservatiivi-liberaaliakselilla (kuva 2).



Kuva 2: Helsingin Sanomien HS sijoitti kuntavaalien ehdokkaat arvokartalle mielipiteiden mukaan –juttuun liittyvä arvokartta

Yle uutiset toteutti huhtikuussa 2013 *Me tiedämme missä asut* - visuaalisen jutun verkkoon. Kokonaisuuteen on upotettu myös trailer-video, joka tiivistää tarinan. Juttu on upea esimerkki digitaalisesta journalismista, joka hyödyntää netin tarjoamia mahdollisuuksia ja koodia. Juttu ei ole datajournalismia, mutta visuaalista journalismia se on. Jutun on toteuttanut yhdeksän hengen tiimi.

Datajournalismin sudenkuopat

Kun useimmat mediatilat vasta harjoittelevat datajournalismin tekemistä verkossa, on selvää, että välillä langetaan sudenkuoppiin. Siksi on hyvä pohtia, onko suunnitellulla jutulla oikeasti merkitystä. Toisaalta ongelma voi olla datajournalistinen toteutus liian epäselvänä, jolloin on selkeä vaara väärintulkintaan. Joskus datajuttu, kuten Yle Uutisten (Rigatelli 2013) *Tässä ovat Suomen parhaat terveyskeskukset* – verkkouutinen on kuvaus hetkestä, eikä se välttämättä puolen vuoden kuluttua pidä enää täsmälleen paikkaansa. Yhteiskuntatieteilijä Petro Poutanen (2013) huomauttaa, että datajournalismi on journalistille valtaisa osaamishaaste; ”haasteeksi nousee

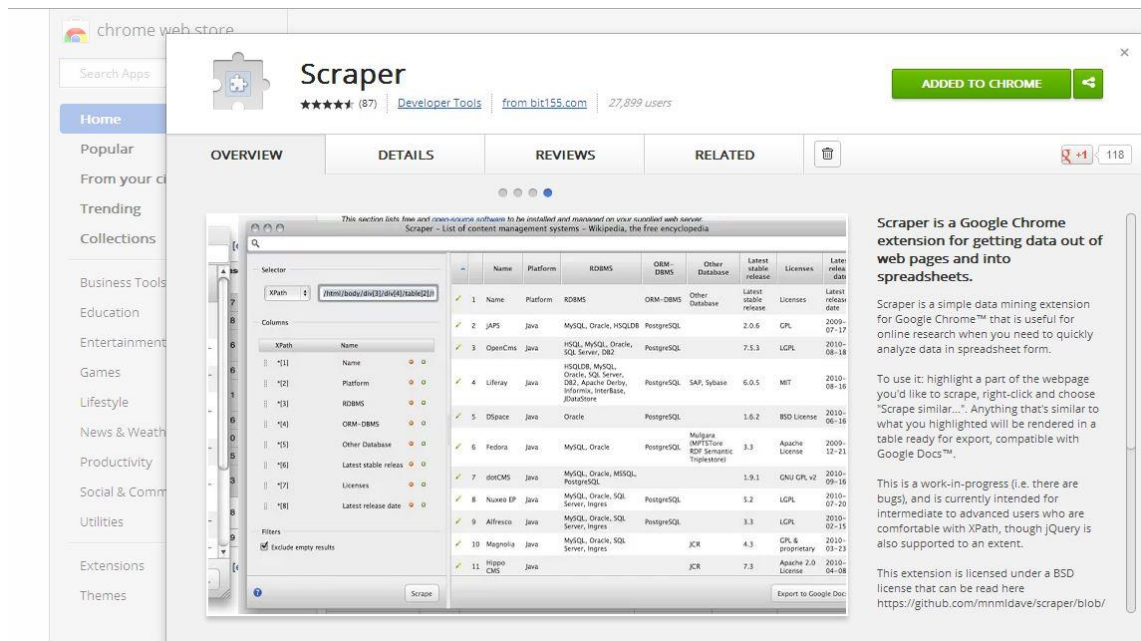
informatiivisten visualisointien tuottamisen ohella numerodatan, kuvien ja infografiikan kriittinen lukutaito”. Tilastotieteen alkeiden tuntemus on välttämätöntä.

3.1 Mistä dataa

Suomen hallitus teki maaliskuussa 2011 periaatepäätöksen julkisen tietoaineiston saatavuudesta niin, että datan tulisi olla avoimesti saatavilla ja uudelleenkäytettävissä yhtenäisin ehdoin. Vuonna 2012 muun muassa Maanmittauslaitos ja Verovirasto avasivat tietovarantojaan. Ilmatieteen laitoksen avoimen datan portaali on vielä beta-vaiheessa. Inspire-direktiivin myötä monet paikkatietoaineistot tulevat saataville yhtenäisessä muodossa kaikista EU:n jäsenmaista. Tilastokeskuksen StatFin-tietokannassa on Suomea koskevaa tilastotietoa eri aihealueilta. Tilastokeskuksen aluetietokanta Altika muuttui vuoden 2012 lopussa maksuttomaksi. Valtakunnan tasolla dataa on avattu viimeisen vuoden aikana reipas määrä, mutta kuntatasolla kehitys kulkee hitaammin. Helsinki Region Infoshare-hankkeessa on avattu Helsingin, Espoon ja Vantaan dataa, muun muassa tilinpäätöstietoja. Myös Helsingin kaupungin päätöksentekokoasiakirjat ovat saatavilla kaikille avoimena datana maaliskuusta 2013 alkaen. Muista kaupungeista Tampere avaa dataa kaikkien käyttöön, mutta lähes kaikkien muiden kaupunkien kohdalla datan hankkiminen on työläämpää, eli se on pyydyttävä tietopyynnöillä. Dataa esimerkiksi rakennushankkeista saa maksusta KTI Kiinteistötiedosta. Katso lisää Liite 2: Ohjeet datan hankintaan. Jos käytät hakukonetta, kuten Googlea datan etsimiseen, voit valita kehittyneen haun ja rajata hakutulokset tiedostomuotokohdasta koskemaan esimerkiksi Excel-tiedostoja.

Yllä lueteltujen avoimen datan lähteiden lisäksi dataa voi hankkia myös julkisuuslain mukaisilla tietopyynnöillä. Silloin on tiedettävä, mitä dataa ko. viranomaisella on, jotta voi tehdä selkeän tietopyynnön. Facebookissa on Tietopyyntö.fi -niminen sivu, jonka tavoitteena on auttaa toinen toistaan ja dokumentoida Facebookiin vietyjen tietopyyntöjen keskusteluketjut.

Avoimen datan ja tietopyyntöjen lisäksi dataa voi hankkia ruudunraavinnalla (screen scraping) ja joukkoistamisella, josta edellä jo mainittiin (liite 2). Helppokäyttöinen ruudunraavintatyökalu on Google Chrome -selaimen ladattavissa oleva scraper-lisäosa (kuva 3). Työkalulla voi kopioida määrämudossa olevaa dataa verkkosivuilta ja viedä se Exceliin tai Google Spreadsheet -taulukkolaskentaohjelmaan.



Kuva 3: Google Chrome - selaimen saatava Scraper-lisäosa

Kun data on pdf-muodossa, voi kokeilla siirtää tiedot 'kopioi ja maalaa' -toiminnolla Exceliin. Jos yksinkertainen tapa ei toimi, Adobe Acrobat Pro - ohjelma tai CometDocs -verkkotyökalu tekee muunnoksen toiseen formaattiin. CometDocsista on ilmaisversio.

Dataa voi hankkia ns. joukkoistamalla. Eräs tapa on pyytää lukijoita/käyttäjiä vastaamaan tutkittavaan kysymykseen verkkokyselyn avulla (liite 2). Edellä kerroin Helsingin Sanomien kiinteistövälittäjien palkkioita kartoittaneesta juttuprojektista, jossa lehti kysyi lukijoiltaan välittäjien palkkioista verkkokyselyllä. Katso, miten kysely tehdään Itä-Suomen Yliopiston ohjeista verkkokyselyn tai -lomakkeen tekemiseen Google Formsilla (2013).

3.2 Datajournalistin työkalut

Datajournalismin tekemiseen sopivia työkaluja on verkossa valtaisa määrä, ja lisää kehitetään koko ajan. Siksi osa ammattitaitoa on seurata alan verkkotyökalujen kehitystä (Mäkinen 15.3.2013). Ilman Excelliä ei kuitenkaan pärjää. Edellisessä kappaleessa kerrottiin datan hankkimisesta. Seuraavat vaiheet ovat: puhdistaminen/käsittely, analysointi ja visualisointi.

Tämän tutkimuksen tavoitteena on esitellä ja perehdyttää keskeisiin työkaluihin. Suosittelemme aloittamaan datajournalismin harjoittelemisen www.datajournalismi.fi -sivuilla olevista tutoriaaleista. Jens Finnäsin sivuilla on myös runsaasti tutoriaaleja,

mm. Google Chromen scraper-lisäosasta (<http://dataist.wordpress.com>). Esa Mäkisen kotisivuilta ja Teemo Tebestin (<http://datajournalismi.blogspot.fi>) blogista löytyy tutoriaaleja ja/tai toimintaohjeita. Googlen työkalujen etu on niiden helppokäyttöisyys ja ilmaisuus. The Data Journalism Handbook - verkkokirja tai perinteinen kirja on perusteellinen opus aloittelevalle datajournalistille.

Puhdista data

Google Refine on hyvä työkalu datan siivoamiseen, esimerkiksi kirjoitusvirheiden löytämiseen ja datan yhdenmukaistamiseen. Samoja asioita voi tehdä Sublime textillä, joka on maksullinen; tosin kokeilu on ilmaista. Muitakin työkaluja on runsaasti.

Analysoi data

Excel on ylivertainen työkalu datan analysoinnissa. Taulukkolaskentaohjelmassa voi järjestää dataa, suodattaa tietoja datan joukosta, ristiintaulukoida, havainnollistaa ym. Jyväskylän yliopiston avoimen taulukkolaskenta-kokonaisuuden avulla pääsee jo tutuksi Excelin kanssa (<http://appro.mit.jyu.fi/doc/tiedonhallinta/taulukkolaskenta/>). The Centre for Investigative Journalism -keskuksen verkkosivuilta löytyy *Data Journalism or Computer Assisted Reporting* - käsikirja. Eturivin datatoimittajat Suomessa ja muualla kertovat Excelin olevan tärkein datatoimittajan työkalu (Esa Mäkinen; Simon Rogers, 2013). Mäkinen käyttää Access-tietokantaohjelmistoa datojen yhdistelemiseen (15.3.2013).

Visualisoi data

Helppoja työkaluja visualisointiin ovat Tableau Public ja Google Fusion Tables. Excel on myös hyvä visualisoinnissa, mutta Excelillä tehdyt visualisoinnit eivät ole suoraan vietävissä verkkoon. Guardianin Simon Rogers korostaa datajournalismin tekemisen helppoutta: ”Verkossa on lukuisia työkaluja, sukella vain mukaan. Toimittaja, joka osaa koodata, on vahvemmassa asemassa. Jos tavoitteena on tehdä upeaa datajournalismia, silloin tarvitaan koodaajaa” (Rogers 10.1.2013). Rogers käyttää visualisointiin etenkin Google Fusion Tablesia ja Datawrapperiä (10.1.2013). Datawrapper ei toimi Suomessa.

Esa Mäkinen on koodaava datatuottaja, ja tämä näkyy hänen käyttämässään työkalupakissa. Hän käyttää paljon html- ja JavaScript-ohjelmointikieliä (Mäkinen 15.3.2013). ”Osa vanhoista työkaluista, kuten Googlen Fusion Tables, on tekniikaltaan jäämässä jälkeen; sen käyttäminen ei enää ole kovin perusteltua ollen viimeiset kaksi vuotta beta-versiona. Samojen asioiden tekemiseen on parempia työkaluja tarjolla” (Mäkinen 15.3.2013).

Abelan (2006; liite 3) kaavio-ohje kertoo, mitkä kuvaajat sopivat eri tilanteisiin. Ohjeessa lähdetään liikkeelle siitä, halutaanko kuvata muuttujien suhdetta, verrataanko lukuja keskenään, jakaumaan tai esimerkiksi osuutta kokonaisuuteen. Millainen visualisaatio sitten on hyödyllisin kuvaamaan kulloistakin datasettiä? Gregor Aischin (Grayn ym mukaan, 2012, 165) mukaan melkein aina on järkevää käyttää jotain muuta visualisointimuotoa kuin taulukkoa. Näin siksi, että taulukko ei sellaisenaan kuvaa datassa mahdollisesti olevia rakenteita (Aisch 2012, Grayn ym mukaan, 2012, 165).

Erilaisiin visualisointimuotoihin kannattaa tutustua työkalujen, tutoriaalien ja ohjeiden kautta. Esimerkiksi Tableau Publicilla voi tehdä monenlaisia visualisointeja, ja ohjelma osaa ehdottaa sopivia visualisointeja, kun olet ladannut tietokantasi ohjelmaan.

Myös tekstiä voi visualisoida perinteisesti sanapilven tapaisella linkkipilvellä. Tämä voi olla hyödyllinen, kun halutaan kuvata poliitikon puheen teemoja tai sanojen esiintymistä datavirrassa. Sanapilvien avulla voi vertailla esimerkiksi nykyisen ja edellisen presidentin puheita. (Bradshaw ja Rohumaa 2011, 62) Bradshaw ja Rohumaa mukaan Twitter-syötteistä voi tutkia ihmisten reaktioita verkossa johonkin aiheeseen.

Poliittisen tieteen ja tilastotieteen professori Edward Tufte (2011, 79) kritisoi grafiikoiden liiallista ”koristelua”. Hän uskoo sen johtuvan siitä, että monet graafikot pitävät tilastoja tylsinä. Tuften mukaan vaarana on, että liioitellaan datan todistepohjaa. (Tufta, 2011, 79). Journalismissa tarinan pitää olla pääasia. Simon Rogers painottaa datajournalismissa ymmärrettävyyttä. Guardianin visualisoinneista iso osa kuvaa yksinkertaisia yhteyksiä datassa, mihin riittää, että osaa laskea prosentuaalisen muutoksen ym. Esimerkiksi korrelaation kuvaaminen on monimutkaisempi visualisointi, jota kaikki eivät välttämättä ymmärrä (Rogers 10.1.2013). Alberto Cairo (2013, 62) kuvaa tiedon visualisointia koskevassa teoksessaan em. ”kädenvääntöä” visualisaatiopyöräksi kutsumallaan mallilla. Cairon (2013, 62) mukaan toimittajat,

graafiset suunnittelijat ja taiteilijat painottavat helppotajuista ja pintapuolista visuaalista esitystä päinvastoin kuin tiedemiehet (kuva 4).

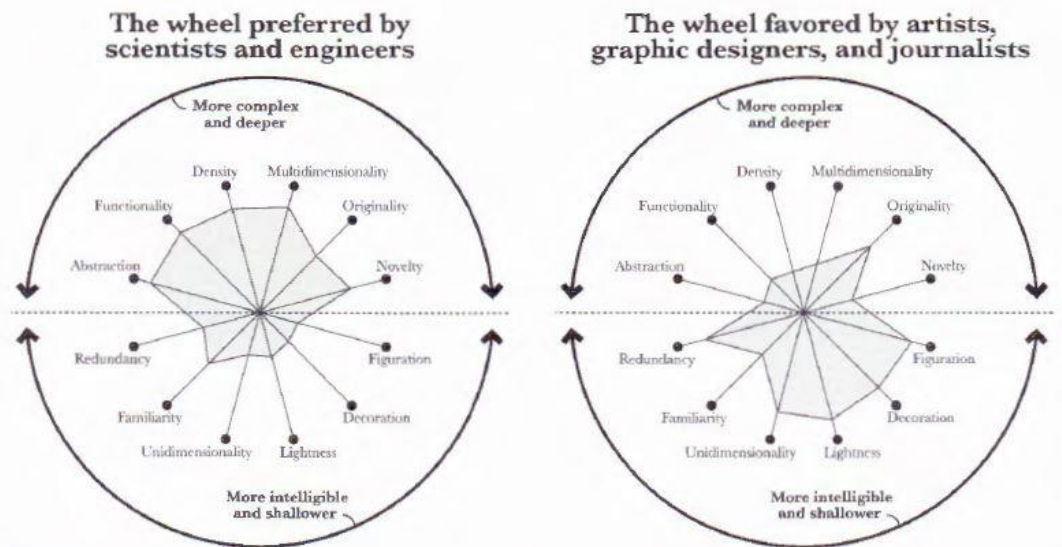


Figure 3.1 Different professional backgrounds, different ways of facing projects.

Kuva 4: Alberto Cairon visualisaatiopyörät kuvaavat eri ammattiryhmien lähestymistapaa

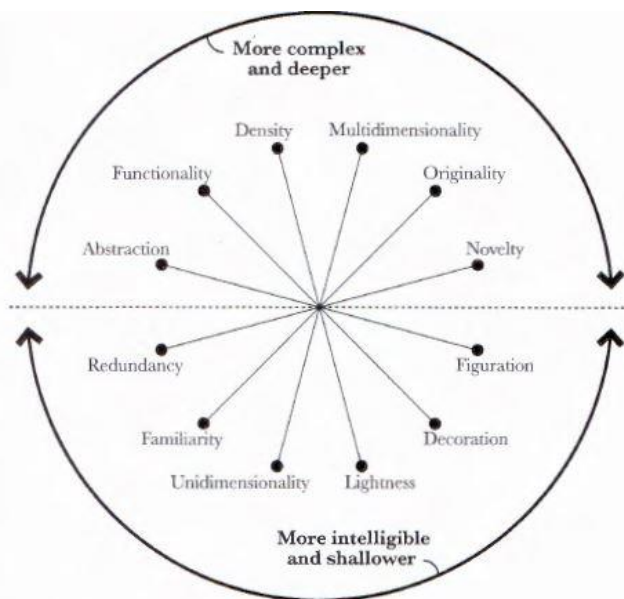


Figure 3.2 The visualization wheel.

Kuva 5: Alberto Cairon visualisaatiopyörä

Visualisaatiopyörän avulla voi arvioida omaa suhtautumistaan datajournalismin visualisointeihin tai ylipäätään tiedon visualisaatioon.

Kun ensimmäiset datajournalismiharjoitukset ja -jutut on tehty, datajournalismin viemistä pidemmälle helpottaa koodaamisen jonkinasteinen lukutaito. Toimittajalle on

hyödyllistä ymmärtää hiukan koodaamista, ja jos kiinnostusta riittää, opetella koodaamaan itse. Jos toimittaja ymmärtää ja osaa koodaamisen alkeet, hänen on helpompi kommunikoida koodareiden kanssa osaten ainakin muokata koodia tarvittaessa. Teemo Tebest suosittelee aloittamaan koodin opettelu JavaScriptistä (2013). Jens Finnäs (7.3.2013) suosittelee koodaamisesta kiinnostuneille toimittajille amerikkalaisen toimittajan Dan Nguyenin oppimateriaalia verkossa (<http://danwin.com/2010/04/coding-for-journalists-101-a-four-part-series/>).

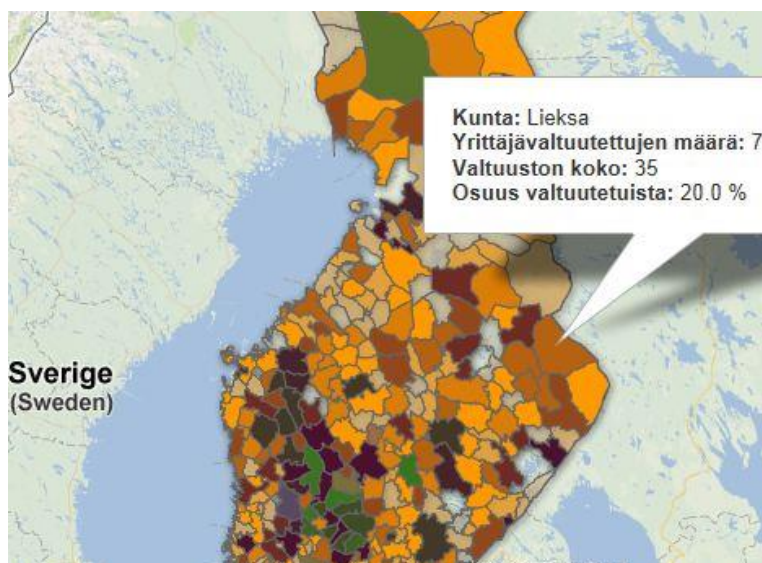
3.3 Datajournalistinen työprosessi ja arviointi

Viiden hengen toimituksessa työnteon rytmi on hyvin hektinen. Tällöin osaamisen parantamiseen tähtäävän kehittämishankkeen on mahdollista niihin resursseihin, jotka on mahdollista saada. Osaamisen kehittäminen on paljon henkilön motivaatiosta kiinni. Kevät 2013 ei ole ollut ajankohtana paras mahdollinen työyhteisöni kehittämishankkeelle, sillä toimituksessa oli ajankohtana sijaisjärjestelyjä. Mutta kaikesta huolimatta tuloksia on tullut.

Sain toimituksen innostumaan datajournalismista, kun aloitimme hankkeen marraskuussa 2012 ja nyt huhtikuussa 2013 itseni lisäksi yksi kollega on ”tahrinut” sormensa dataan ja toteuttanut datavisualisoinnin. Kehittämishankkeen yhtenä tavoitteena on, ettei datajournalismin tekeminen ole vain yhden ihmisen varassa. Datatyökalujen opiskelu on jokaisen oman aktiivisuuden varassa. Olen iloinen, että toimituksessamme on näkynyt enemmän Excel-taulukoita kuluneen puolen vuoden aikana. Eräs osin kehittämishankkeen tuoma ansio on, että mietimme sopivien juttujen kohdalla, millaista dataa aiheesta on saatavilla. Tosin kaikissa näissä tapauksissa kyse ei ole ollut varsinaisesti datajournalismista, mutta datan hankinta on lisääntynyt joka tapauksessa, mikä on tavoitteiden suuntaista kehitystä.

Ensimmäinen datajuttumme koski kunnallisvaalien tulosten selvittelyä yrittäjänäkökulmasta. Teimme verkkokartan, joka sisältää tiedot yrittäjävaltuutettujen osuuksista kunnanvaltuustoissa jokaisessa Manner-Suomen kunnassa.

Juttu ilmestyi printissä samana päivänä aukeaman kokoisena karttana.



Kuva 6: Yrittäjävaltuutettujen osuus kunnanvaltuustoissa (lähde: yrittajat.fi)

Yrittäjävaltuutettuja koskeva datajournalistinen juttu oli SY:n toimitukselle hyvä harjoitus, joka käsitti vaiheet datan siivoamisesta ja analysoinnista datan visualisointiin. Tässä jutussa tekijöitä oli kaksi, raakadatan saimme järjestöosastoltamme, missä tiedot oli kerätty. Tällä jutulla saavutettiin myös datajournalismille ja ”hitaalle” journalismille asettamamme tavoite erottautua muista medioista ja muiden medioiden jutuista. SY:n toimituksessa ”hitaita” datajournalismijuttuja tehdään aiheista, joita voidaan hyödyntää niin verkossa kuin printissä. Yrittäjäsanomien ja yrittajat.fin päätoimittaja Kimmo Koivikko korosti yrittäjävaltuutettujen kuntakohtaisia tietoja esittävän verkkokartan juttunäkökulman onnistumista.

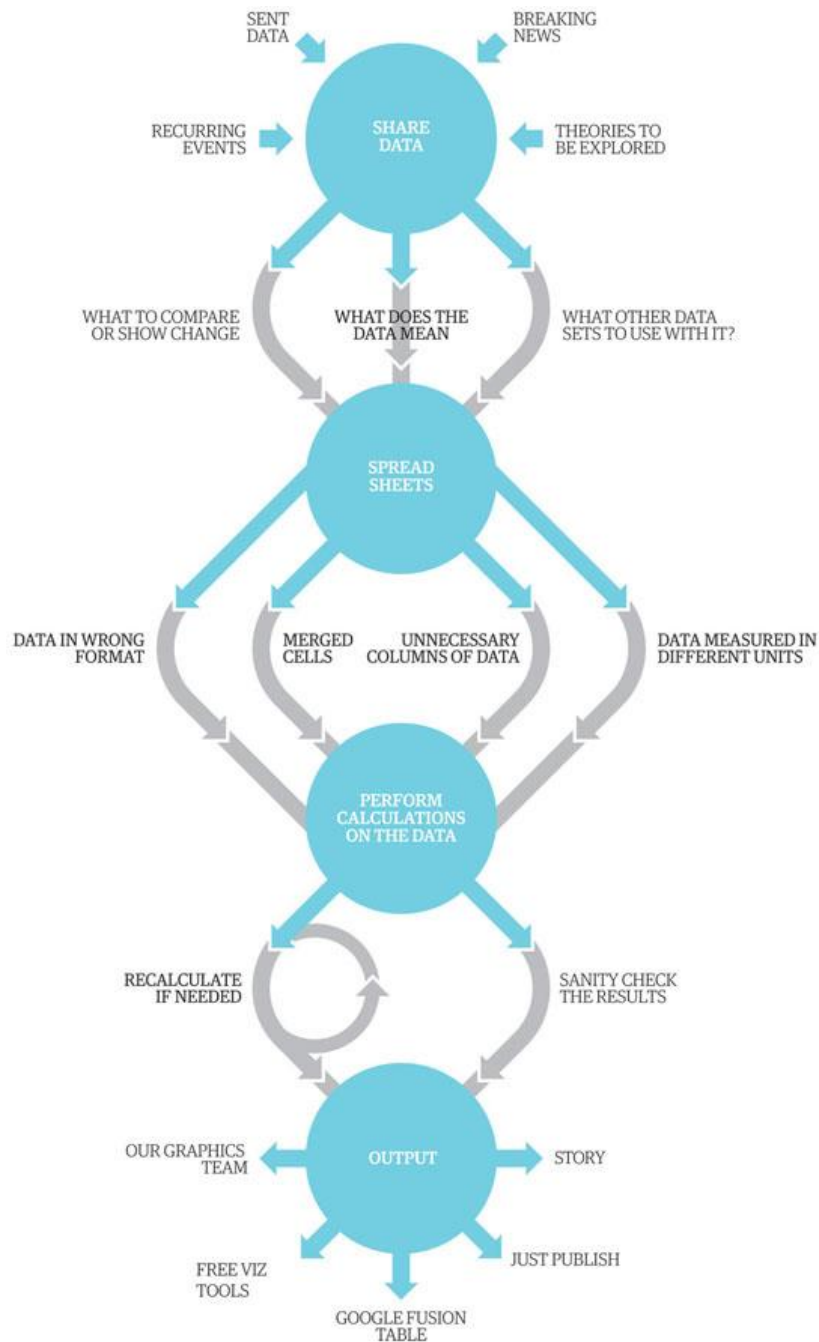
Kuntadata oli tietokantana oivallinen, koska meillä oli kaikista kunnista materiaali. Lehden funktio oli nostaa tiettyjä ilmiöitä esille, esimerkiksi missä kunnassa on eniten naisvaltuutettuja ja miksi ym. Verkossa jokainen pystyi katsomaan, miltä tilanne näyttää omassa kunnassaan. (Koivikko 2013)

Datajournalismijutuissa korostuu toimittajan ja graafikon yhteistyö, joskus prosessin jossakin vaiheessa mukana voi olla myös koodaaja.

Työprosessi

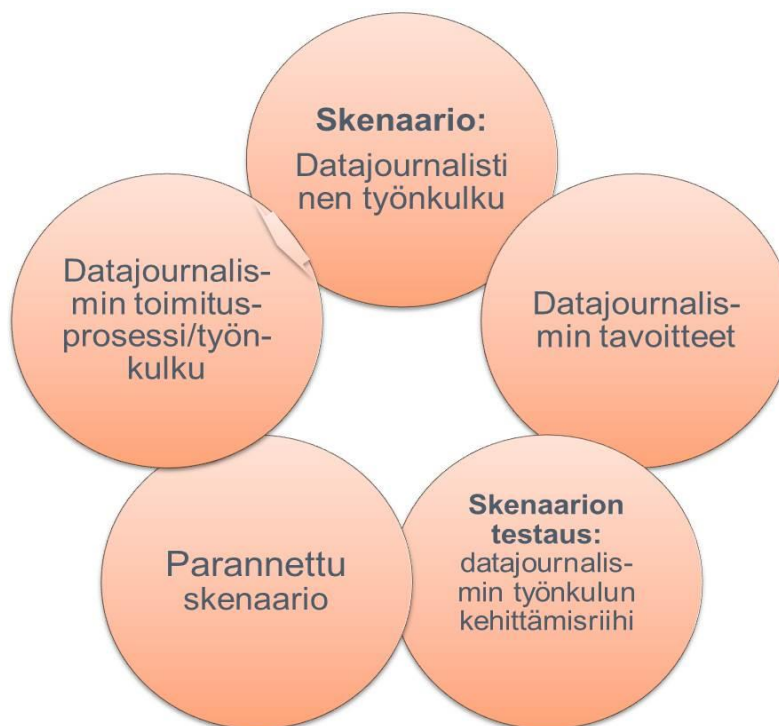
Käyn läpi hyväksi havaitun Guardianin datatiimin toimitustyön prosessin vaiheet (guardian.co.uk 2013). SY:n datajournalistisissa jutuissa Guardianin malli on todettu päteväksi. Tosin emme järjestömediana ole kokeneet tarpeelliseksi tässä vaiheessa avata ja jakaa käyttämäämme dataa Guardianin tapaan. Opinnäytetyön liitteenä on

avoin data - aktiivi Antti Poikolan Journalisti-lehdelle Guardianin mallista muokkaama toimitusprosessin kuvaus (liite 4).



Kaavio 1: Guardianin Datablogin toimituksellisen prosessin visualisaatio (The Guardian)

John M. Carroll (1999) kirjoittaa tieteellisessä artikkelissaan skenaariopohjaisesta suunnittelusta ja siitä, miten em. suunnittelun perinteessä käytetään ihmisten tehtävien suorittamisen kuvauksia suunnittelussa. Käytin toimintatutkimuksen yhtenä menetelmänä datajournalistisen työprosessin skenaariota, jonka esittelin marraskuussa kehittämishankeen aloituspalaverissa.



Kaavio 2: Skenaariomenetelmä SY:n kehittämishankkeessa

Poimin tässä keskeisiä SY:n toimituksen näkemyksiä datajournalistisen työprosessin suuntaviivoista. Toimintatutkimuksen arvioinnin läpikäynti jatkuu luvussa 4. Näkisin, että useissa pienissä uutistoimituksissa tai järjestömedioissa lienee samankaltainen lähtöasetelma ja tämän kehittämishankkeen opit ovat hyödyksi myös muille. Toimitusprosessi (kaaviossa työnkulun skenaario) lähtee liikkeelle joko ideasta, jota halutaan tutkia tai valmiista datasetistä, jota tutkitaan. Totesimme, että jutunteon elinkaari vaihtelee ja riippuu muun muassa datan hankinnan kestosta. Suunnitelmallisuus alusta alkaen on tärkeää juttuprosessin nopeuttamiseksi. Helsingin Sanomien datatiimi tekee sekä yhdessä päivässä datajuttuja, joissa reagoidaan uutistilanteeseen, että kuukausittain pari isoa viritystä, jotka saattavat vaatia useita viikkoja työtä (Mäkinen 15.3.2013). Datan pyörittely ja analysointi voi johtaa myös siihen, ettei juttua löydy. Gregor Aisch (Grayn ym mukaan, 2012, 165) suosittelee lähestymään dataa tutkimalla, mistä se kertoo ja millaisia oivalluksia dataa analysoimalla ja visualisoimalla huomaa. ”Hyvä journalisti voi taidokkaasti punoa datasta löytyneet ’oivallukset’ tarinaksi” (Aisch 2012, 165). Esa Mäkinen (15.3.2013) muistuttaa, että tietysti joskus käy niin, ettei datasta löydy tarinaa. Sen huomaa yleensä aika nopeasti; silloin siirrytään seuraavaan juttuun (Mäkinen 15.3.2013). Toisaalta usein datasta jää jotain myöhempää varten (Finnäs 7.3.2013). Omat kokemukseni ovat vielä vähäisiä, mutta julkaistavaksi päätyneissä jutuissa datan hankintavaiheessa on tiedetty suunnilleen, mitä datasta voi löytyä. Toki em. kuntia koskevien datasettien

kuntatason tarkastelussa saattoi eteen tulla ”yllätyksiä” analysointi- ja visualisointivaiheissa. Tällä tarkoitan, että dataseiteissä tiedettiin olevan varmaa juttuaineista.

Toimituksemme kokee suunnitelmallisuuden tärkeäksi myös visualisointien suunnittelussa. ”Pääasia visualisoinnissa on, että lukija pääsee hyödyntämään jutun maksimaalisesti” (Koivikko 2013) Kun data on hankittu, siivottu ja analysoitu, 80 prosenttia prosessiin kuluvasta ajasta on mennyt. Jutun visualisointi ja kirjoittaminen ovat tavallisesti prosessin nopeimmat osat toteuttaa. Keskustelimme toimituksen mahdollisuuksista ostaa ulkopuolista koodausosaamista jutuntekoon. Viesti oli, että lähtökohtaisesti pyrimme välttämään tätä, sillä ulkopuolisen perehdyttäminen juttuun vie myös aikaa. Silloin kun päätetään palkata henkilö tekemään joku osuus kokonaisuudesta, työn pitää olla jatkohyödynnettävissä ja monistettavissa myöhemmin (Pärnänen 2012). Tämä voisi tarkoittaa verkossa juttuun liittyvän avoimeen lähdekoodiin pohjaavan laskurin tms toiminnallisuuksien tai ulkoasun muokkaamista koodaamalla.

Korostimme toimittaja-graafikko-työpariaselman keskeisyyttä, jotta toteutukset verkkoon ja printtiin ovat parhaat mahdolliset. Realismia kuitenkin on, että työlään datajournalismijutun pitää tuottaa sen verran enemmän lisäarvoa normaaliin juttuun verrattuna, että sitä kannattaa pienilläkin resursseilla tehdä (Koivikko 2013). Verkkovisualisointien suunnittelussa otetaan huomioon interaktion toteuttaminen. On mietittävä, mitä käyttäjä haluaa esityksessä tehdä. (Pärnänen 2013)

Uudelleenkäytettävyys ja systemaattisuus

Jens Finnäs (7.3.2013) ja Antti Poikola (26.2.2013) neuvovat toimittajia miettimään seuraamiltaan aihealueilta, millaisia tietokantoja he voisivat rakentaa. Finnäs (7.3.2013) vinkkaa, että esimerkiksi kaikki kulttuuriarvostelut voisi tallentaa tietokantaan, kuten myös yritysten vuosikertomuksia ja urheilutuloksia. Uudelleenkäytettäviä aiheita ovat sellaiset, jotka toistuvat: vaalit, olympialaiset, säätiedotteet. Näitä systemaattisesti keräämällä voi tehdä jutun hyödyntämällä kerättyä dataa, kun aihe tulee uudelleen ajankohtaiseksi. Näin esikäsittelytyötä voi hyödyntää useammin ja datajournalismin vaatima työmäärä alkaa myöhemmin ruokkia lisää juttuja. (Poikola 26.2.2013) Toimituksen keräämissä tietokannoissa on runsaasti mahdollisuuksia tehdä täydennysjuttuja myöhempänä ajankohtana, kun dataa on

kerätty systemaattisesti. Kun tiettyjä aihealueita koskevia datasettejä on kerääntynyt riittävästi, niitä jatkokäyttämällä ja yhdistelemällä on saatavissa kokonaan uusia ja yllättäviäkin näkökulmia aiheisiin. Tällaisia tullaan näkemään jatkossa enemmän, sillä datajournalismin tekeminen on Suomessa varsin nuori journalismin muoto ja vie aikansa, että toimituksille kertyy datavarastoja. Yhtenä esimerkkinä datajuttujen päivittämisestä ja jatkokäytöstä voi mainita Ylen ja hs.fi:n vaalikoneiden vastausdatan päälle tehdyt jutut tyyliin 'katso toteutuivatko vaalilupaukset'.

4 Analyysi datajournalismijutusta, case: kaupunkien sotepalveluostot

Helmikuun loppupuolella ideoimme toimituksen ja verkkotuottajan kesken potentiaalisia datajuttuaiheita; päädyimme toteuttamaan omaan datan hankintaan pohjaavan jutun maakuntien keskuskaupunkien ja pääkaupunkiseudun kaupunkien sosiaali- ja terveysterveystoista yksityiseltä sektorilta. Kehittämishankkeen aikana varsinaisia palaverieja oli neljä kappaletta, sillä iso osa työyhteisön välisestä keskustelusta oli ad hoc - luonteista. Päätimme toimituksen palaverissa, mitä haluamme kysyä kaupunkien sosiaali- ja terveysjohtajilta. Halusimme mukaan numeraalisen datan lisäksi kaupunkien linjaukset siitä, paljonko tänä vuonna vähennetään tai lisätään ostoja yksityiseltä sektorilta sekä mitä palveluita tullaan ostamaan enemmän. Pyydetävät tiedot olivat senkaltaisia, että vuoden 2012 luvuista vastaukset sai helpoiten kaupungin talouspäälliköltä, mutta linjaukset tai alustavat linjaukset kuluvaan vuoteen ostoista taas piti kysyä sosiaali- ja terveysjohtajalta (tehtävänimikkeet vaihtelevat). Toivoimme, että kysymyspatteristo tuo meille näkökulmaa siihen, miten sosiaali- ja terveysterveystoimintamarkkinat ylipäättään kohdentuvat.

Datan hankkiminen oli pitkä prosessi. Aluksi tavoittelimme vastauksia Suomen keskisuurilta ja suurilta kaupungeilta sekä sotekuntayhtymiltä webropol-kyselyllä, mutta vain reilut 15 prosenttia tietopyynnön saajista vastasi. Vastaajajoukossa oli mukana 67 henkilöä, pienen vastausprosentin vuoksi jouduimme muokkaamaan mukaan otettavien kaupunkien määrää. Näin siksi, että kyseinen juttu vietiin läpi yhden henkilön työpanoksella muiden töiden ohella. Tämänkaltaisia muutoksia juttuprosesseihin tulee jatkuvasti, silloin katsomme, miten menettelemme, jotta alkuperäinen juttuidea ei vesity. Päädyimme ottamaan mukaan maakuntien keskuskaupungit ja pääkaupunkiseudun kaupungit, mikä kuitenkin tuo maantieteellisesti mielekkään kattavuuden. Tästä eteenpäin hankimme datan sähköpostiyhteydenotoin ja puhelinsoitoin, joista jälkimmäinen oli toimivampi tapa tavoittaa ihmiset ja saada haluttu

tieto. Jos data kerätään kyselylomakkeella, kuten Google Formsilla, suunnitteluvaiheessa on tarpeen miettiä, saadaanko haluttu data monivalintakysymyksillä. Monivalintakysymyksiä voi käsitellä Excelissä, vapaiden vastauskenttien vastausten käsittely on asia erikseen.

Datajutun elinkaari on tavallista pidempi, sitä voi julkaisun jälkeen päivittää lisäämällä siihen uutta tietoa (Poikola 2013). Jotta tehty datajournalismijuttu saa maksimaalisen näkyvyyden verkossa, on syytä miettiä, miten voidaan jakaa jutun sisältämää tietoa. Yksi tapa on jakaa jutun yksityiskohtia, kuten paikallista tietoa (verkkokartalla kaupunkikohtaisia tietoja) esimerkiksi Facebookissa. Toimituksen keräämä data kannattaa jalostaa tietokannaksi, jota päivitetään kyseistä aihetta koskevilla uusilla tiedoilla, kun sellaisia on saatavilla (Poikola 26.2.2013; Finnäs 7.3.2013). Näin päätimme toimia Yrittäjäsanomissakin soteostoja koskevan datan kanssa. Datajournalismi alkaa ruokkia itse itseään sitten, kun toimitukselle kertyy enemmän datasettejä, jolloin voidaan tehdä täydennysluonteisia jatkojuttuja.

Anttilan mukaan (2007, 47) kehittämisprosessi kulkee sykleittäin eteenpäin. Kehittämisprosessille on ominaista, että syklin aikana tehdään ”havaintoja asian sujumisesta” ja tavoitteiden suuntaisia korjausliikkeitä (Anttila 2007, 47). Myös toimituksen kehittämishanketta on arvioitu tulosten näkökulmasta ”kierroksen” (suunnittelu, toiminta, havainnointi ja reflektointi) saavutettua maaliviivan. Arviointipalaverissa huhtikuussa pohdimme datajournalismihankkeessa saavutettuja tuloksia ja oppeja ”summatiivisen arvioinnin” kaltaisesti. (Anttila 2007, 47) Keskeisimmät havainnot Yrittäjäsanomien näkökulmasta ovat: suunnitelmallisuuden tärkeys alusta alkaen, sillä sitä vähemmän juttuprosessi vie työaikaa. Toisena toimitus korosti printin ja verkon funktioiden pohtimista, niin että lisäarvoa syntyy kummassakin kanavassa. Tämä onnistui tyydyttävästi. Kolmantena havaintona oli, että dataa sisältävät aiheet on otettava mukaan toimituksen juttupalaverihin, että taitomme karttuvat ja tuotamme säännöllisesti datajournalismijuttuja. Teimme jutut eri näkökulmilla printtiin ja verkkoon. Säästimme verkkojuttua varten osan datasta, sillä lehdessä on rajattu tila ja näin saimme uutta asiaa verkkojuttuun. Soteostoja koskevan visualisoinnin miinuspuolena on, että kahden kaupungin antamat vastaukset olivat puutteelliset, niistä puuttui arvio tämän vuoden ostoista. Vielä peräänantamattomampi soittaminen olisi voinut toimia. Kimmo Koivikko (2013) nosti esiin havaintoja omasta datan hankkimisesta. Datan hankinnassa on laitettava tärkeysjärjestykseen, millaista

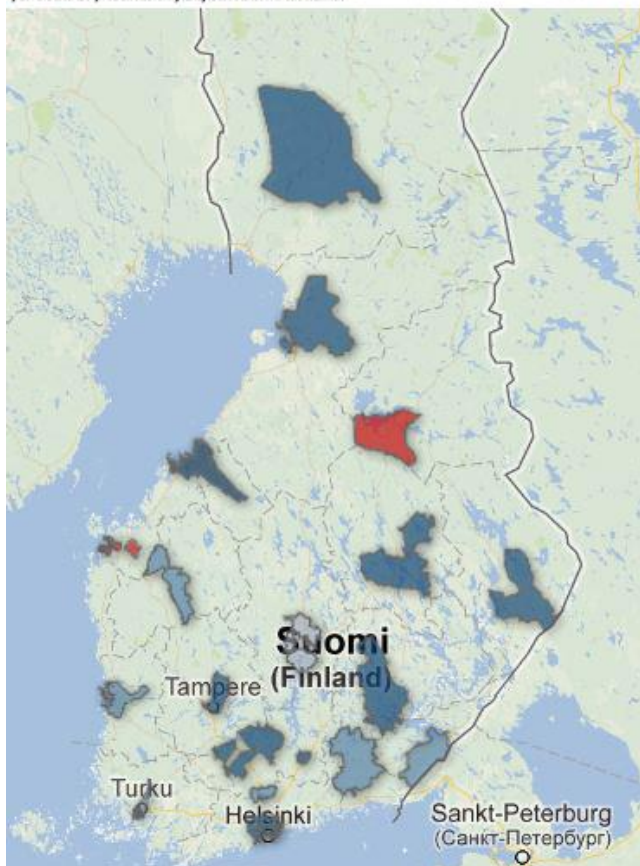
otantaa ryhdymme tavoittelemaan, ja otettava huomioon toimituksessa olevat resurssit (Koivikko 2013).

Juttua ideoidessamme tiesimme, että kunnat julkaisevat tilinpäätöksensä edellisvuodelta maaliskuuhuhtikuussa. Juttu oli ilmestyessään huhtikuun Yrittäjäsanomissa ajankohtainen ja hyvin tuore. Lisäksi se kosketti koko Yrittäjäsanomien kohderyhmää tavalla tai toisella.

25.4.2013

Näitä palveluita kaupungit aikovat ostaa eniten yksityisiltä - katso kartasta

Yrittäjäsanomien tekemän selvityksen mukaan maakuntien keskuskaupunkien primus on Kokkola, jonka sote-palveluista 26 prosenttia on yksityisen sektorin tuottamia.



Kuva: Maakuntien keskuskaupunkien soteostot - verkkokartta (lähde: www.yrittajat.fi)

kiinnittää huomiota. Datajournalismi on yksi digitaalisen journalismin käytössä olevista keinoista tehdä perusteellista journalismia, jota lukijat arvostavat. Esimerkki journalistisesti mielenkiintoisesta datavisualisoinnista, on hs.fi:n tekemä ”Lapsia yhtä päiväkotipaikkaa kohden Helsingin kaupunginosissa” kuvaava verkkokartta. Helsingin Sanomat yhdisti väestötietoja ja päivähoitopaikkojen tietoja koskevat tietokannat toteuttaakseen helppolukuisen ja monitahoisen jutun. Esa Mäkisen mukaan (15.3.2013) tulevina vuosina tehdään entistä parempaa datajournalismia, koska dataa tulee enemmän saataville, ja yhdistelmistä syntyy mielenkiintoisia asioita.

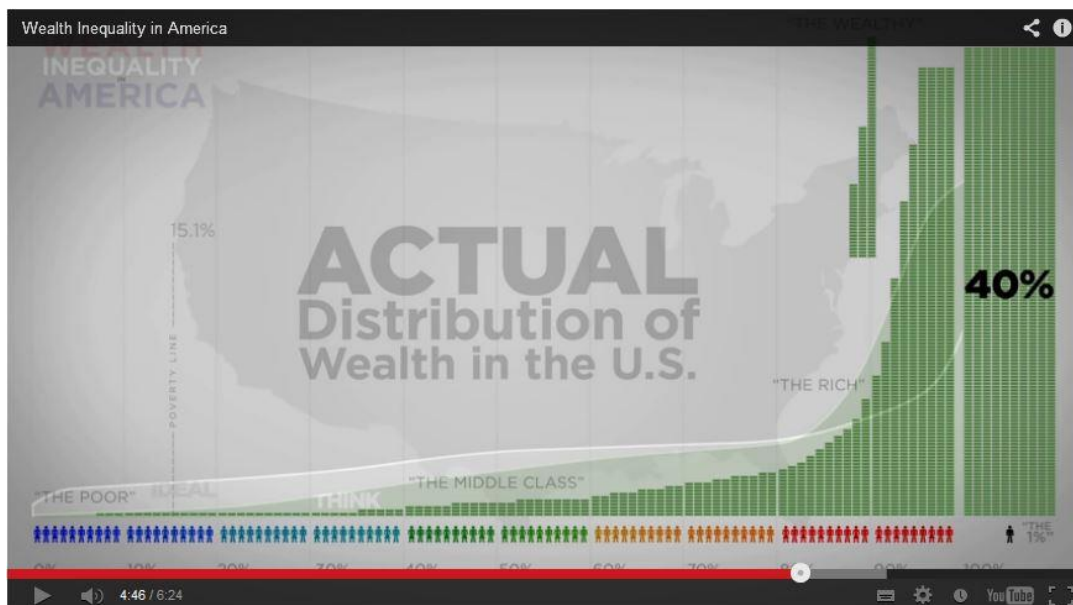
Datajournalismin kehittämishankkeen läpivieminen työyhteisössäni on ollut mielenkiintoinen projekti. Oma oppimisprosessini datajournalismin parissa jatkuu. Asetin itselleni tämän opinnäytetyön tavoitteeksi kehittää erikoisosaamista datajournalismissa sekä kehittyä digitaalisen journalismin tekemisessä. Nämä tavoitteet olen saavuttanut, tosin paljon on vielä opittavaa.

Tutkimuksen alkuvaiheessa kävi ilmi, että toimituskunta tarvitsee monipuolisesti oppia datajournalismin tekemiseen. Tähän tarpeeseen opinnäytetyö osaltaan vastaa. Yrittäjäsanomien toimituksen kehittämishanke kehitti valmiuksiamme tehdä datajournalismia, siksi uskonkin, että tästä työstä on apua kollegoilleni. Tutkimukseni erityinen ansio on, että se tuo datajournalismin lähestyttäväksi myös pienille toimituksille.

Datajournalismilla on annettavaa tänä päivänä, kun uutismediat etsivät uusia tapoja esittää juttuja verkossa ja hyödyntää digitaalisen alustan tuomat mahdollisuudet. Mediatiloille datajournalismi kannattaa myös siksi, että datavisualisoinnit keräävät huomiota pitkän aikaa, ns. pitkä häntä. Toimittajan kannattaa pitää huolta ammattitaidostaan, ja datajournalismiin perehtyneelle toimittajalle on varmasti kysyntää työmarkkinoilla. Tutkimuksessa kävi ilmi, että päätoimittajien kannattaa motivoida toimituskuntaansa korostamalla sitä, että datajournalismissakin juttu on se pääasia. Taidot ja työkalut ovat opetettavissa ja uutisvainu on kaiken a ja o (Rogers 2013). Kehittämishanke osoitti konkreettisesti, että suunnitelmallisuus datajournalistisessa juttuprosessissa alusta alkaen on tärkeää juttuprosessin nopeuttamiseksi. On selvää, että toimituksen juttupalavereissa pitää totutun lähestymistavan lisäksi kysyä myös, millaista dataa kulloisenkin juttuidean taustalla on. Jos data ei ole mukana juttupalavereissa, on vaarana, että datajournalismin tekemisestä tulee luonteeltaan satunnaista.

Joukkoistaminen on lisääntynyt suomalaisessa mediassa. Näkisin, että Britanniassa käytössä oleva joukkoistamisalusta ”Help me investigate” voisi toimia myös Suomessa. Sivustolla voi aloittaa kiinnostavan aiheen ”penkomisen”, ja kiinnostuneet voivat kantaa kortensa kekoon. Sivuston tarkoituksena on tutkia aiheita, jotka ovat julkisen kiinnostuksen kohteena. Isot mediatalot, kuten Helsingin Sanomat, toteuttaa joukkoistamishankkeensa onnistuneesti hs.fissä, mutta pienemmän mittaluokan mediatalot hyötyisivät joukkoistamisalustasta.

Datajournalismi voi tulevina vuosina saada uudenlaisia muotoja. Yksi vähemmän käytetty esitystapa on videografiikka, selostettu ja animoitu grafiikka videon muodossa. Oheisessa esimerkissä on ruutukaappauskuva paljon katsojia keränneestä videografiikkajutusta.



Kuva: Videografiikka varallisuuden epätasa-arvoisuudesta Yhdysvalloissa - ruutukaappaus

Videografiikka on kuitenkin vain yksi datan esitysmuoto, muut datavisualisoinnit pysyvät varmasti mukana ja niiden käyttö lisääntyy avoimen datan lisääntyessä. Jonesin ja Salterin mukaan ”Journalismin tulevaisuus piilee siinä, että yhdistetään kaikki digitaalisuuden ja verkon tuomat mahdollisuudet, ja parannetaan sekä uudistetaan journalismia tulevia sukupolvia varten” (2012: 173).

Suomessa yksi mielenkiintoinen journalismin tulevaisuutta ideoiva foorumi on Helsingin Sanomain Säätiön Uutisraivaaja, joka on innovaatiokilpailu. Uutisraivaaja-kilpailulta voimme odottaa edelleen uusia journalistisia konsepteja, joista enemmistö kohdistunee digitaaliseen journalismiin.

Lähteet

Antikainen, Hannele 2012. Hyperlocal Content. Next media tuloskooste nro 2, 8.

Luettavissa osoitteessa:

http://virtual.vtt.fi/virtual/nextmedia/uutiskoosteet/Next_Media_tk_2_2012.pdf (Luettu 30. syyskuuta 2012)

Bradshaw, Paul ja Rohumaa, Liisa 2011. The online journalism handbook: skills to survive and thrive in the digital age. Harlow : Longman.

Cairo, Alberto 2013. The Functional Art. Berkeley: New riders.

Carroll, John M. 1999. Five Reasons for Scenario-Based Design. Proceedings of the 32nd Hawaii International Conference on System Sciences.

Charles, Arthur 2010. Analysing data is the future for journalists, says Tim Berners-Lee. Luettavissa osoitteessa: <http://www.guardian.co.uk/media/2010/nov/22/data-analysis-tim-berners-lee> (luettu 5. huhtikuuta 2013)

Finnäs, Jens (2011-13) Tutorial. Luettavissa osoitteessa: <http://dataist.wordpress.com> (luettu 7. maaliskuuta 2013)

Haara Paula 2012. Viestintäalan ja - ammattien tulevaisuuden osaamistarpeita.

Luettavissa osoitteessa:

http://www.oph.fi/download/141442_Selvitys_viestintaalan_laadullisesta_ennakoinnista.pdf (luettu 28. maaliskuuta 2013)

Harju, Auli 2012. HS Open hackathon –luento. Open Knowledge Festival Helsinki. 18. syyskuuta 2012.

Help Me Investigate – verkkosivu. About. Luettavissa osoitteessa:

<http://helpmeininvestigate.com/> (Luettu 28. huhtikuuta 2013)

Helsingin Sanomat. HS perustaa datajournalismin ryhmän. 2012a. Luettavissa osoitteessa:

<http://www.hs.fi/kotimaa/HS+perustaa+datajournalismin+ryhm%C3%A4n/a1305567392362> (luettu 2. huhtikuuta 2013)

Hurrell, Bella ja Leimdorfer, Andrew 2012. Data Journalism at the BBC. The data Journalism Handbook. Sebastopol: O'Reilly.

Itä-Suomen Yliopisto. Miten teen verkkokyselyn tai -lomakkeen Googlessa? 2012.

Luettavissa osoitteessa: <https://wiki.uef.fi/pages/viewpage.action?pageId=15466918> (luettu 25. huhtikuuta 2013)

Jones, Janet ja Salter, Lee 2012. Digital journalism. Lontoo: SAGE.

Kanninen, Jaana 2013. Tuoko ryntäys kaivoksen naapuriisi – verkkokartta. Luettavissa osoitteessa:

http://yle.fi/uutiset/tuoko_ryntays_kaivoksen_naapuriisi_katso_kartalta/6476988 (luettu 18. huhtikuuta 2013)

Kuula, Arja 1999. Toimintatutkimus. Kenttätyötä ja muutospyrkimyksiä. Tampere: Vastapaino.

Marshall, Sarah 2013. 10 things every journalist should know in 2013. Luettavissa osoitteessa: <http://www.journalism.co.uk/news/10-things-every-journalist-should-know-in-2013/s2/a551648/> (luettu 28. maaliskuuta 2013)

Mäkinen Esa 2012b. HS selvitti kiinteistönvälittäjien hinnat. Luettavissa osoitteessa: <http://www.hs.fi/talous/HS+selvitti+kiinteist%C3%B6nv%C3%A4litt%C3%A4jien+hinnat/a1305606469642> (luettu 19. huhtikuuta 2013)

Poikola, Antti 2013. Johdattelua datajournalismiin. Luettavissa osoitteessa: <http://www.slideshare.net/fullscreen/apoikola/20120123-johdattelua-datajournalismiin/1> (luettu 13. huhtikuuta 2013)

Poutanen, Petro 2013 Nettikyselyjen sudenkuopat - blogikirjoitus. Luettavissa osoitteessa: <http://organisaatioviestinta.com/tag/datajournalismi> (luettu 21. huhtikuuta 2013)

Rigatelli, Sara 2013. Tässä ovat Suomen parhaat terveyskeskukset. Luettavissa osoitteessa: http://yle.fi/uutiset/tassa_ovat_suomen_parhaat_terveyskeskukset/6357493 (luettu 21. huhtikuuta 2013)

Rogers, Simon 2012. Data Journalism in the Newsroom -luento. Open Knowledge Festival Helsinki. 18. syyskuuta 2012.

Tebest, Teemo 2012. Näkökulma: Datajournalismi ennen, nyt ja tulevaisuudessa. Luettavissa osoitteessa: <http://datajournalismi.blogspot.fi/2012/06/nakokulma-datajournalismi-ennen-nyt-ja.html> (luettu 6. huhtikuuta 2013)

Tebest, Teemo 2013. Datajournalismi – luento. Luettavissa osoitteessa: http://teelmo.info/teelmo/app/seminar-2013-datajournalismi_otavan_opisto/datajournalismi.html#/datajournalism (luettu 19. maaliskuuta 2013)

Tufte Edward 2011. The Visual Display of Quantitative Information. 2nd edition. Cheshire, CT: Graphics Press.

Valkama, Meri 2012. Tulevaisuuden tekijät. Journalisti. 6. syyskuuta 2012.

Valkeeniemi, Hannele ym. 2013. Me tiedämme missä asut. Luettavissa osoitteessa: http://yle.fi/uutiset/me_tiedamme_missa_asut/6590408

Vehkoo, Johanna 2011. Painokoneet seis! kertomuksia uuden journalismin ajasta. Helsinki: Kustannusosakeyhtiö Teos.

White, Charlie 2013. Wealth inequality in America. Katsottavissa osoitteessa: <http://mashable.com/2013/03/02/wealth-inequality/> (katsottu 28. huhtikuuta 2013)

Vis, Farida 2012. What do data visualizations want? Tieto näkyväksi – seminaari. 17. syyskuuta 2012.

yrittajat.fi. Yrittäjävaltuutettujen osuus valtuustosta. 2012. Luettavissa osoitteessa: <http://www.yrittajat.fi/fi-FI/kunnallisvaalit2012/> (luettu 18. huhtikuuta 2013)

Haastattelut

Finnäs, Jens 2013. Datatoimittaja. Free-toimittaja/Ruotsi. Otavan opiston kurssin järjestämä Adobe Connect-haastattelu: 7. maaliskuuta 2013.

Koivikko, Kimmo ja Hallamaa Hannu 2013. Päätoimittaja; toimittaja. Suomen Yrittäjien Sypoint Oy. Kehittämishankkeen arviointikeskustelu: 22. huhtikuuta 2013.

Mäkinen, Esa 2013. Datatuottaja. Helsingin Sanomat. Haastattelu: 15. maaliskuuta 2013.

Poikola, Antti 2013. Konsultti ja tutkija. Haastattelu: 26. helmikuuta 2013.

Pärnänen, Janne ym. 2012 Verkkotuottaja. Suomen Yrittäjien Sypoint Oy. Kehittämishankkeen aloituspalaveri: 6. marraskuuta 2012.

Rogers, Simon 2013. Datatuottaja. The Guardian ja The Guardian Datablog. Skype-haastattelu: 10. tammikuuta 2013.

Liite 1: Datajournalismin työkalun kehittämisiin

Suomen Yrittäjien Yrittäjäsanomien ja yrittajat.fin henkilökunnan kehittämishankkeen alkupalaveri 6. marraskuuta 2012.

Avoim data uutishankinnan tukena – datajournalismin työkalun kehittämisiin

Opinnäytetyö: Datajournalismi
uutistoimituksen työkalupakkiin

Riika Koskenranta 6.11.2012

Tausta

- Mitä tarkoitetaan avoimella datalla?
 - julkinen data, johon kohdistuu vaatimuksia avoimuudesta
 - julkinen data tässä tarkoittaa laajoja, usein numeerisia tietoaaineistoja, tarjolla on nykyään sovelluksia, joilla tietoaaineistoja voi siivota yhdenmukaisiksi jne
 - dataa on myös paremmin saatavilla kuin ennen, esimerkiksi HRI
- Näin: julkisen datan yhdistäminen SY:n omaan tietoon -> esimerkiksi yrittäjävaltuutettujen suhteellinen osuus kunnanvaltuustoista
- Ilmaisia työkaluja (karttadataa Fusion Tables, zoomaus)

Riika Koskenranta 6.11.2012

Datajournalismin työnkulun skenaarion testaaminen

1. Miten valitsemme dataa, jota haluamme tutkia ja esittää visuaalisin keinoin?
2. Paljonko prosessi vie keskimäärin aikaa?
3. Millaisissa tapauksissa voisimme käyttää ulkopuolisen koodaajan apua?
4. Millaisia tavoitteita meillä on datajournalismille ja datan hyödyntämiselle toimitustyössä?
5. Datajournalismin visualisoinnit toimivat myös verkossa hyvin, kumpaan datajournalismista haetaan lisäarvoa ensisijassa (yrittajat.fi/YS)?

Liite 2: Ohjeet datan hankintaan

Neljä olennaisinta tapaa hankkia dataa ovat: avoimen data, julkisuuslain mukainen tietopyyntö, ruudun raavinta (screen scraping) ja joukkoistaminen.

1. Avoin data (Suomea koskevaa dataa)

- ❖ Tilastokeskuksen StatFin-tilastotietokanta - Suomea koskevaa tilastotietoa
- ❖ Helsinki Region Infoshare datakatalogi pääkaupunkiseutuun liittyvistä aineistoista
- ❖ Findikaattori
- ❖ Maanmittauslaitos (Avoimien aineistojen tiedostopalvelu)
- ❖ Paikkatietoikkuna.fi
- ❖ Terveyden ja hyvinvoinnin laitoksen tilastopankki SotkaNet - SOTKANet - tilastoja suomalaisten hyvinvoinnista ja terveydestä sekä väestö- ja terveystilastoja EU-tasolla
- ❖ Datavaalit hankkeen aineistosivu (mm. kunnallisvaalien ehdokasdataa)
- ❖ Helsingin ympäristökeskuksen tilastoja
- ❖ Kelasto - Kelan tilastotietokanta
- ❖ Ulkomaankauppatilastot – Tulli
- ❖ KTI Kiinteistötieto (maksullisia aineistoja)
- ❖ Matilda - maataloustilastot (maksullisia aineistoja)
- ❖ Vero.fi - Yhteisöjen tuloverotuksen julkiset tiedot

b. tietopyyntö

- ❖ Jyväskylän yliopiston lista viranomaisrekistereistä - <https://koppa.jyu.fi/avoimet/viesti/viranomaisrekisterit>
- ❖ Facebook-ryhmä Tietopyyntö.fi

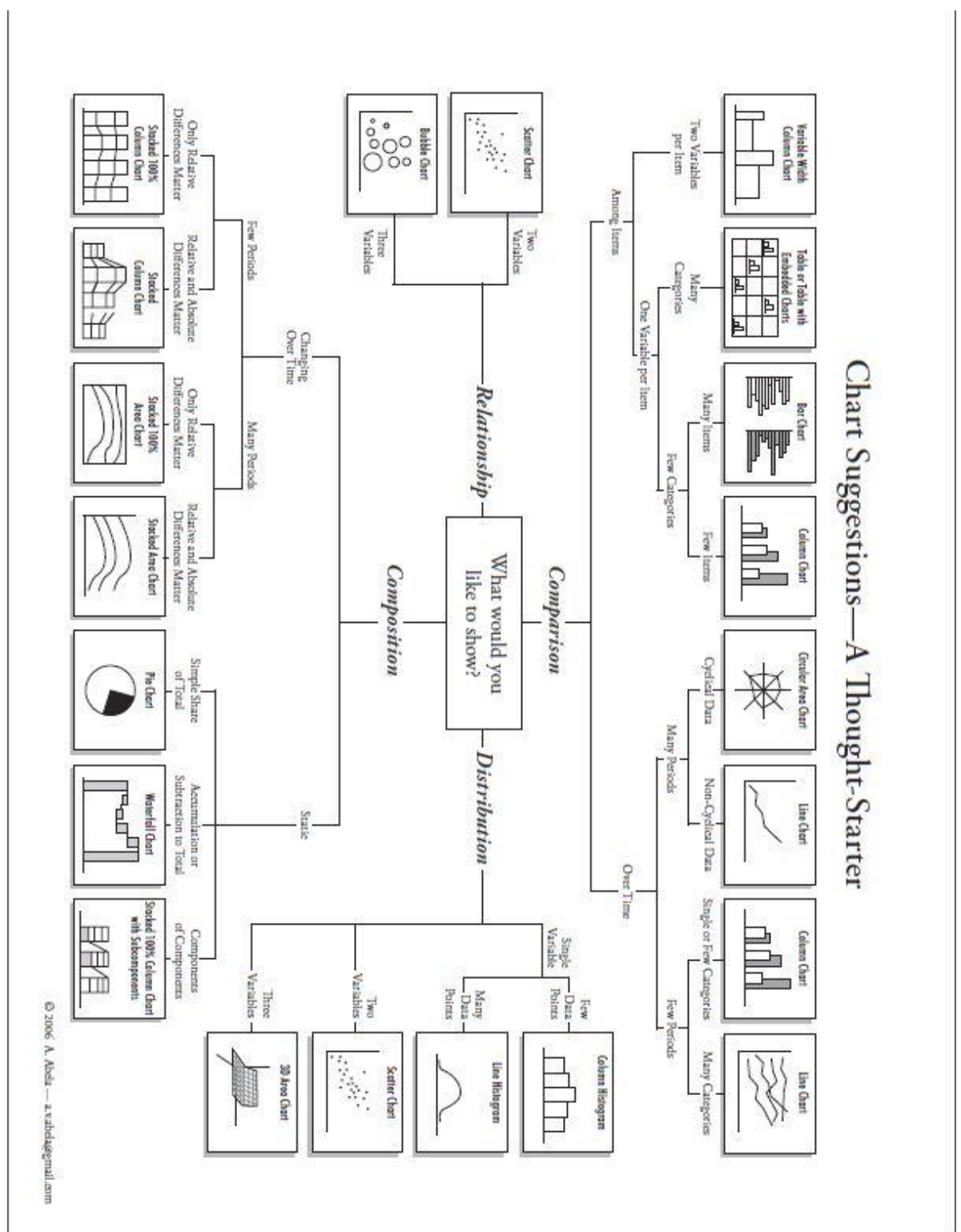
c. ruudun raavinta (screen scraping)

- ❖ Google Chrome -selaimen scraper lisäosa (katso tutoriaali, esim. listoja voi "raapia" verkkosivulta ja viedä taulukkolaskentaohjelmaan)
- ❖ CometDocs (maksuton) – muunna mm. pdf, excel ja word - tiedostoja toiseen muotoon
- ❖ Sublime text (maksullinen, tosin kokeilu ilmaista) - datan siivoaminen
- ❖ jne

d. joukkoistaminen

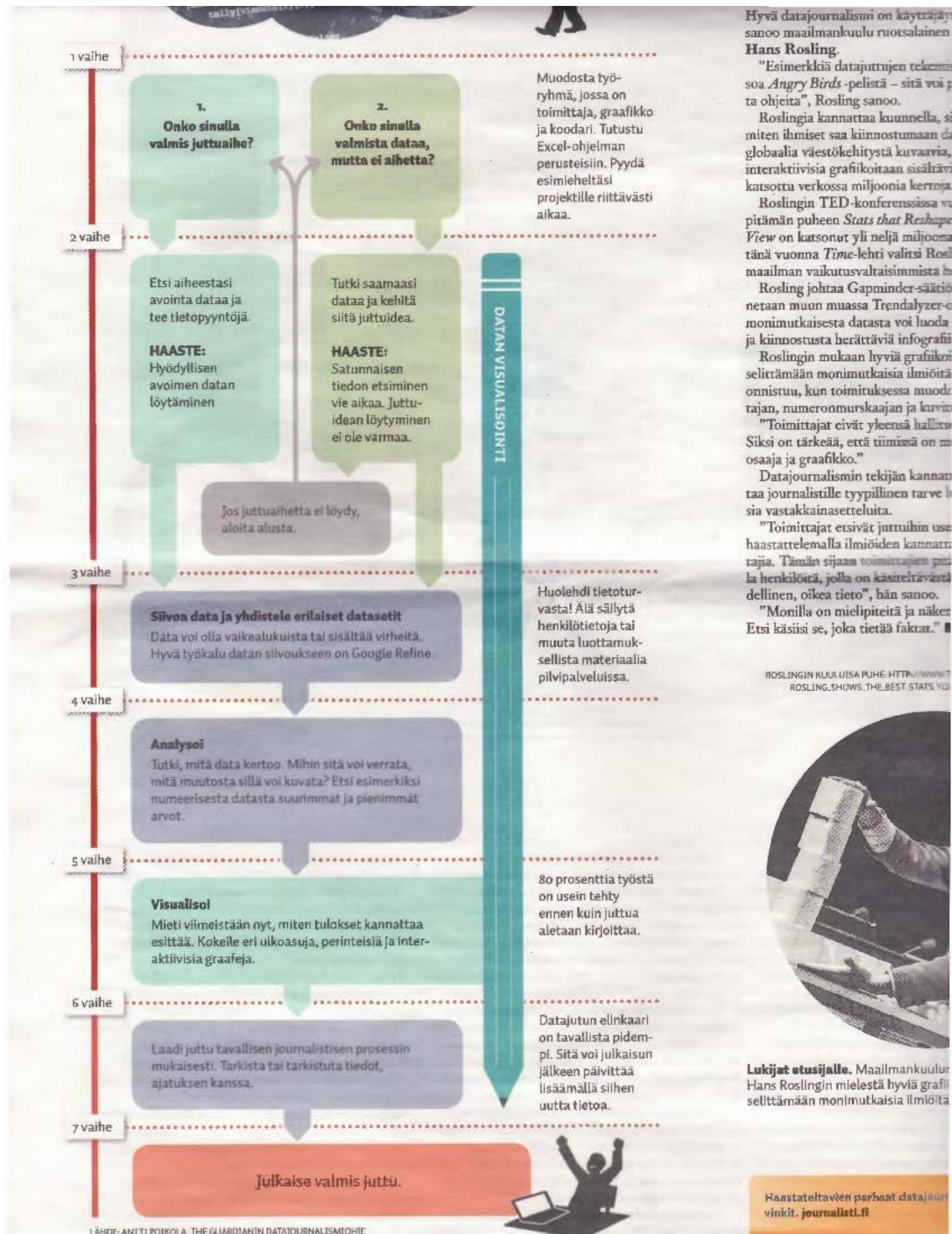
- ❖ joukkoistamisen tärkeimmät seikat: 1) saadaan ainutkertaista dataa, mutta samalla siinä hetkessä kerättyä (mielipiteet voivat vaihtua) ja 2) uniikkia (lähde: Teemo Tebest 2013)
 - > Google scriptin avulla osoitteet saa koodattua paikkatiedoksi
- ❖ Itä-Suomen Yliopiston ohjeet verkkokyselyn tai -lomakkeen tekemiseen Google Formsilla (2013)

Liite 3: Excel-kuvaajan valitseminen



Lähde: 2006 A. Abela – a.v.abela@gmail.com

Liite 4: Datajournalismin toimitusprosessin malli (lähde: Antti Poikola)



Hyvä datajournalismin on käyttänyt sanoo maailmankuulu ruotsalainen Hans Rosling.

"Esimerkkiä datajuttujen tekemiseen soa *Angry Birds* -pelistä – sitä voi p... ta ohjeita", Rosling sanoo.

Roslingia kannattaa kuunnella, si... miten ihmiset saa kiinnostumaan da... globaalia väestökehitystä kuvaavia, interaktiivisia graafikoitaan sisä... katsottu verkossa miljoonia kertoja.

Roslingin TED-konferenssissa va... pitämän puheen *Stats that Reshape View* on katsonut yli neljä miljoonaa tänä vuonna *Time*-lehti valitsi Rosl...

maailman vaikutusvaltaisimmista b... Rosling johtaa Gapminder-säätiö... netaan muun muassa Trendalyzer-o... monimutkaisesta datasta voi luoda ja kiinnostusta herättäviä infografi...

Roslingin mukaan hyviä grafiikoita selittämään monimutkaisia ilmiöitä onnistuu, kun toimituksessa muode... tajan, numeromurskaajan ja kuvien...

"Toimittajat eivät yleensä hallitse... Siksi on tärkeää, että tiimissä on m... osaja ja graafikko."

Datajournalismin tekijän kannan... taa journalistille tyypillinen tarve h... sia vastakkainasetteluita.

"Toimittajat esiväät juttuihin use... haastattelemaalla ilmiöiden kannatt... taja. Tämän sijaan toimittajan puo... la henkilöitä, jolla on käsiteltävistä...

dellinen, oikea tieto", hän sanoo.

"Monilla on mielipiteitä ja näköt... Etsi käsiisi se, joka tietää faktat."

ROSINGIN KUULUISA PUHE: [HTTP://WWW.TED.COM/TALKS/HANS_ROSLING_SHOWS_THE_BEST_STATS_YOU...](http://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ever_ignored)



Lukijat otusijalle. Maailmankuulun Hans Roslingin mielestä hyviä grafiikka selittämään monimutkaisia ilmiöitä

Haastateltavien parhaat datajournalistit. journalisti.fi