

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Uskali, Turo; Kuutti, Heikki

**Title:** Models and Streams of Data Journalism

**Year:** 2015

**Version:**

**Please cite the original version:**

Uskali, T., & Kuutti, H. (2015). Models and Streams of Data Journalism. *The Journal of Media Innovations*, 1(2), 77-88.

<https://www.journals.uio.no/index.php/TJMI/article/download/882/1160>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

## Models and Streams of Data Journalism

Turo Uskali

Department of Communication, University of Jyväskylä  
turo.i.uskali@jyu.fi

Heikki Kuutti

Department of Communication, University of Jyväskylä  
heikki.kuutti@jyu.fi

### ABSTRACT

This paper presents the initial results of a two-year research project, Data Journalism Work Practices, which focuses on newsrooms in Finland, the UK and the US. Data journalism or data-driven journalism has been defined simply as journalism based on large data sets (Rogers 2011; Bounegru et al. 2012.) According to our ongoing research on data journalism work methods, we can claim this has been an oversimplification. Based on six interviews of leading Finnish, American and British data journalists we can claim that there are already at least three different models for organizing data journalism work practices, and two main streams of data journalism, not just one.

### KEYWORDS:

Data journalism, work practices, newsrooms, investigative reporting, interviews

The Journal of Media Innovations 2.1 (2015), 77–88

<http://www.journals.uio.no/index.php/TJMI>

© Turo Uskali & Heikki Kuutti 2015.

## INTRODUCTION

Data, data everywhere. Now that we're deep into the information age, it's time for everyone to accept that the amount of information in our lives is only going to keep growing. (Mark Briggs 2013, p. 233)

Without question large digital data sets will challenge practices in many industries, also in news media.

One of the prime tasks of journalism is to monitor and evaluate the actions of government and other power holders. This can be achieved mainly through the access of official files produced or received by authorities. Traditionally official documents have been in the paper format. But as an element of information management technologies, digital records comprising database systems have become common tools for journalists.

The transformation to digitalization started more than 15 years ago, but only recently have large digital data sets been available. WikiLeaks' secret data sets have often been mentioned as the starting point for contemporary data journalism activities. The Afghan war logs, Iraq war logs and the US embassy cables have since 2010 forced many prominent news organizations, all over the world, to start

their initial data journalism activities (Rogers 2011; Bounegru et al. 2012; Hewett 2013.)

We argue that 'data journalism' and 'data-driven journalism' have been accepted rapidly and easily as new concepts among journalists. As Jonathan Hewett (2013, 3) from City University, London, remembers "using the term 'data journalist' only five years ago would have produced puzzled faces'. Data journalism or data-driven journalism has often been defined as journalism based on large data sets, otherwise known as 'big data' (Rogers 2011; Bounegru et al. 2012). In a similar vein, a data journalist could be defined as a person who creates news stories based on large data sets. In addition, Parasie and Dagiral (2012), when summarizing the evolution of data-driven journalism in Chicago, have used the terms "computer-assisted-reporters" and "programmer-journalists". Even so, here we use data journalists.

This new working method consists of analyzing large databases by spreadsheets or database software programs. Information data may be searched, categorized, arranged, counted, compared and cross-indexed rapidly and precisely. Information created this way gives journalists "new horizons" in their work. The data journalism approach enables journalists to build new kinds of entities and to see

the broader patterns of the covered topics. Using data journalism practices, journalists are able to uncover the connections and reasons behind seemingly separate issues, which could not be revealed using only individual data.

Traditionally, journalists have looked for separate and detailed information in documents and write their stories based on them. A data-based approach means that journalists can receive the entire register or database from an authority and do their own computer runs of the data in those registers.

News companies around the world have started their data journalism activities gradually. Parasie and Dagiral (2012, p. 9) detected the first data-driven initiatives under the banner of open government advocacy beginning in 2005. Powers (2012, p. 27) has noticed that there are often tensions among occupational subgroups in the newsrooms, when new forms of journalism are tested and developed. His early example is the introduction of photojournalism in the early 20<sup>th</sup> century. Later, at the turn of the new millennium, Powers reminds us that the online journalists were first seen as "second-class workers" in the newsrooms. Therefore, integrating online and print staff required a lot of effort.

In a similar way, data journalism is once again challenging newsrooms. Data desks try to assimilate

late journalists, graphic designers and programmers into joint projects.

After all of this, we can claim that the first big wave of data journalism started in 2010 and included for example, Britain's *The Guardian* and the US newspapers like the *New York Times* and *Chicago Tribune*. The UK and the US news media organizations have been at the forefront in adapting and adopting new digital technologies for journalistic processes many times before. (OECD 2010; Kleis Nielsen 2012; Parasie & Dagiral 2012; Briggs 2013.)

In Finland, the first data journalism news desk was founded in the summer of 2012 by *Helsingin Sanomat*, and then by the Finnish Broadcasting Company, *YLE*, in 2013. Prior to these data desks appearing, individual journalists in Finland tested data journalism practices (Mäkinen 2013; Finnäs 2013).

So far, most of the literature about data journalism has been aimed at working journalists and is authored by data journalists focusing mainly on their best practices (Holovaty 2006; Rogers 2011, Bounegru et al 2012; Mair 2013.)

Data journalism literature has focused on big themes like the conflicts in Afghanistan and Iraq, offshore tax-havens, the Olympics, horse-meat trafficking, immigration, tobacco smuggling, crime

in big cities, pollution in certain areas, and traffic jams (Bounegru et al 2012).

The very first Data Journalism awards by the Global Editors Network in 2012 have also emphasized the importance of the investigative reporting tradition in data journalism. It is also worth mentioning that the average amount of time used to compose a story was several months, the longest period was seven years (Burn-Murdoch 2012).

As Nygren et al. (2012, p. 82) notice, there is scant academic research on data journalism. Some of the early pioneers were researchers from the University of Tampere, Finland, where Sirkkunen et al. (2011) conducted seminal research on data journalism in several countries. Among their conclusions was a prediction that "as the cost of data analysis continues to decrease, the skill requirements of data analysis keep growing, and open data is made more available by governments, companies and organisations, more and more data journalism will occur" (p. 16).

Interestingly, even in 2014 many researchers appear to view data journalism as a monolith rather than several different work streams (Appelgren & Nygren 2014). This paper will challenge this by further exploring and updating how information and communication technology and big data has changed the core practices of data journalism, fo-

cusing on how some prominent newsrooms in the US, the UK, and Finland have adopted data journalism.

## INVESTIGATIVE REPORTING TRADITION

We open this section by presenting the necessary context for data journalism with a brief history of traditional “shoe-leather” investigative reporting in Western democracies. Data journalism has its roots in Computer Assisted Reporting (CAR) and in the tradition of investigative journalism (Meyer 1973; Ettema & Glasser 1998).

The Europeans have openly admitted that the US newsrooms were greatly ahead of them in using CAR methods (Hewett 2013, p. 4). Beginning in the 1960s, Phillip Meyer introduced social science research methods into American journalism via Detroit Free Press’s prize-winning investigation into the causes of riots in Detroit. Meyer directed quantitative survey research with a team of 30 interviewers, two university professors and – a computer programmer. Even before this, in 1952, the US television network CBS used a powerful computer to predict election results (Meyer 1973; Hewett 2013, p. 6).

Along with the fast adaption of new information and communication technology in newsrooms, the US has benefited from the Freedom of Information Act (FOIA) since 1967. In brief, the Act declares that every federal agency shall post for public view

information regarding its operation, procedures, opinions and orders, staff manuals, and indices to these materials (Schudson 2014, p. 1). In comparison, the freedom of information legislation did not come into force in the UK until 2005 (Hewett 2013, p. 6).

In addition, it is important to note that Sweden was the first country in the world to create freedom of information legislation as early as 1766. According to the Finnish Act on the Openness of Government Activities, documents are to be in the public domain regardless of the format or unless there is a specific reason for withholding them. Public information in data format may be an entry to a register or a register as a whole (as a collection of entries).

Without Freedom of Information legislation, there are not any open access data sets, and, essentially no proper data journalism. As Schudson (2014, pp. 10–11) asserts, the drive to make government more open has lasted over the past half century, but only recently produced results. He mentions several examples from the *New York Times*, where journalists’ use of FOIA has made front-page news.

Filing a FOIA request is indeed one of the most important tools for a data journalist, even if some critics have feared that journalists can be tied too much to computers in assessing published data.

Some have also mentioned that one possible unintended consequence of the FOIA may be a ‘chilling effect’, meaning that decisions could go unrecorded or sanitized due to the fear of coming requests (Riddell 2014, p. 24; Worthy & Hazell 2014, p. 37).

## INTERVIEWING LEADING DATA JOURNALISTS

The primary source of data in this paper is the first round of interviews of data journalists for the research project *Data Journalism Work Practices*. There were six interviewees, two Finnish, two Americans and two British. All the interviewees could be defined as leading data journalists in their area. The journalists worked for *The Financial Times* (FT), *Los Angeles Times* (LAT) and *Pro-Publica* (PP) and *Helsingin Sanomat* (HS). Two of the interviewees were freelancers publishing their stories in several outlets in Finland/Sweden and in the UK.

Interestingly, all the interviewees, in addition to working as data journalists, were also often educating other journalists in their own or other newsrooms. One data journalist, Paul Bradshaw from the UK, was even lecturing at two universities. He had also published several handbooks and articles on data journalism, for example *Scraping for Journalists* (2013a) and *Data Journalism Heist* (2013b).

The interviews were conducted during May 2013 and January 2014 by the authors of this paper. Two of the interviews were conducted at the interviewee's work places, one in Helsinki and another in Birmingham. The rest of the interviews

were conducted during a data journalism seminar and conference breaks in Stockholm (*The First Data Journalism Conference in Nordic Countries*, Södertörn Högskola, 22-23.11.2013) and (*Data Journalism: Mapping the Future*, Adam Street Private Members Club, 22.1.2014 London). All the interviews were recorded and later transcribed in order to better analyze the content.

Since none of the interviewees requested anonymity, all their names are used in this paper.

The main starting questions for the interviews were: 1) Is data journalism only for specialists or for every journalist? 2) How do you describe the evolution of data journalism in your newsroom and country, and 3) How do you see the near future of data journalism?

The interviews lasted from 10 minutes to over one hour. Longer interviews, of course, included more themes than the shorter ones, like several examples of data journalism work processes in more detail and also important issues about data journalism ethics. Because of space limitations, only a few parts of the interviews are presented in this paper, focusing on the different models and streams of data journalism.

In addition to the interviews, the authors recorded some panel discussions and keynote talks by the interviewees and other data journalists,

which were used as source material for this paper. Furthermore, recent books and articles about data journalism were used as background material.

## IS DATA JOURNALISM FOR EVERY JOURNALIST?

The interviewees did not totally agree with one another about whether or not data journalism is for only specialists or for everyone. Sisi Wei from *Pro-Publica* was the most eager supporter of the claim that data journalism is for everyone.

I think no matter what kind of journalism you are doing, having that skill to be able to fact check your sources. It is just another empowering tool. So, I think it is probably for everyone, knowing using Excel or use Python, it does not matter how little of how little programming you know as long as you know how to use it. To do this kind of thing. It is definitely everybody. (Sisi Wei)

In a similar fashion Paul Bradshaw believed that data journalism is for almost every journalist.

If you are a journalist, you have to deal all types of information then you have to work with that. Speed and accuracy are the two key assets for a journalist. And if you deal with data, you have to do with that. There will be some sort of journalism that does not deal with the data. It will be increasingly small area. (Paul Bradshaw)

Ben Welsh from the Los Angeles Times wanted to choose both sides of the argument.

Well, I think it is a little bit of both. There are more difficult, require more technical skills that you need people to specialize to do. - - But there is a lot of data work that is really looking stuff up and like doing basic math, and I think that it is things that any smart journalist is really capable of. (Ben Welsh)

One of the initial ideas for the data desk in *Helsingin Sanomat* was that its data journalists would also educate other journalist in doing data journalism.

One of my prime aims is to minimize my own work and help others to do more data work. (Esa Mäkinen)

There are already several examples of successful data journalism projects in *Helsingin Sanomat* led by “common journalists” and only assisted by the newspaper’s data desk. One of those was the ‘lobbyists in the Finnish parliament story’ by Tuomas Peltomäki, which won the very first Nordic award on data journalism in the fact category (Djurberg 2013).

Finnish-Swedish freelancer Jens Finnäs was in the opposite camp, arguing that not every journal-

ist could be a data journalist. He believes a real data journalist is a person who can code.

Not every journalist should have coding skills, but all journalists could at least know one person, who can code, and that he or she can ask for help. (Jens Finnäs)

During a panel discussion on data journalism in London, Martin Stabe from the FT wanted initially to criticize data journalism as being already a new buzzword.

I think there is the tendency, especially because it has become a buzzword in the industry to kind of mystify or fetishize data. What is data? Data is really just documents. (Martin Stabe)

Later on, however, he argued very strongly that journalists just had to keep up with other professions, and industries, which were more capable of doing independent quantitative analysis based on raw data.

If we cannot do our own independent analysis, we are at the mercy of the PR machine, which has this firepower behind them. There is this absolutely es-

sential need to be able to do independent quantitative analysis based on raw data. -- So I think that is really the key in journalism. It is not only the flashing new things that you can do fancy graphics. It is not only numbers. It is about responding to the way information is now stored, controlled, processed and allowing journalism to fulfill its mission, independently accessing those things, and without data journalism that is going to be impossible. (Martin Stabe)

In many respects, these citations already demonstrate that even a small number of data journalists see the mission of data journalism slightly differently. This can be the result of different definitions of data journalism, or/and different newsroom realities. Finnäs referred to coding skills, whereas Wei was happy with journalists using data as a fact checking tool.

These differences already force us to rethink the concept of data journalism. At this point, however, we want to continue to present more insights from the interviewees.

### “AWAY FROM DATA DESKS”

In our interview, FT’s Martin Stabe surprised us by saying that actually the paper does not have a data desk or special data journalists, even if it does produce journalism based on large data sets.

There is no central unit, it’s kind of diffused to our newsroom. And those people kind of come together, and work on projects a lot, but very much informally. You could not look at FT’s organizational chart and find [any] data journalists, but now you can find – one. (Martin Stabe)

The FT had just recruited one journalist with the title of ‘data journalist’. Actually, FT wants to use the phrase “interactive news team” instead of “data desk”. According to Stabe FT has more than 10 people working on data projects.

Also Ben Welsh, LAT, describes his newsroom as a flexible data hub, where new data projects develop without any central command post.

We have a network of people through our newsroom. We are doing computer programming, data analysis, data visualizations in an effort to tell stories about Los Angeles, California and the World. (Ben Welsh)

By contrast, *Helsingin Sanomat* not only established a data desk, but also employs three people; a journalist coder, a graphics designer/coder, and a graphic designer (Mäkinen 2013). In a similar vein, *Trinity Mirror* in the UK started its data journalism activities in 2013 with a data desk (Ottewell 2014).

An argument can be made that in “older” data journalism newsrooms, over time, the data desk merges into the common newsroom practices. The more other journalists are able to do data journalism activities, the less the data desk is needed.

Paul Bradshaw has seen the evolution of data journalism in the UK, and says that indeed “the move is away from data desks”.

It is almost as if, organizations innovating in particular direction and when they get the peak of that, they are not able to track back explore things of doing things better. So, different organizations lead the way. Everyone else have adopted the best practices. (Paul Bradshaw)

Bradshaw had special concerns about the future of *The Guardian*’s data journalism, because the newspaper had just lost two of its leading data journalists to the US. One of those, Simon Rogers, now works in *Twitter*. According to Bradshaw, invest-



ments in data journalism at both the BBC and the *Daily Telegraph* have already peaked. By contrast, regional newspaper chains like *Trinity Mirror* and many magazines have only just started investing in data journalism in the UK.

Bradshaw also recounted that some of his students are interested in creating online data services, for example around traffic data. Esa Mäkinen emphasized that *HS* was not in the business of creating data services, but news stories. Bradshaw had the same observation in the UK news outlets. This, of course, creates new opportunities for new data service startups.

On-the-fly learning is an important aspect of being a data journalist. Freelancer Jens Finnäs emphasizes that working outside any newsrooms gives more freedom and flexibility to constantly learn new data skills. He aims to be the leading Nordic data journalist, and uses a lot of his time for learning, in particular new coding languages. Indeed, it is easy to understand that working in a newsroom is more hectic than working as a freelancer. Of course, one has to be good enough to attract sufficient clients to earning a living as a freelancer.

This section could be summarized as follows. The interview data suggests there are three different models for organizing data journalism activities in newsrooms. The first is *the traditional data desk*

*model*, the second and more advanced is *the flexible data projects model*, and the third is *the entrepreneur model*. The entrepreneur model could also be called *the sub-contractor model*, because the data journalist often works for several news organizations and with many different projects.

Naturally, newsroom practices of data journalism can still differ a lot from one news organization to another, even if they belong to the same category of model. This will require even more extended examination next.

## TOWARDS REAL-TIME DATA JOURNALISM

Some other trends could also be identified. According to our interviews, data journalism is going to be faster. In addition to the traditional, investigative reporting style of data journalism with FOIA requests, there is an increasing tendency towards “real-time data journalism”.

Once again, the US newsrooms have been at the forefront of this innovative trend. Sisi Wei from *ProPublica* was educated at Northwestern University, which is the home for Narrative Science, a company that uses algorithms in order to automatically create news stories. Narrative Science started with sports events like baseball, but is now heading towards financial news.

We do some of that as well in ProPublica. We write articles about every school, school data base, using a similar algorithm, and filling the blank, here is the data.- - Computers made us faster. This is kind of development make journalists be able to focus more on things that really need human mind. Other things tossed away like automated. (Sisi Wei)

The *Los Angeles Times*’ fast data journalism is probably the closest to the ideal of real-time data journalism. Ben Welsh has created two automatic

data journalism feeds based on real-time official press releases and emails. One feed is from the Los Angeles Police Department, another from the Earthquake Monitoring Department.

They (police and earthquake monitoring departments) are doing the same system that we are doing. They have a little program, which writes the email addresses to their data base. And it almost comes with the same suddenness. If there is an earthquake of these magnitudes, located in these places, in these times, and then we have a code that catches that email, [and] parses it. It is trained to know the patterns, so it picks up bits of data that we want to analyze. And there is just small amount of computer code, like ten lines, which says to pull the number out of the email, if the number is greater than this, look the longitude and latitude of the location. It is also in that email. We have box codes all around California, and if both those things are yes, then start doing our own little math thing. The whole thing is probably two hundred lines of computer code, it is not that much, but of course you need to know how to write the code. (Ben Welsh)

Also *The Financial Times* is looking for real-time financial data journalism, but at present they lack resources.

I think sort of instant reactions to market trends will be something, certainly organizations like us have to pay attention to. That's what the banks can do, but we cannot do so, there is the question, whether we are going ever have resources for that that is very difficult. (Martin Stabe)

As for the Finnish case, real-time data journalism does not exist.

Next, we argue for two main streams for data journalism.

## TWO MAIN STREAMS OF DATA JOURNALISM

Based on our interviews, observations and analysis on data journalism, we define two main types of data journalism: *investigative data journalism* (IDJ) and *general data journalism* (GDJ), both of which have originated in investigative reporting, but have developed slightly different ethos and practices. We provide samples of 'paired' IDJ *versus* GDJ results to support our arguments:

- 1) In IDJ, the journalists have plenty of time for their work, which can take months and even years. By contrast, in GDJ the journalists have only hours or a few days to finish their data stories.
- 2) In IDJ, data skills are used at advanced levels, including coding. This could be done in team work. In GDJ, the data skills are at the basic level, for example, the capability of using Excel and some data visualization and analyzing tools.
- 3) In IDJ, the topic and the point of view of the story define the data to be gathered, whereas in GDJ, datasets are the starting point for the story and based on their content, define the

story's topic and the point of view.

- 4) In IDJ, unofficial and confidential information and data leaks often have great significance in starting and directing investigations, whereas in GDJ, massive public and open access data sets ("big data") decide the direction of the investigation.
- 5) In IDJ, journalists collate pieces of information from various sources in order to understand the broad picture (or pattern) of a news story, but in GDJ the data sets provide almost all the source material.
- 6) In IDJ, journalists need to confirm, cross-check and verify all the gathered information because they cannot trust only a single source. In contrast, GDJ data sets are often uncritically accepted as trustworthy sources.
- 7) In IDJ reporting, stories are based on gathered facts and journalists' own interpretations. In GDJ, stories are based almost solely on computer assisted analysis.
- 8) In IDJ, the newsroom has the power to in-

dependently prioritize which of the data sets to use for the stories, and request copies of them. In GDJ, priority is given to generally distributed and easily accessible, open data.

In addition to these two main streams of data journalism, we also recognize a third, still rather weak, stream called *Real-Time Data Journalism*, which is already in practice in several US newsrooms, but not in Europe. At Columbia University this is also known as sensor journalism (TowCenter.org 2014). Further research should focus more on this interesting new data journalism stream.

## CONCLUSIONS AND DISCUSSIONS

This paper, based on six interviews of leading data journalist from the US, the UK and Finland, argued that more than one type of data journalism exists.

According to the interviews, there are three different organizational structures for data journalism. The first is *the traditional data desk model*, the second, and more advanced is *the flexible data projects model*, and the third is *the entrepreneur model*. The entrepreneur model could also be called as *the sub-contractor model*.

We also detected two main types of data journalism streams and called them *Investigative data journalism* (IDJ) and *General data journalism* (GDJ). In IDJ the journalists have plenty of time for their work, have advanced data skills, as well as coding skills, and have strong links to the ethos of investigative reporting. For them, data journalism is really about what is "hidden in the shadows". By contrast, in GDJ the journalists have only hours or a maximum of a few days to finish their data stories. Their data skill set is somewhat modest and their professional attitudes towards data journalism less ambitious. In many newsrooms both data journalism streams operate side-by-side.

In addition, we also recognized a new, still rather weak, third stream called *Real-Time Data Jour-*

nalism. It is based on algorithms that automatically create news from data sources. This stream exists only in the US.

In conclusion, we argue that data journalism practices, especially the use of large data sets, could be used in many newsrooms. There is a tendency especially in the UK for data journalism practices to diffuse to the regional press and magazines.

Furthermore, it is possible that at least general data journalism (GDJ) will be the future of all journalism, as for example Sir Tim Berners-Lee, the inventor of the World Wide Web, has predicted on many occasions (Bounegru et al. 2012.) It is probable that once the methods of GDJ have been implemented as a common practice in newsrooms, IDJ will remain as an extravagance for a few prominent and best resourced newsrooms like *The New York Times*, *ProPublica* and *The Guardian*.

As Sirkkunen et al. (2011, p. 6) remind us, only ten years ago it was difficult to use large data sets in reporting and doing so required skills that common journalists did not have. Nowadays, free and easy-to-use data tools are widely available. The new problem seems to be finding sufficient time to explore all the new applications and their possibilities. In a way, data journalism has entered, at least in theory, a new phase from having a scarcity of data, and data tools, to an abundance of both. This

new phase creates pressures and challenges for data journalism education, both in the newsrooms and journalism schools.

Finally, a clear limitation of this paper is that it is based only six interviews. Therefore, further research should critically test the results of this paper.

#### ACKNOWLEDGEMENTS

We gratefully thank the Helsingin Sanomat Foundation for funding this research project.

#### REFERENCES

- Appelgren, E. & Nygren, G. (2014). Data Journalism in Sweden. Introducing new methods and genres of journalism into “old” organizations. Published online before print in *Digital Journalism*.
- Bounegru, L., Chambers, L., and Gray, J. (Eds.). (2012). *The Data Journalism Handbook*. EJC: Brussels. <http://datajournalismhandbook.org/1.0/en/>.
- Bradshaw, P. (2013a). *Scraping for Journalists*. Lean Publishing.
- Bradshaw, P. (2013b). *Data Journalism Heist*. Lean Publishing.
- Burn-Murdoch, J. (2012). Winners of the inaugural Data Journalism Awards announced. The Guardian Online, 31.5.2012, <http://www.theguardian.com/news/datablog/2012/may/31/data-journalism-awards-winners>, retrieved 20.1.2015.
- Djurberg, C. (2013). *Nordic Data Journalism Award till Sverige, Norge och Finland*. Sveriges Radio Online, 23.11.2013, <http://sverigesradio.se/sida/artikel.aspx?programid=3938&artikel=5713423>, retrieved 20.1.2015.
- Ettema, J. S. & Glasser, T. L. (1998). *Custodians of conscience: investigative journalism and public virtue*. Columbia University Press: New York.

- Hewett, J. (2013). Learning in progress: From computer-assisted reporting to data journalism via freedom of information, open data and more. In Mair, J. et al. (Eds.) *Data Journalism: Mapping the future*. Abrams Academic Publishing: Suffolk, 3–14.
- Holovaty, A. (2006). *A fundamental way newspaper sites need to change*. Holovaty.com, <http://www.holovaty.com/writing/fundamental-change/>, retrieved 20.1.2015.
- Kleis Nielsen, R. (2012). *Ten Years that Shook the Media World. Big Questions and Big Trends in International Media Developments*. Reuters Institute, Oxford. [https://reutersinstitute.politics.ox.ac.uk/fileadmin/documents/Publications/Working\\_Papers/Nielsen - Ten Years that Shook the Media.pdf](https://reutersinstitute.politics.ox.ac.uk/fileadmin/documents/Publications/Working_Papers/Nielsen_-_Ten_Years_that_Shook_the_Media.pdf), retrieved 20.1.2015.
- Mair, J., Keeble, R. L., Bradshaw, P. and Beleaga, T. (Eds.). (2013). *Data Journalism: Mapping the Future*. Abrams Academic Publishing: Suffolk.
- Meyer, P. (1973). *Precision journalism: A Reporter's Introduction To Social Science Methods*. Indiana University Press: Bloomington and London.
- Nygren, G., Appelgren, E., and Hüttenrauch, H. (2012). Datajournalistik – ett växande område. *Nordicom Information* 34(3-4), pp. 81–88.
- OECD. (2010). *The Evolution of the News and the Internet*. <http://www.oecd.org/internet/ieconomy/45559596.pdf>, retrieved 20.1.2015.
- Parasie, S. & Dagiral, E. (2012). Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists” in Chicago. *New Media Society*, 15(6), 853–871.
- Powers, M. (2012). “In Forms That Are Familiar and Yet-to-Be Invented”: American Journalism and the Discourse of Technologically Specific Work. *Journal of Communication Inquiry*, 36(1), 24–43.
- Riddell, P. (2014). *Impact of Transparency on Accountability*. In Bowles et al. (Eds.) *Transparency in Politics and the Media. Accountability and Open Government*. Reuters Institute for the Study of Journalism & I.B. Tauris: Oxford.
- Rogers, S. (2011). *Facts are sacred: The power of data*. Guardian shorts. The Guardian: London.
- Schudson, M. (2014). Origins of the Freedom of Information Act in the United States. In Bowles et al. (Eds.) *Transparency in Politics and the Media. Accountability and Open Government*. Reuters Institute for the Study of Journalism & I.B. Tauris: Oxford, 1 - 18.
- Sirkkunen, E., Aitamurto, T., and Lehtonen, P. (2011). *Trends in Data Journalism*. NextMedia D.3.2.1.2.B, Helsinki.
- Worthy, B. & Hazell, R. (2014). The Impact of the Freedom of Information Act in the UK. In Bowles et al. (Eds.) *Transparency in Politics and the Media. Accountability and Open Government*. Reuters Institute for the Study of Journalism & I.B. Tauris, Oxford, 31 - 48.