



## D.3.3.2.1 EXPERIENCES FROM COLLABORATION BETWEEN NEWSROOM AND DEVELOPERS

Author(s): Harju, Auli & Haara, Paula  
Confidentiality: Public  
Date and status: 21.12.2012 draft to be accepted

This work was supported by TEKES as part of the Next Media programme of TIVIT (Finnish Strategic Centre for Science, Technology and Innovation in the field of ICT)

**Version history:**

Version	Date	State (draft/ /update/ final)	Author(s) OR Editor/Contributors	Remarks
0.1		draft to be accepted	Auli Harju	

Participants	Name	Organisation
Tampere Research Centre for Journalism, Media and Communication (COMET)	Auli Harju	<b>University of Tampere</b>
	Paula Haara	
	Esa Sirkkunen	
	Antti Poikola	
Helsingin Sanomat	Esa Mäkinen	<b>Sanoma</b>
	Santtu Parkkonen	
	Tuukka Muhonen	

Next Media  
[www.nextmedia.fi](http://www.nextmedia.fi)  
[www.tivit.fi](http://www.tivit.fi)

## Johdon tiivistelmä

Tutkimuksessa tarkasteltiin ja arvioitiin HS Open –tapahtuman käytäntöjä, sen järjestämisen haasteita ja mahdollisuuksia. Keräämällä tapahtumaan osallistujien kokemuksia sekä haastatteleamalla HS Openiin vuosina 2011 ja 2012 osallistuneita keskeisiä toimijoita pyrittiin arvioimaan Helsingin Sanomien luomaa konseptia, jossa toimittajista, koodareista ja graafikoista muodostuvat tiimit kehittävät tarjolla olevan datan pohjalta uutissovelluksia. Lisäksi tutkimuksessa kartoitettiin näkemyksiä suomalaisen datajournalismin nykytilasta ja arvioita sen kehityksestä jatkossa.

Tutkimusaineiston muodostivat nettikysely toukokuussa 2012 järjestetyn vaalidata-aiheisen HS Openin osallistujille sekä tutkimushaastattelut. Tutkimus vastaa seuraaviin tutkimuskysymyksiin:

- Millaiset tekijät motivoivat osallistumaan HS Openiin?
- Millaisia kokemuksia osallistujilla on HS Openista, mitkä ovat tapahtuman keskeiset haasteet ja mahdollisuudet?
- Miten haastateltavat arvioivat datajournalismin ja avoimen datan nykytilaa ja tulevaisuudennäkymiä Suomessa?

Tulokset osoittavat, että Helsingin Sanomat ja HS Open ovat onnistuneet kokoamaan avoimesta datasta ja datajournalismista kiinnostuneet saman pöydän ääreen ja tukeneet näin kyseisen yhteisön muodostumista. Tapahtuma on tuonut alan osaajat yhteen ja mahdollistanut verkottumisen sekä toisilta oppimisen.

HS Openin käytäntöjen nähtiin haastatteluissa kehittyneen alkuvaiheista. Suurimmiksi haasteiksi nostettiin ajanpuute sekä tiimien työnjako työprosessin eri vaiheissa. Yksipäiväisen tapahtuman aikana tulee haastateltavien mukaan kiire kehittää toimiva demo uutissovelluksesta, etenkin jos tiimi lähtee liikkeelle ilman etukäteisideointia tai muuta valmistautumista. Työskentelyn aikana koodareitten koettiin olevan avainroolissa, he ovat usein ylityöllistettyjä HS Openin työprosessin aikana, kun taas toimittajien rooli korostuu alkuvaiheessa ja graafikoille puolestaan tulee kiire sovelluksen viimeistelyssä. Ratkaisuksi haastatteluissa ehdotettiin etukäteen valmistautumisen ottamista osaksi HS Openin ”virallista” työprosessia. Osa osallistujista ehdotti tapahtuman laajentamista kaksipäiväiseksi. Erityisesti toimittajien roolia toivottiin kehitettävän siihen suuntaan, että journalistinen näkökulma ohjaisi paremmin sovellusten kehittämistä ns. teknisen lähestymistavan sijasta.

Datajournalismin arvioitiin olevan yhä kokeiluvaiheessa Suomessa, mutta haastateltavat katsoivat kiinnostuksen avointa dataa ja datajournalismia kohtaan kasvaneen merkittävästi parin viime vuoden aikana. Datajournalismin arvon nähtiin olevan erityisesti demokratian ja läpinäkyvyyden lisäämisessä yhteiskunnassa, mutta haastateltavat myös uskoivat datajournalismin journalistiseen ja taloudelliseen merkitykseen.

## Executive Summary

This research describes and evaluates the practices, benefits and challenges of data journalism by exploring the experiences of HS Open hackathons organized by Helsingin Sanomat during the years 2011 and 2012. The research on data journalism, and especially on hackathon-type of events, focusing on the practicalities of working with the data, combining the work of different professionals in making a datajournalistic story etc. is still rare despite the increasing interest in data journalism and open data, and the existence of the tradition of data-driven journalism.

The research was conducted by observing HS Open hackathons and conducting a survey and research interviews with participants. The research questions were as follows:

- What motivates the participants in taking part of HS Open?
- What kind of experiences they have of the hackathons i.e. what are the benefits and the challenges of HS Open hackathons?
- How they see the situation of data journalism and open data in Finland at the moment and in the future?

The results prove that HS Open has been in the key role in organizing the community of interest around data journalism and open data. The hackathons have during these two years managed to gather the skilled professionals – either because of their own interest or sent by their employers – and enthusiastic open data activists. Besides combining their expertise in creating interesting pieces of data journalism they also use the event for networking and learning from each other.

The practices and working processes of HS Open has improved since the beginning. According to research data, the main challenges at the moment concern the timing of the work during the hackathon and the roles of the different professionals in working together. Some participants had experienced that one day is too short time for producing a functioning demo of a news application and they had noticed that the programmers of the app are overloaded with work whereas the journalists and graphic designers have smaller roles in the process. The solutions suggested included preliminary work before the hackathon, making HS Open a two-day event, and increasing journalists' role and adding journalistic emphasis on the working process.

At the moment, data journalism was considered to be still in an experimental phase in Finland, but the informants recognized the increasing interest for the field. They considered that data journalism supports transparency and democracy in the society and thus underlined both its societal and journalistic, but also potentially economic significance.

## Table of Contents

Johdon tiivistelmä.....	2
Executive Summary .....	3
Table of Tables .....	4
Introduction.....	5
1.1 Research Methods.....	5
Main results from the web survey and the interviews .....	6
1.2 Background .....	6
1.3 Why to participate in HS Open? .....	7
1.4 Brainstorming is team work, the rest is up to individuals .....	8
1.5 Meeting people and learning from them is crucial .....	9
1.6 Developing HS Open: suggestions.....	10
1.7 The significance of data journalism .....	14
1.8 Roles of the audience and the journalists .....	15
1.9 Data journalism in Finland now and in the future .....	16
Conclusions .....	17
Appendix 1: Research interview questions (in Finnish) .....	19
Literature .....	20

## Table of Tables

Table 1 Problems and suggested solutions.....	13
-----------------------------------------------	----

## Introduction

In the mid-2000s, especially in US and UK, news organizations have hired programmers – or so called programmer-journalists – in the newsroom. (Parasie & Dagiral 2012) These programmers are often involved in open source communities and are also advocates of open government. In Finland, the similar kind of development has been emerging, especially due to the enthusiasm of open data activists, starting from the Apps for Finland competition that was organized for the first time in the year 2009, leading to the organizing of Open Knowledge Festival (<http://okfestival.org/>) in Helsinki in September 2012, and founding of Open Knowledge Finland association in December 2012.

Helsingin Sanomat has been among the early developers of Finnish data journalism by organizing HS Open hackathons, the first of them taking place already in the spring 2011. So far there has been six HS Open hackathons, out of which the last one was organized during the Open Knowledge Festival in Helsinki with international participants. In one HS Open there are from 50 up to 80 attending, altogether the number of participants rises over 200.

In this research we explored the practices and experiences of HS Open hackathons by observing the hackathons, by conducting a web survey and research interviews with the participants of HS Open. We wanted to find out

- What motivates the participants in taking part of HS Open,
- What kind of experiences they have of the hackathons i.e. what are the benefits and the challenges of HS Open hackathons and
- How they see the situation of data journalism and open data in Finland at the moment and in the future.

The research has been conducted in collaboration with Helsingin Sanomat, especially Esa Mäkinen, Janne Kajjärvi, Santtu Parkkonen and Tuukka Muhonen, and from the open data activist side with Antti Poikola, who also took part in the research by observing the hackathons and conducting the web survey with the researcher. More information of HS Open and the news applications created in the hackathons, please visit: <http://blogit.hs.fi/hsnext/> in the Helsingin Sanomat website.

---

### 1.1 Research Methods

The research data consists of 6 interviews and a web survey conducted during the Election Data HS Open in May 2012. The interviews were conducted in May-June 2012 with HS Open participants (including a journalist, 2 graphic designers and 3 coders). They were asked about their experiences and ideas on HS Open, their thoughts on open data and data journalism in general and their expectations of the future developments of the field (see Appendix 1 for the questions).

The web survey form was created in Google Docs. The researchers participating in Election Data HS Open delivered the url ([http://bit.ly/datavaalit\\_kysely](http://bit.ly/datavaalit_kysely)) of the survey for the participants and also actively interviewed some of the participants by asking the survey questions and filling in the format in behalf of the informants according to the answers. The survey material includes 18 answers for questions about the election data and HS Open Hackathon.

In the web survey there were both multi-choice and open questions, altogether 10 questions. There were 18 participants who answered and half of them (9) were journalists. The others were coders (4), graphic designers (3) and others i.e. observers (2). The informants reported how many times they had attended HS Open Hackathon and the reasons they attended this particular hackathon. They named the most interesting election data sets available at the hackathon as well as got to suggest a missing data sets if they thought there should have been different type of data in the event. They also were asked about their interest in Election Data theme in general and finally, they could suggest some ways to further develop the HS Open Hackathon event.

Only four of the informants attended the HS Open Hackathon for the first time, most of the informants had participated at least once before the Election Data HS Open.

## **Main results from the web survey and the interviews**

### **1.2 Background**

HS Open Hackathon is a workshop type of event, where everyone interested in doing data journalism is welcomed. The event has had different themes based on the data sets available, such as 'Culture Data', 'Election Data', and 'Municipal Data'. The emphasis is on doing data journalism i.e. creating a journalistic story in a format of a visual application that combines data and presents in a visualized and often interactive way.

The participants sign up to the hackathon beforehand. Usually there are between 50-80 participants in one event, counting all the participants in the 6 HS Open hackathons the number rises over 200. There are always newcomers present, but many of the participants have attending more than one time. The participants are mainly journalists, coders and graphic designers, but also representatives of data providers, local and national administration, small enterprises, and research institutions. Work in HS Open hackathons are based on teams of three: The participants either arrive with their own team members or the organizer divides the participants into groups that include one coder, one journalist, and one graphic designer. These teams get to know the data sets provided by the organizer (the use of other data is also possible), brainstorm the ideas, and produce during the day so called news application that tells a visual story based on the data. The application is 560 x 400 pixels of size and it should work in Iframe and can be embedded and thus published in the HS.fi. In the end of the day all the teams present their app,

either a fully working version or at least a draft version of it. Helsingin Sanomat has bought and published some of the best applications.

The HS Open day starts with the presentation of the data sets available. The teams then gather in the working spaces and start brainstorming ideas. In the intensive working hours in the afternoon, a demo of the news application is created, and the hackathon ends with a session, where all the teams present their work, and an aftermath in a pub nearby.

### 1.3 Why to participate in HS Open?

In the web survey, the most common motivations to participate in the hackathon were (out of the given options) the possibility of learning something new (6 votes) and that the employer suggested participating (5). The election theme had inspired three of the informants, the networking possibility got two votes and the good earlier experiences of HS Open event motivated one of the informants.

In the open question, some of the informants elaborated further their motivations to participate. In these answers the learning from each other in the hackathon was considered valuable. Meeting new people, getting new ideas for one's own work, for instance for journalistic stories of election theme and for infographics, were also listed. The demos created in the event were well-liked. These answers emphasized the value of the hackathon in developing the professional skills of both journalists and graphic designers. Some of the informants described their general interest in open data and hackathons – they saw HS Open as an opportunity to get to know more and follow the development of the field.

According to the interviews, the reason for participating in HS Open was a combination of already existing interest in data journalism and expectations of the hackathon. Some of the interviewees wished to meet new people; the others had been working with open data and knew the potential participants beforehand.

For some of the interviewees coding and working with data is both a work and a hobby, for some more work than a hobby, but they all share an interest in data journalism and are interested in meeting and sharing ideas with the like-minded.

*“The strength of HS Open is that it collects [people –AH]. It is like a data journalism conference and a workshop at a same time.” (coder)*

*“You have to make difference between where you get your money from and what you are doing in your life. I've been trying to keep them as separate as possible but this [data journalism –AH] belongs to my societal-media-political life sphere. But all this is kind of cloud that has to do with each other.” (coder)*

As the quote above shows, especially the coders interviewed described, how open data and data journalism is a mixture of their work and their hobby. Their main work is something else, but it too links to the data journalism and open data in one way or another. And even if they consider attending HS Open as a free-time



activity, it still has positive effects on their working life, for instance in one interview the informant described how the collaboration in HS Open had led him traveling abroad to present Finnish open data cases. In a couple of other interviews, the interviewees mentioned “funny side projects” and work contacts that had been initiated in HS Open.

In the interviews, HS Open was also considered important for its outcomes and the ideas it produces. The demos are created in collaboration when different ideas and view points are brought together, and the results often end up being interesting, informative, and entertaining:

*“It is like up to the imagination, it demands that you open your mind, that how little data and little information you can end up with so interesting results. And if not interesting at least funny.” (coder)*

Learning was also considered as a motivation to participate in the hackathon. Since the participants work in different places and professions, the event gathers a lot of skills and experience around the table – a good opportunity for everyone to learn from each other:

*“Well, this shared playtime helps that people learn to do things together. I bet everyone got something out of yesterday [in the HS Open – AH]. If not in any other way, at least new working methods, new programs, a name one can contact.” (graphic designer)*

#### 1.4 Brainstorming is team work, the rest is up to individuals

The team members collaborate in creating the news app, and especially in the brainstorming phase their input is equal. However, the interviewees characterized the different roles in the team when describing the process of creating a news app:

**The journalist** is described as the one who is asking questions like what is interesting in this data. S/he is the one who knows journalism and where the audience’s interest lies, so the role for the journalist is often to search for the scoop. Sometimes the journalist also coordinates the work of the team, but not in all teams.

**The coder** has the skill of programming, s/he knows the tools and how to “talk computer”. The interviewees described the coders having a hobby-mentality towards the work in a good way, they are enthusiastic and come to the events because they like coding and solving the technical challenges. The coders’ work starts when the idea has been crystallized. They work with the data combining it and creating the functionalities of the app.

**The graphic designers** are the final pieces in the process. They give the idea for the visualization and finalize the app. Their skills lie in the understanding the design and visuality, and they enable the application’s appeal to the audiences.

The experiences of the team work varied but the interviewees thought that the teams functioned well in general, and that the number of participants in a team was sufficient, though one of the interviewees mentioned there could be more members in a team. However, according to one informant, the teams of three included “*enough people for brainstorming and not too much for making decisions*”.

The working processes in creating the app are, besides brainstorming, is somewhat sequential. This was raised as a challenge in several interviews:

*“(...) when there are teams of three, consisting of three expertise areas, necessarily there are times when someone doesn’t have work to do, you just have to wait. (...) Of course there could be discussion but when the work and the thinking process starts, it is really something you do alone. The expert does it. But you could intensify and improve it [the process – AH]. When you work in a team with strangers the working processes are not perfect, there are gaps.” (coder)*

The role of the coder was considered strategic in the interviews. The central role of the coders means that their workload is bigger than the journalists and the graphic designers during the hackathon. Furthermore, the graphic designers have to work in a hurry in the end of the day, in order to get his/her work done before the demo presentations. However, the informants did not find out any concrete solutions to improve the division of labour.

The experiences the interviewees had of the team work in HS Open had an influence on the development suggestions. In some of the answers the team work was considered to work well, in other answers the informants hoped for more motivated participants (“people who are not there only because their employer has demanded them to participate”). The overall development of HS Open hackathons was recognized, the organizing (facilities, timetable, data sets) had improved on the course of the events, and the gender division balanced from the beginning.

## 1.5 Meeting people and learning from them is crucial

The interviewees also raised the issue of the interaction between the teams. The interaction was seen important, for different teams could be dealing with similar problems and trying to find out the solutions in their separate rooms instead of collaborating between teams and sharing the solutions. One interviewee stated that the concrete space where the hackathon takes place does not support the communication between the teams, for the teams are working in separate rooms, and meet each other only during the lunch break. On the other hand, another interviewee stressed that the communication is not even possible for the limited time resources: the teams are so busy in each phase of the work that they have not time to discuss with the others.

The interviewees highlighted the importance of the workshop method: they appreciated working together. It enabled learning from the others, and also

observing how the others work with the data and different tools. Since there were three different professionals in the teams, the interviewees felt they got new view points and understanding of each others' work. Moreover, one of the interviewees described that in the teams there was a new collaborative working culture emerging.

Meeting new people was also considered rewarding. One of the interviewed coder especially enjoyed working in a team in which he does not know the members beforehand:

*"I find it [working with strangers -AH] appealing. In my opinion it is not so fruitful to go there in the HS Open to work with prepared results and the team. For me it is more meaningful to go and meet new people and new tools."*

The networking and brainstorming continue after the hackathon in the aftermath session in a pub. That is the place for discussions about new forms of collaboration, new ideas, developing the apps further, sharing information about new tools and describing examples of open data and data journalism elsewhere.

The well-functioning team work is also important in order to create the news apps during the one day of HS Open. All the team members manage their own area of expertise, and combining the knowledge in collaboration is necessary: not everyone has to know everything, and few actually do in the field of data journalism, as one of the interviewees stated. Newcomers are also welcomed: one of the graphic designers suggested that a team could consist of the experienced and of those who want to learn. In this way the knowledge and skills would transfer to novices as well.

Some of the interviewees mentioned that they make use of the solutions people invented during the hackathon - or knew already from some other context - later on in other projects. In this sense, HS Open seems to expand the participants' professional skills: they take the new ideas and knowledge with them and use them in their work and other projects. This seems to apply especially for coders and the technical challenges that they face when coding.

## 1.6 Developing HS Open: suggestions

In the interviews, though the interviewees gave good marks for HS Open in general, there were several issues that were considered problematic and worth further developing. Next we point out some concrete problems experienced by the participants during HS Open. The problems and suggestions for the solutions are also presented in a table format in the end of part 3.4.

There were only few concrete development suggestions in the web survey. Some informants hoped that the teams of three would meet before the hackathon. Coming to the event "with empty hands" was a problem for one of the informant, who, knowing that the organizers recommend discussions before the hackathon, suggested that the organizers would concretely increase the collaboration and

interaction between team members. Those, who did plan their work earlier together, considered it beneficial and suggested it to become an official part of the HS Open working process.

In addition, some of the informants wished more time for discussions and questions, especially when presenting the demos in the end of the day.

The informants wished for more integration to the daily journalism. Since the demos presented in the end of the hackathon were considered interesting and well-done, one informant wanted to see more of them published in Helsingin Sanomat. One informant wished more media publicity in general for the event.

The interviewees pointed out some imbalances in the working processes. According to the journalist interviewee, since the journalist's role is mostly focused in the beginning of the work, s/he should prepare as much as possible before the hackathon. She was a bit annoyed for the lack of work the journalists experienced towards the end of the hackathon. The interviewee pondered if it would be possible to increase the role of the journalist somehow during the HS Open. This could also strengthen the connection between HS Open and daily journalistic work.

One informant emphasized that the team composition should be well planned: in every team there should be someone "who knows what to do", the new-comers then easily get the idea from the experienced. According to the answers, the new-comers' presence was welcomed; the hackathon had remained open and not formed an insider circle.

**More emphasis on journalism and visualizations.** The interviewees noted that HS Open is too much concentrated on technical side i.e. programming. Coders are the most active people who participate and the informants hoped that journalists and graphic designers would get more involved. It could change the character of the event in a positive way. The journalist suspected that if more journalists participated it could "*clarify those starting points and journalistic doing*". At the same time, one graphic designer complained that graphic designers haven't been too enthusiastic to participate in HS Open.

One coder assumed that the event may interest specifically coders because their job already includes working with the data. Data journalism requires that you understand data and technology, and consequently doing data journalism requires from journalists that they adopt new working methods. On the other hand, it's worth discussing why HS Open attracts only few graphic designers even if they too work with the data.

**More participants outside Helsinki and HS.** One interviewee suggested that it could be possible to get participants from different areas of expertise if there were co-operation with some learning institution. One informant said that some universities already give credits if you participate in HS Open. There have been some student groups attending in HS Open at least from Aalto University and University of Tampere. In HS Open the students can learn how to solve problems

in practice. The question is how one can combine theoretical side of data journalism and the work in HS Open. According to one interviewee, it could be one focus of development.

The interviewees hoped that people would participate also from other organizations than Sanoma. One informant mentioned that some of the Sanoma's journalists and graphic designers are there mostly on behalf of their job, not because of their passion for data journalism. This may reflect in the working atmosphere.

Two informants suggested that HS Open could also take place in another city than Helsinki. However, it was said that the event should be near Helsinki. Tampere was mentioned as an example of a good location for organizing HS Open. Organizing the event elsewhere than in Helsinki could be a way to get new people involved.

**More time, a two-day event?** The most common problem that the interviewees mentioned was that the time runs out during the hackathon. That's the main reason why the outcomes remain unfinished. The outcomes of the HS Open get quite a little attention because people don't want to offer them for publishing when they still are unfinished. Informants hoped that there would be more time to use. At the same time they noted that many people don't have time to sit even a one whole day in a workshop.

If the event lasted two days, there could be more time to prepare and work. Many of the interviewees supported the idea of a two-day event. One interviewee added that it could make the participation more difficult for those who live further away from Helsinki, for they have to arrange accommodation, dining and evening program.

Beforehand preparation was mentioned as another way to make sure that the time wouldn't run out. One could also increase the quality of the outcomes by preparing in advance. On the other hand, one informant explained how his own working methods do not include preparation in advance for it does not always lead to good results:

*"I have used to that there can be within an hour warning a big piece of news and that you just have to react then. If you plan a lot in advance, it probably won't come through in a way I originally thought. Or then the schedule won't hold. [...] The most ironic thing was [in a HS Open hackathon – AH] that the team which started already on Wednesday had the most unfinished demo."* (graphic designer)

The informants mentioned also difficulties concerning preparation in advance. Again the biggest problem concerned people's limited resources of time. It's difficult to find time to do it in the pressures of daily work. The fact that the team members do not always know each other beforehand will not make it easier either.

*“It’s very difficult as a group, if the team members don’t know each other beforehand. I’ve found it very difficult. [...] There is a very big difference in quality [of the outcome – AH] when you make some preparations. (coder)*

When it comes to improving the interaction between the teams it was suggested that one of the organizers could be a mediator taking care of the communication between the teams.

**Data cleaned up better.** The data available was evaluated as interesting but some of the interviewees criticized the quality of the data. They wished the data to be better cleaned beforehand, for otherwise it may take a lot of time for the coder to clean up the data, in order to be able to use it. The interviewees however pointed out that the problem of the messy data can be difficult to solve, for it would take a lot of resources and preparation from the organizer and data providers to make available clean ready-to-use data sets.

**Table 1 Problems and suggested solutions**

Issue	Problem	Reasons	Solution
Interaction	Lack of communication between teams	Lack of time during the day Teams working in different rooms	More time for discussions A mediator person between the teams
Participation	Not enough new faces Lack of enthusiasm among some participants Not many graphic designers attending	Small circles of the interested Some coming because of the employer’s demands	Co-operation with universities, inviting students, adding theoretical viewpoints Novices teaming with the experienced Inviting journalists from different media houses Organizing the hackathon in some other city than Helsinki More publicity for HS Open in the media
Division of labour	Coders busy most of the time, others having time doing nothing (journalists in the end, graphic designers in the beginning)	Technical issues (coding) dominate	Increasing the role of the journalist in the hackathon Preparing beforehand

Working process	Lack of time	Starting from the scratch Only few hours time to work	Beforehand preparation as part of the hackathon process Two-day-hackathon
Results	Demos unfinished Demos lacking connection to daily journalism	Not enough time during the hackathon No resources (time, motivation) after the hackathon	Better preparing beforehand Helsingin Sanomat publishing more of the demos created in HS Open
Data	Data is not in a ready-to-use format	Quality is poor because of the providers	Data cleaned up better beforehand (needs resources)

## 1.7 The significance of data journalism

In the research interviews many of the informants thought that you cannot do data journalism without understanding the data and its features. However, it was stated that data journalism requires know-how that a single person can rarely master well. Data journalism combines data usage to journalistic aspects, and then there is the technical part that must be handled in one way or another, either with easy-to-use tools that a journalist can manage, or then the media houses need to hire programmers to do the coding.

*” The challenge of Helsingin Sanomat is of course that it [data journalism - AH] certainly requires new know-how because basic journalist and graphic designer can’t code. It means that you can’t work that data straight away. So, either journalists or graphic designers learn how to code or they employ guys who take care of the data transformation. Surely everybody should know a bit about everything there too. So that they would understand the potential.” (coder)*

Besides technical ability, the interviewees emphasized that you should have a vision of what you want from the data. On the other hand, it was also said that you cannot always know beforehand which data will be interesting or will be meaningful. One interviewee said that you can get confusing ideas by using confusing data. In other words, you don’t always have to know beforehand what you want from the data and sometimes the data ends up being not so interesting after all.

**Social, economic and journalistic significance.** The informants were of the opinion that data journalism has social significance. Opening of data was seen as a way to increase transparency and democracy in the society. Especially information that is produced by public organizations with tax resources was considered the kind of data that should be opened. It was also argued that citizens

have the right to know which data is behind the legislative decrees. One coder thought that transparency in the society would also increase if media opened the backgrounds and the data of their news reporting to the audience. On the contrary to this, another coder said that in his opinion journalists should not necessarily have to reveal everything to the audience in order to preserve the status of journalism.

Besides social significance, the informants considered the economic opportunities of data journalism. One coder for example said that data journalism is a rewarding way to do that kind of applications which are also profitable. On the other hand, he noted that the outcomes of data journalism are not always applications or visualizations. The final result can also be a story produced by a journalist after doing the digging in the data. Another interviewee mentioned that data journalism could develop business activity to small firms.

The interviewees also considered the ways in which media companies could benefit from data journalism. Open data was seen to offer interesting possibilities for doing journalistic stories. Data journalism was also mentioned to be a way to stand out of the crowd of competitors because you cannot be copied as easily as regular news information.

The journalist regarded data journalism as something that will have revolutionary consequences. For her processing of open data is a possibility to access information which may have journalistic value. A graphic designer on the contrary stated that data journalism isn't something new per se. It is just one way to get information. He presumes that utilizing the data is not going to change journalism, for it only deepens journalism and gives it added value. Using data may not be as central part in the work of journalists, whereas graphic designers' work has already included working with the data.

## 1.8 Roles of the audience and the journalists

Many of the interviewees thought that the audience is not interested in data journalism or the background information about making it. Instead, the audience is interested in the stories which are produced using datajournalistic means. For example, a coder-interviewee working in a media house said that their datajournalistic articles are read more compared to the other news articles.

One coder mentioned crowdsourcing as a method for data journalism. There exists a vast quantity of data and any data could be interesting to some people. When the audience digs into data they may find something interesting. This kind of crowdsourcing requires that the data is open to everyone who is interested in it. In other words, open data is needed for it allows anyone to do a data journalistic story.

Another way to approach journalism that engages audiences is to emphasize journalist's role. A coder told that he prefers that journalists choose topics for the stories because they know better what interests the audience than he does. In that



way data journalism is a tool for journalists who create the stories for readers. The journalist-interviewee hoped that the data could be open in particular for journalists to use because they – and journalism as an institutionalized profession - have the journalistic responsibility:

*”Those, who sit on the data, should maybe think who they rather give that data to, if someone will anyway take it. Because data security will never be so good (...) journalists have codes for what you can and what you cannot do. So better this way, open everything to journalists. Or do you involuntarily open it to hackers? Rather to those journalists.” (journalist)*

Hence, there are two different views about who should be doing data journalism. The first quoted coder hoped the democratizing of the data and that everyone could take part in doing data journalism. The latter coder and the journalist considered data journalism from the point of view of the journalistic profession. Journalists’ professional status could be threatened if everyone did (data) journalism.

## 1.9 Data journalism in Finland now and in the future

Data journalism in Finland was considered in the interviews as a new field which still is in an experimental phase. However, the interest in data journalism has grown in a short period of time, for example among journalists and students. The interviewees mentioned The Guardian and The New York Times as the media houses where doing data journalism is already an advanced practice.

The interviewees perceived that data journalism in Finland depends on few experts at the moment. A graphic designer noted that doing journalism and telling socially meaningful stories by using data journalistic means still happen rarely in Finland.

The growing interest in data journalism in Finland is not alone enough for data journalism requires special skills and resources. A graphic designer pointed out that data journalism could benefit especially small media houses because it is quite a cost-effective way to do stories. Sometimes accomplishing a story may only take a minor effort. However, smaller media houses do not necessarily have sufficient resources or experts for doing data journalism.

Several interviewees mentioned that open data and open tools are necessary for data journalism to develop. At the same time it was noticed that data journalism can also be based on closed data that the journalist is digging out by doing data inquiries.

However, the interviewees hoped that those who administer data would make decisions to open it. They mentioned Statistics Finland, the City of Helsinki, the National board of patents and registration of Finland, the government, local authorities, Yle and companies as examples of data holders. Reasons that were mentioned as obstacles for opening data included privacy, financing and pleading

licenses. The informants also mentioned other problems concerning open data in Finland, such as a need to advance the ways you can publish data, find data, clarify data and sell applications. One coder presented an idea about centralized data site where you could also make surprising discoveries. Another interviewee mentioned that there is a demand for trading centre that could ease the selling of the applications and finding partners for co-operation. He also considered if media houses should create a network for co-operation in data-journalism.

**Openness is an important value.** All the informants advocated open data and they took positive attitude to openness. One graphic designer questioned openness as an absolute value. He said that data should be easily available whenever somebody needs it but you also have to know how to interpret it and how to find meanings in the data in order to benefit from it.

The interviewees saw no problem or only slight problems concerning open data but they associated openness with unexpectedness and experimenting. They reminded that it is difficult to foresee all the consequences of data journalism and open data for they can also generate surprising results. It was also noted that these surprising results are not necessarily exclusively positive. When you combine different data you may be able to conclude something that you should not publish because of the protection of the personal data and privacy. Nevertheless, the interviewees hoped that those who administer data would open it quite carelessly.

## Conclusions

All in all, HS Open is considered important and organization of the event was evaluated as successful. The overall opinion among the informants was that the work processes were well thought through and the event organizations had evolved from the beginning. Since each HS Open hackathon draws new people to attend, it was considered as a proof that the event still offers interesting themes and data, and thus motivates people to participate. However, some of the informants felt that it is a small circle of people that is interested in data journalism and thus attending HS Open.

HS Open was considered important for the data journalism and open data community as well as for personal development. The informants enjoyed the informal atmosphere of the event, which makes meeting new people and networking easy. They especially appreciated the fact that there are leading experts present which enables learning from each other while working in the teams and thus developing one's own professional skills.

The main challenges of HS Open had to do with the short time available for working on the news application ideas. The division of work during the hackathon was experienced somewhat imbalanced: the coders are the key figures in creating the news app which leads to other team members having partly lack of work (journalist), partly pressure to do their work in a hurry (graphic designer). However, the coders' good attitude and enthusiasm was well recognized in the interviews.

The lack of time causes that the outcomes are left unfinished and the work for finalizing them does not continue after the hackathon. The demos were liked and the informants wished to see more of them published. Furthermore, some of the informants predicted that the interest in attending HS Open may decrease if the data journalism community remains small and forms an insider circle in which you find it difficult to enter as a newcomer.

The development suggestions for HS Open included more time for the event and/or stronger demand for beforehand preparations. The informants wished to see new participants and suggested that Helsingin Sanomat could invite students and journalists from different media houses to attend. Newcomers could then team up with the experienced in order to learn the skills and practices. Wishes for the further development also included data sets cleaned better by the data providers and stronger connection between HS Open and daily journalism. The latter could be done by increasing the role of the journalist in the hackathon.

Most of the interviewees linked data journalism to their interest in the society. They saw that opening public sector data can increase transparency in the society. The connection between data journalism and the society is strong for many reasons. First, data journalistic stories tend to target on societal issues, for the data sets are usually numbers and facts on the society. Secondly, most of the data available comes from the public sector. And finally, open data movement can be described as a social movement and its demands for open data aims at increased transparency in the society. In a way, open data movement has started to change the position of journalism in the society. Journalist are not alone anymore in their work serving the democracy and being on the watch for corruption, for open data and open government activists have stepped in by actively searching and publishing information – which used to be the privilege of professional journalists.

## Appendix 1: Research interview questions (in Finnish)

### HSOpen kehittäjäyhteisö: muokattu kysymysrunko

#### 1) Kokemukset HSOpen-tapahtumista

- \* miksi osallistuit? Mikä datajournalismissa sinua kiinnostaa?
- \* kuinka monta kertaa olet osallistunut HS Openiin?
- \* mitä olet oppinut?
- \* millaisia kontakteja olet luonut?
- \* mitä etuja ja haittoja koet siitä, että Helsingin Sanomat on tilaisuuden järjestäjä?
- \* mikä on tärkein osuus HS Openissa, miksi? (alun esitykset; sovellusten ideointi&toteutus; sovellusten demoaminen; oluen juominen...)
- \* Miten **työprosessit** ovat toimineen HSOpenissa (eritele missä järjestetyistä)?
  - miten ”kolmikkomalli” on mielestäsi toiminut? (missä roolissa olit itse?)
  - ovatko työmäärät olleet sopivia kaikilla?
  - miten työnjako ja työskentely on sujunut mielestäsi?
- \* Miten HSOpenin esivalmistelut (myös kolmikoiden) tulisi järjestää? Mitä kanavia kautta vuorovaikutus tulisi järjestää, entä miten aiheiden suunnittelu ja valinta tulisi ottaa käsittelyyn ennen varsinaista tilaisuutta?
- \* Mitkä ovat olleet käytännön työskentelyn ongelmakohtia? Kenen vastuulla niiden ratkaisemisen tulisi olla? Miten ongelma voitaisiin ratkaista?
- \* Ovatko HS Open ja siellä tehdyt työt saaneet tarpeeksi julkisuutta?
- \* Mitä mieltä olet (kevään 2012) sovelluskilpailusta? Oliko onnistunut ideana, miten toimi käytännössä?

#### 2) Näkemykset datajournalismista

- \* Mikä on mielestäsi HSOpenin kehittäjäyhteisön koko. Ketkä viisi nostaisit keskeisiksi henkilöiksi yhteisössä?
- \* Mitkä asiat (ilmiöt, teot, tahot) näet merkittäviksi, jos puhutaan datajournalismin kehittämisestä Suomessa? Miksi juuri nämä?
- \* Entä maailmalla?
- \* Miksi datajournalismi on tärkeää? Kenelle se on erityisen merkittävää (kuka on kohdeyleisö?)
- \* Mitä avoimuus datajournalismissa sinulle merkitsee? Mikä siinä on hyvää? Millaisia ongelmia datan avaaminen voi saada aikaan?
- \* Mihin suuntaan datajournalismi on sinusta kehittymässä? Miten haluaisit vaikuttaa / mitä haluaisit muuttaa tässä kehityksessä? Miksi?

#### 3) Kehittämisideat ja ehdotukset

- \* Miten HSOpenia tulisi tapahtumana kehittää? Miksi? Mitä lisää, mitä pois, mitä muuttaisit?
- \* Millaiseksi koet osallistujien ja Helsingin Sanomien suhteen? Miten sitä voisi kehittää?
- \* Mitä odotat Helsingin Sanomilta jatkossa?
- \* Mitä itse toivoisit voivasi tehdä/ tarjota jatkossa?
- \* Tiesitkö, että HS maksaa julkaistuista töistä? Mitä mieltä olet käytännöstä?
- \* Haluaisitko myydä työtäsi Helsingin Sanomille tai muille mediayhtiöille? Oletko myynyt? (Millaisia kokemuksia?)
- \* Millaisia bisnesmalleja ja työnjakoa haluaisit datajournalismin ja avoimen datan yhteyteen syntyvän, miksi? Mikä olisi oma roolisi tässä?
- \* Tiedätkö joitakin olemassaolevia käytäntöjä, mitä? Mikä niissä hyvää, entä ongelmallista?
- \* Millaista avointa dataa / datajournalismia ehdottomasti kaipaisit juuri nyt?

## Literature

Parasie, S. & Dagiral, E. (2012) Data-driven journalism and the public good: “Computer-assisted-reporters” and “programmer-journalists in Chicago. *New Media & Society Online First*, 1-19. (<http://nms.sagepub.com/content/early/2012/11/15/1461444812463345.full>)  
Retrieved: November 18, 2012).

---