Seznam příloh

Příloha č. 1: Mapa cholery doktora Snowa (obrázek)

Příloha č. 2: Rozdíly tradiční a datové žurnalistiky (tabulka)

Příloha č. 3: Obrácená pyramida datové žurnalistiky (graf)

Příloha č. 4: Komunikační strategie datové žurnalistiky (graf)

Příloha č. 5: Postup zhotovení článku datové žurnalistiky (graf)

Příloha č. 6: Mapa výsledků voleb na Ukrajině v roce 2012 (obrázek)

Příloha č. 7: Faktory ovlivňující úspěch „big data“ v organizaci (graf)

Příloha č. 8: Tematická skladba článků (graf)

Příloha č. 9: Forma vizuálních výstupů (graf)

Příloha č. 10: Podíl interaktivních vizuálních výstupů (graf)

Příloha č. 11: Mapa MHD (obrázek)

Příloha č. 12: Srovnání podmínek v České republice a v Austrálii (tabulka)

Příloha č. 13: Srovnání podmínek v České republice a v Austrálii (graf)

Příloha č. 14: Porovnání postavení datové žurnalistiky v ČR a v AU (tabulka)

Příloha č. 15: Srovnání týmů v České republice a v Austrálii (tabulka)

Příloha č. 16 Softwarové nástroje pro zpracování dat (tabulka)

Příloha č. 17: Osobní rozhovor s Nickem Evershedem (text)

Příloha č. 18: Osobní rozhovor s Craigem Buttem (text)

Příloha č. 19: Osobní rozhovor s Marcem Moncriefem (text)

Příloha č. 20: Osobní rozhovor s Margaret Simons (text)

Příloha č. 21: Osobní rozhovor s Janem Cibulkou (text)

Příloha č. 22: Osobní rozhovor s Marcelem Šulkem (text)

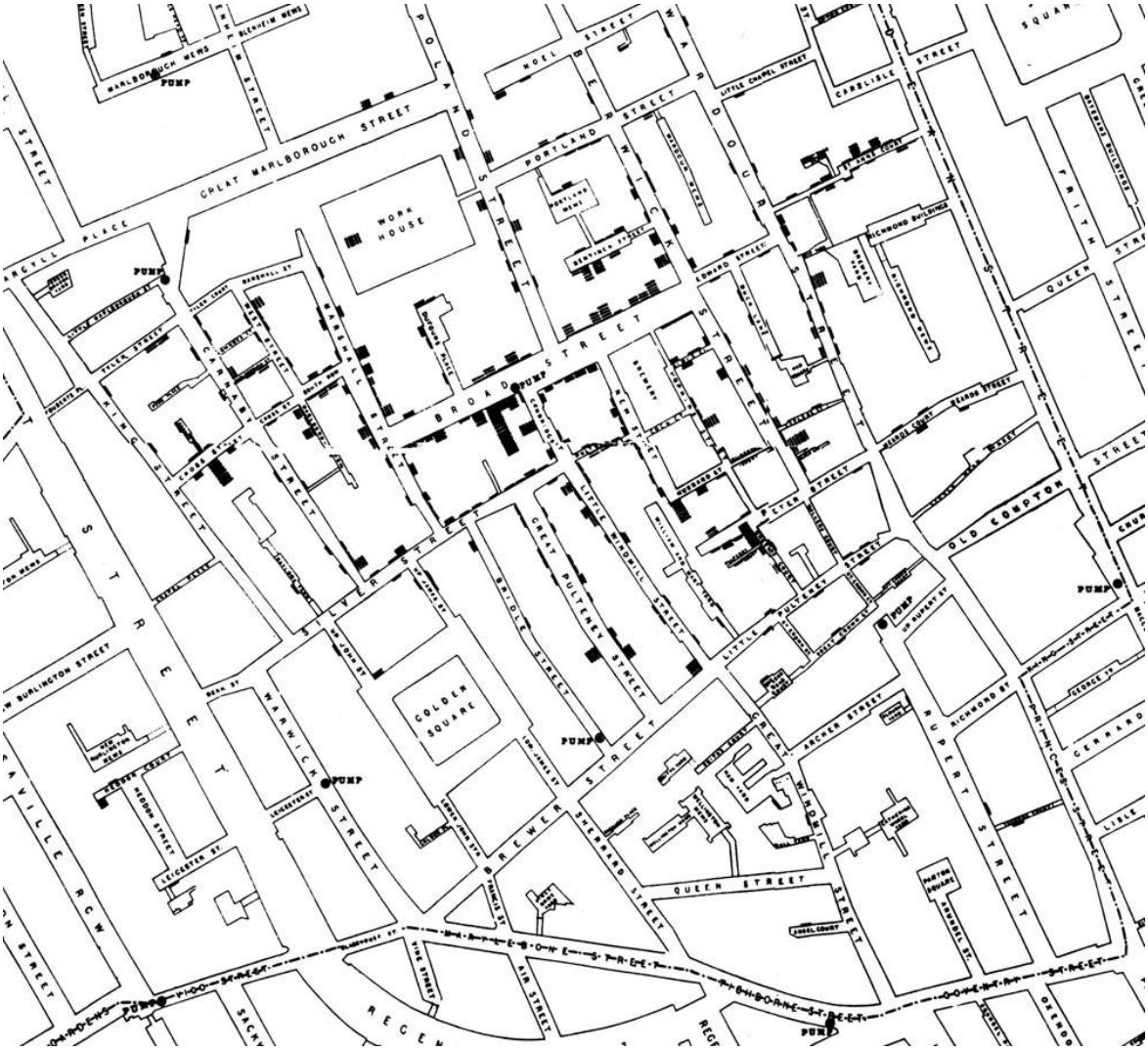
Příloha č. 23: On-line korespondence s Filipem Rožánkem (text)

Příloha č. 24: On-line korespondence s Petrem Kočím (text)

Příloha č. 25: Osobní rozhovory se studenty žurnalistiky (text)

Přílohy

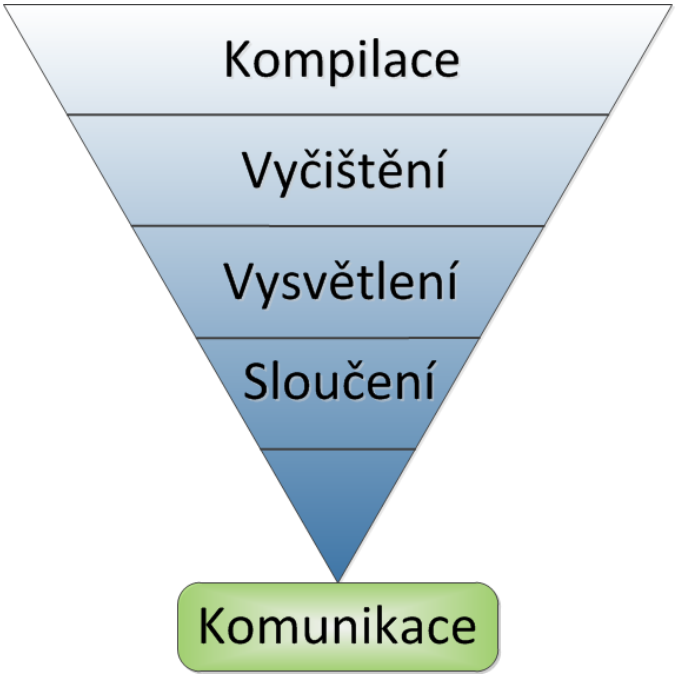
Příloha č. 1: Mapa cholery doktora Snowa (obrázek)



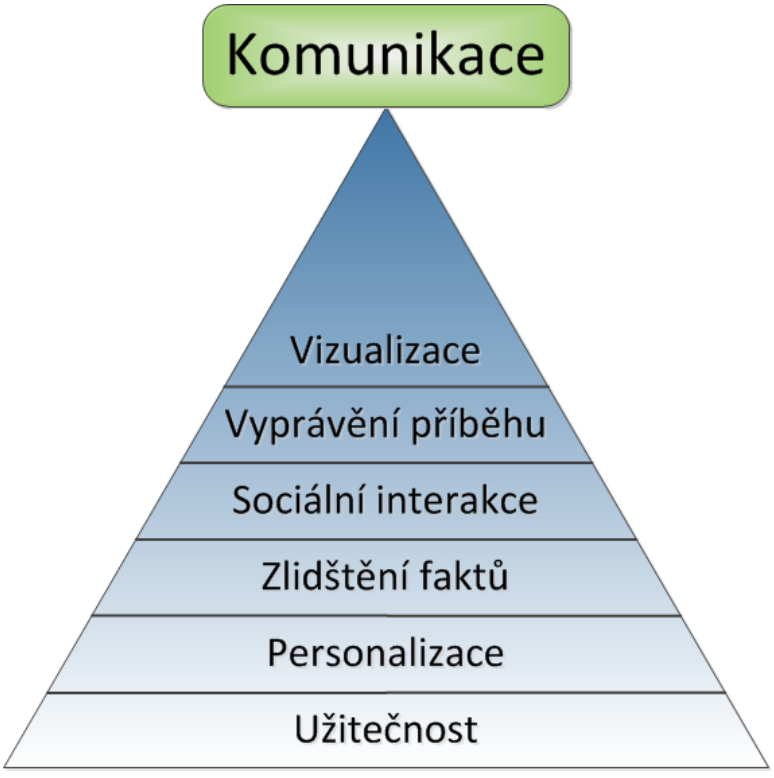
Příloha č. 2: Rozdíly tradiční a datové žurnalistiky (tabulka)

Rozdíly tradiční a datové žurnalistiky			
Faktory		**Tradiční žurnalistika**	**Datová žurnalistika**
Zdroje		Primárním zdrojem jsou rozhovory.	Primárním zdrojem jsou data.
Vzhled článku		Primární výstup je text, který může být doprovázený fotografií nebo videem.	Škála vizualizací – od grafů po interaktivní aplikace, jež mohou být doprovázené textem.
Role čtenáře		Čtenář je pasivní. Informace přijímá.	Čtenář je aktivní. Selektuje informace, o které má zájem.
Komunikace		Komunikace je jednosměrná.	Komunikace je více interaktivní. Dvousměrná.
Role žurnalistiky		Referuje o něčem, co se stalo.	Odkrývá nová témata, sleduje vývoj a vysvětluje souvislosti.
Dopad		Informuje.	Otevírá diskusi.
Objektivita		Objektivita závisí na důvěryhodnosti zdroje, který může informace zabarvit.	Větší míra objektivity. Odkrývání souborů dat a metodologie čtenářům.
Dovednosti novináře		Dovednost psát, selektovat informace, kritický pohled, umět se ptát na otázky.	K tradičním dovednostem novináře programování, vytváření vizualizací, základy statistiky.
Tvorba	**Nástroje**	Textový editor.	Různé nástroje, programy, databáze.
	Lidé	1 novinář, případně více novinářů.	Tým složený z novináře/ů, programátora, statistika, designéra.

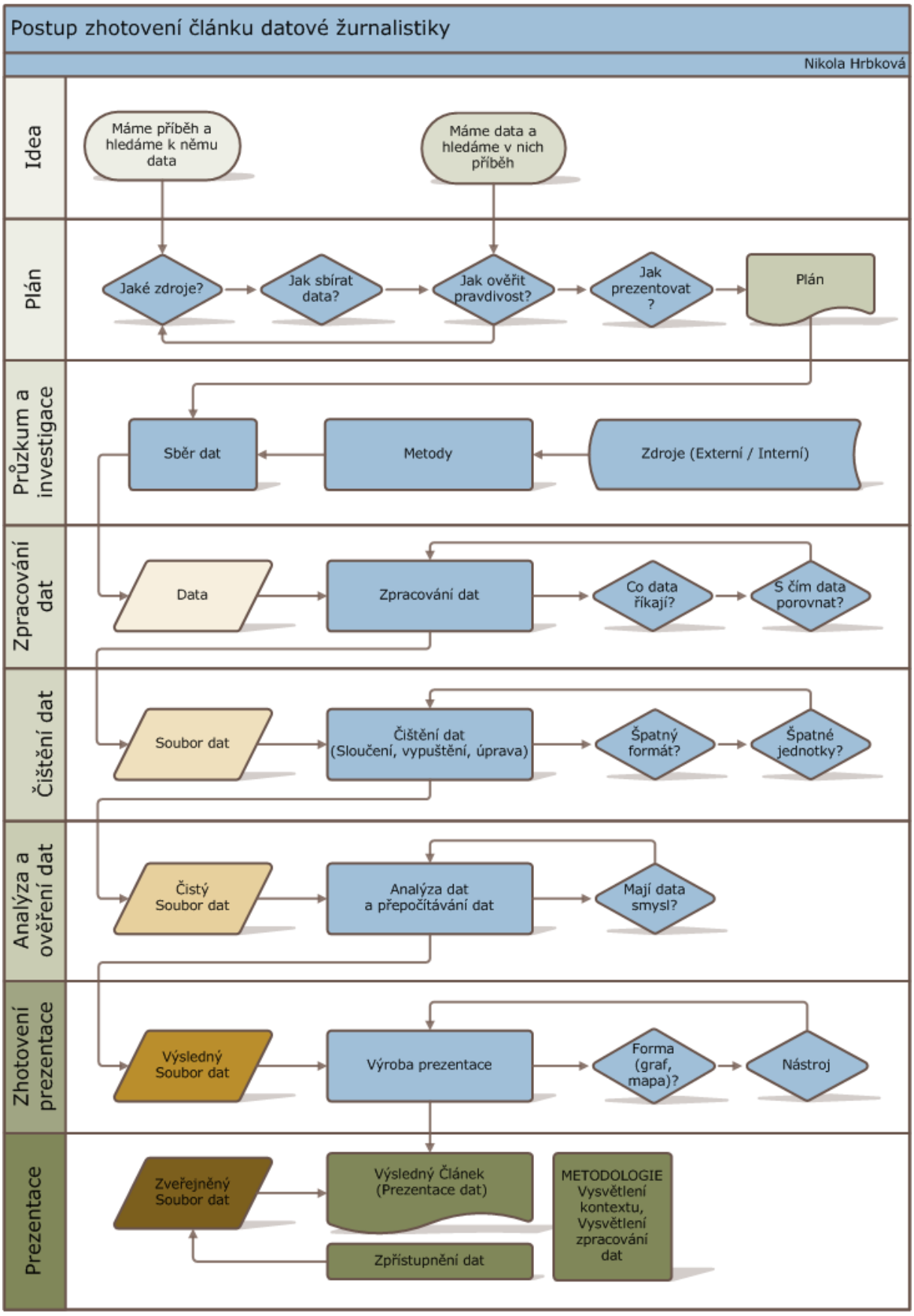
Příloha č. 3: Obrácená pyramida datové žurnalistiky (graf)



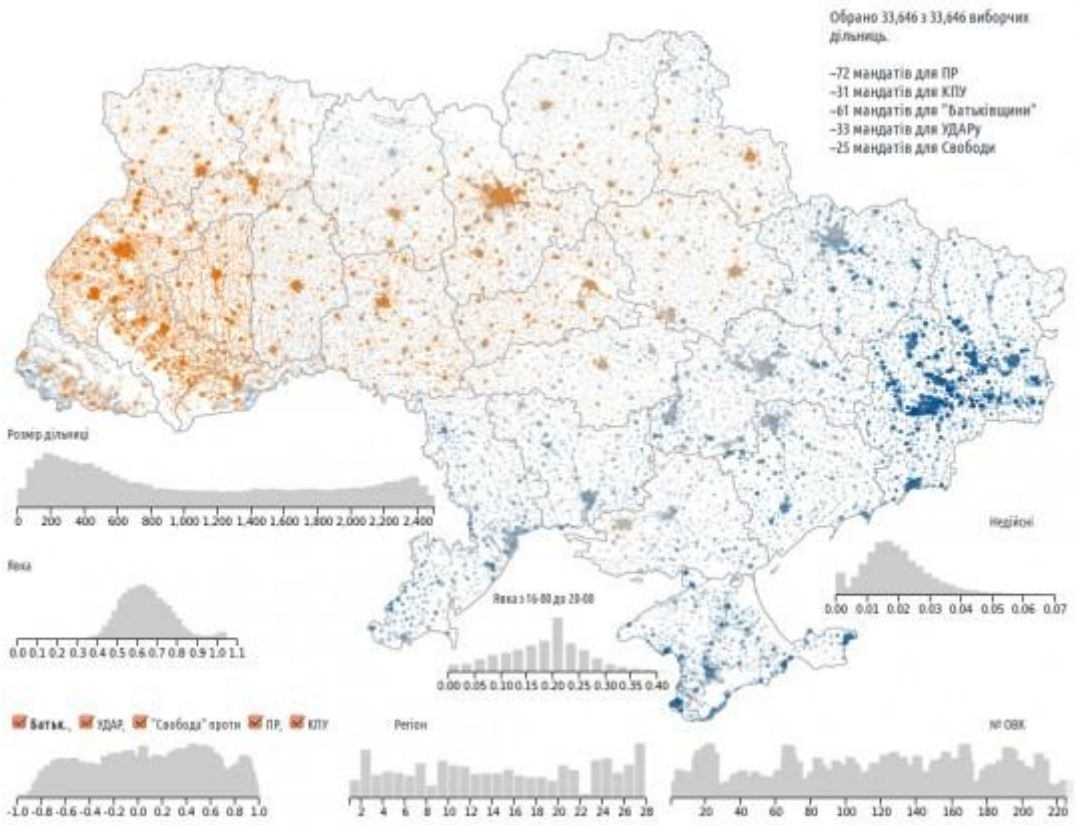
Příloha č. 4: Komunikační strategie datové žurnalistiky (graf)



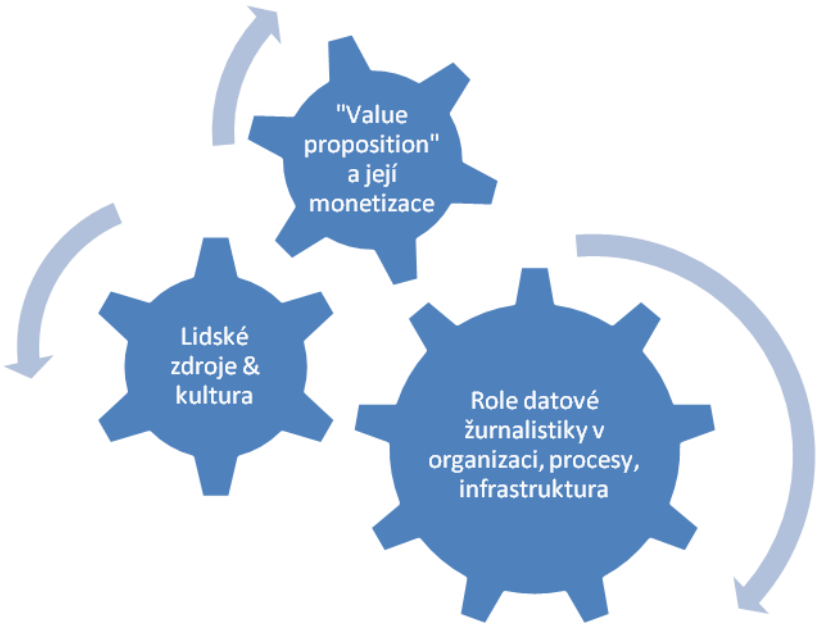
Příloha č. 5: Postup zhotovení článku datové žurnalistiky (graf)



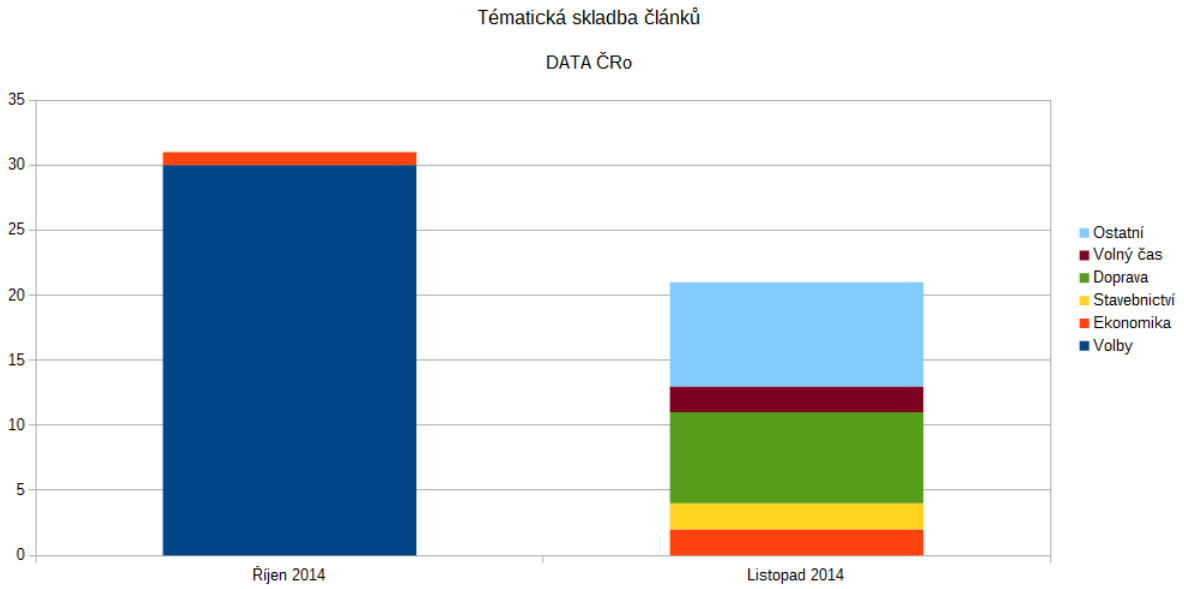
Příloha č. 6: Mapa výsledků voleb na Ukrajině v roce 2012 (obrázek)



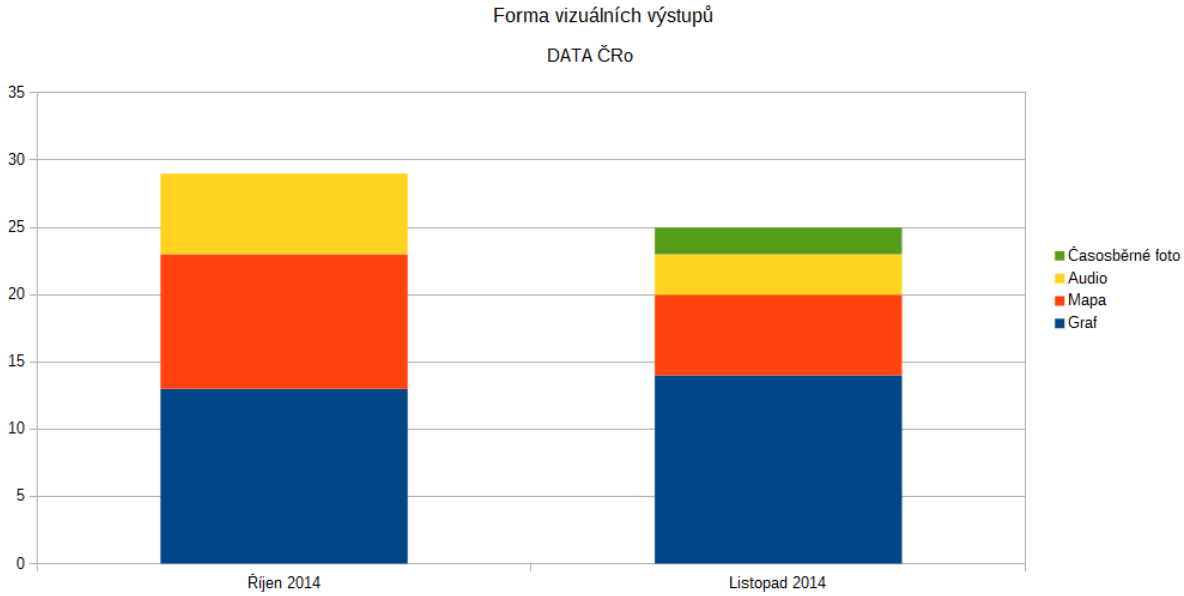
Příloha č. 7: Faktory ovlivňující úspěch „big data“ v organizaci (graf)



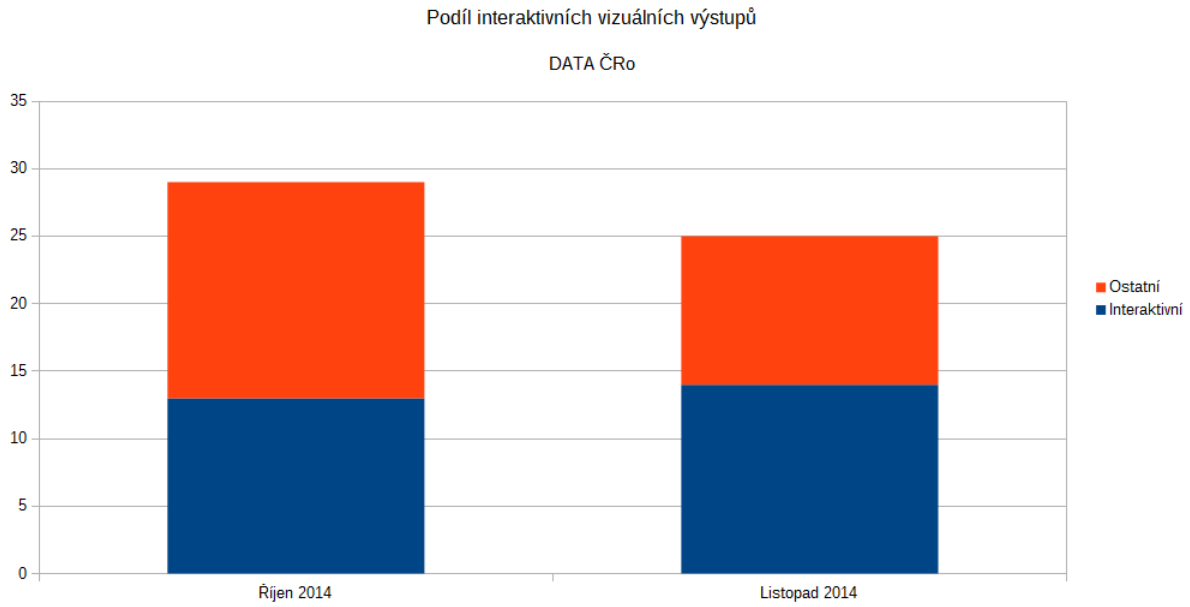
Příloha č. 8: Tematická skladba článků (graf)



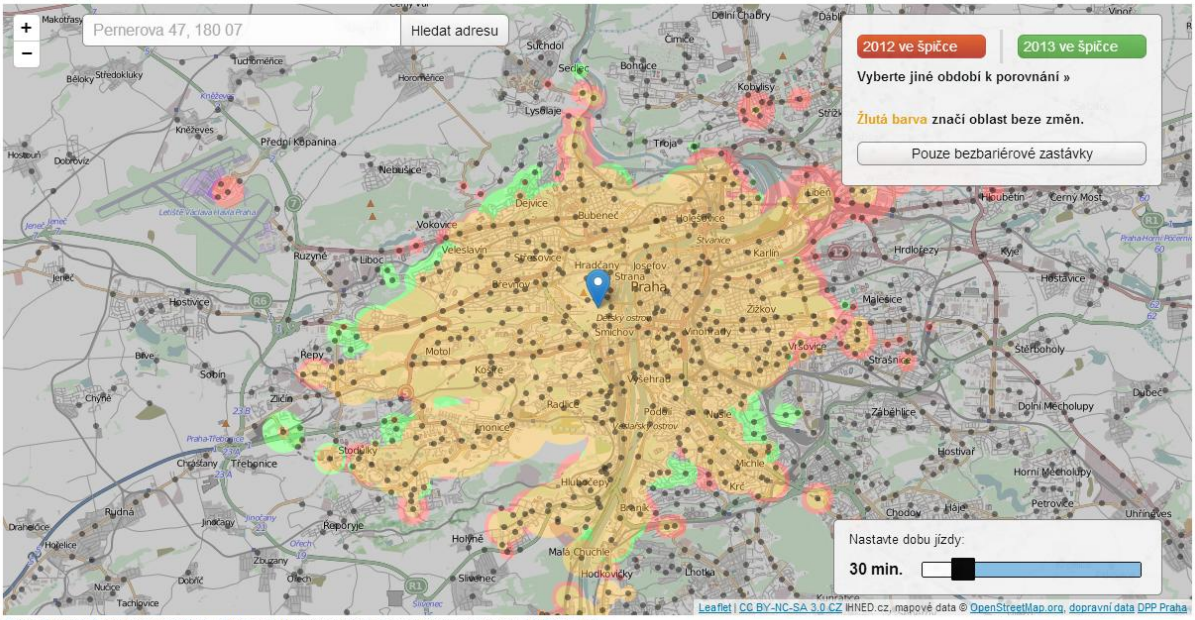
Příloha č. 9: Forma vizuálních výstupů (graf)



Příloha č. 10: Podíl interaktivních vizuálních výstupů (graf)



Příloha č. 11: Mapa MHD (obrázek)

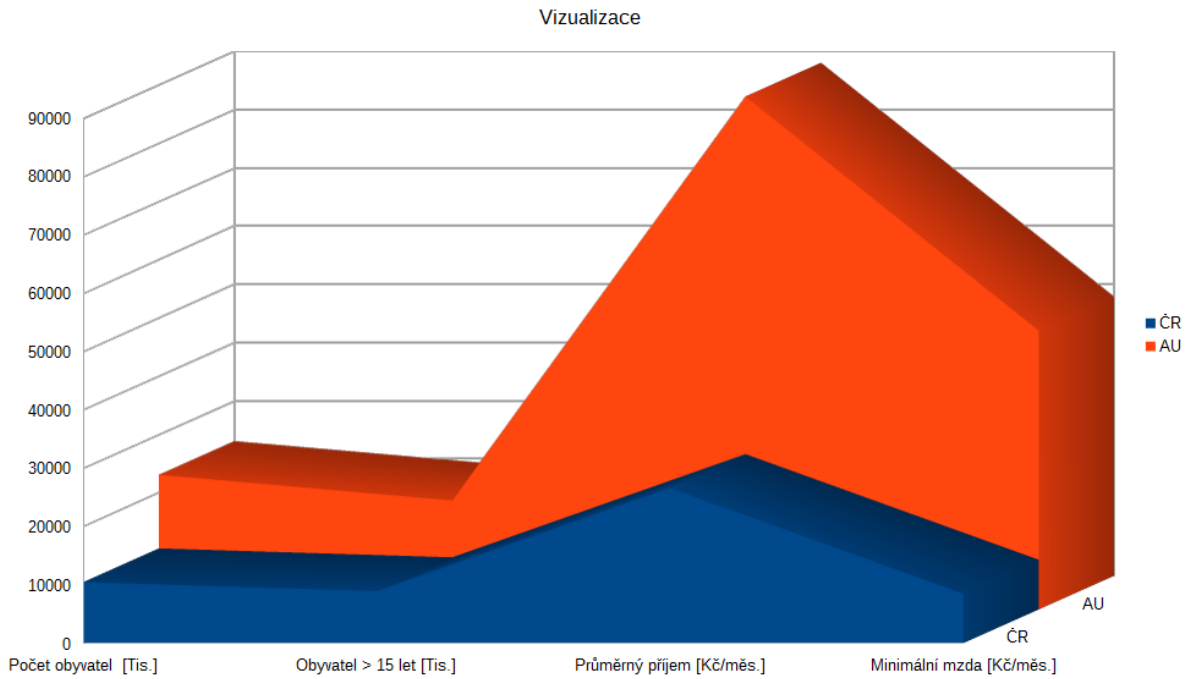


Zdrojem dat je Dopravní podnik hl. m. Prahy. Databáze neobsahuje spojení příměstskými vlaky (linky S).

Příloha č. 12: Srovnání podmínek v České republice a v Austrálii (tabulka)

Srovnání podmínek			
Rok 2013	**Česká Republika**	**Austrálie**	**Poměr AU:ČR|**
Počet obyvatel celkem [tis.]	**10.511**	**23.126**	**2,2**
Počet obyvatel nad 15 let [%]	**85%**	**81%**	
Počet obyvatel nad 15 let [tis]	**8.934**	**18.732**	**2,1**
Průměrný měsíční příjem [Kč]	**26.637**	**88.073**	**3,3**
Minimální měsíční mzda [Kč]	**8.500**	**47.877**	**5,6**

Příloha č. 13: Srovnání podmínek v České republice a v Austrálii (tabulka)



Příloha č. 14: Porovnání postavení datové žurnalistiky v ČR a v AU (graf)

Porovnání postavení datové žurnalistiky v ČR a Austrálii			
Země	**Česká Republika**	**Austrálie**	
Týmy	**Datablog IHNED.cz / DATA Český rozhlas**	**The Age**	**The Guardian AU**
Reference na rozhovory	Jan Cibulka, Marcel Šulek, Petr Kočí, (Filip Rožánek).	Craig Butt, Marc Moncrief.	Nick Evershed.
Složení týmu	Čtyři (Tři novináři, jeden programátor).	Dva novináři.	Jeden novinář.
Vlastnictví	Economia (privátní); od 10/2014 Český rozhlas (veřejnoprávní).	Fairfax Media (privátní).	Guardian Media Group (privátní).
Přínos pro média (subjektivně)	1. Posílení renomé. 2. Inovace postupů.	1. Delší čas strávený u interaktivních aplikací (důležité pro inzerenty). 2. Posílení důvěry. 3. Zvýšení konkurence-schopnosti (trend). 4. Prestiž.	1. Posílení důvěry. 2. Navazující tradice The Guardian.
Největší omezení (subjektivně)	1. Chybějící znalost tématu. 2. Nekoncepční politika státu a z toho plynoucí rozdílné údaje a formáty dat ve veřejných rejstřících. 3. Nedostatečné dovednosti.	1. Chybějící znalost tématu. 2. Získávání dat. 3. Nedostatečné dovednosti.	1. Chybějící znalost tématu. 2. Získání dat. 3. Nedostatečné dovednosti.
Pohled žurnalistů na situaci v zemi (subjektivně).	Málo subjektů se zájmem o datovou žurnalistiku, malá čtenářská obec, malý objem fin. prostředků.	Malé týmy bez specializací, publikum se fragmentuje.	Malé týmy. Vzorem dobrého modelu je VOX, The Upshot.
Pohled žurnalistů na budoucnost (subjektivně).	Stane se součástí běžné novinářské praxe.	Stane se součástí běžné novinářské praxe.	Datová žurnalistika zůstane oddělená jako samostatný žánr

Příloha č. 15: Srovnání týmů v České republice a v Austrálii (tabulka)

Srovnání týmů v ČR a Austrálii			
	ČR	**Austrálie**	
Týmy	**Datablog IHNED.cz/ DATA Český rozhlas**	**The Age**	**The Guardian AU**
Členové týmu	Tři datoví novináři a jeden programátor.	Dva datoví novináři.	Jeden datový novinář.
Původní vzdělání členů týmu	Žurnalistika a technický obor (ČVUT).	Žurnalistika a věda (chemie a fyzika).	Věda (biologie).
Trénink	Praxe a samostudium. Konference.		
Způsob práce	Všichni píší a všichni programují. Společné projekty, kde si práci rozdělují, i projekty, na kterých pracuje každý sám. Krátkodobé i dlouhodobé projekty		
Umístění v redakci	Tým sedí spolu v kanceláři.	Open space, nikdo nemá svůj stůl. Posadí se podle projektu spolu s ostatními novináři	V jedné místnosti s ostatními novináři.
Spolupráce	Ad hoc spolupráce s kolegy, redaktory z aparátu rozhlasu, odborníky.	Ad hoc spolupráce s žurnalisty a grafickým týmem novin. Odborníky a programátory	Ad hoc spolupráce s žurnalisty a grafikem novin. Odborníky a programátory.
Články	Témata, počet článků a formu prezentace určuje tým a vedení redakce.		
Zdroje dat	Lokální databáze (uzavřený objem dat za vybrané období), strojové zpracování vzdálených dat ve strukturovaném formátu (rejstříky vládních úřadů a organizací, škol, nemocnic, policie i soukromých společností) atd. (Rozdíl mezi ČR a Austrálií je v zákonném svobodném přístupu k informacím a míry otevřenosti dostupnosti informací.)		
Nástroje	Převážně OpenSource, případně i placený SW (výčet uveden v tab. níže).		

Příloha č. 16 Softwarové nástroje pro zpracování dat (tabulka)

SW nástroje pro zpracování dat		
Účel	**Open Source**	**Placené**
Pro získávání dat	OpenRefine Adobe Acrobat (pro PDF)	Outwit Hub
Pro analýzu	LibreOffice (calc)	MS Excel
Pro čištění dat	LibreOffice (calc), Jazyk R	MS Excel
Pro zpracování dat	Jazyk R, Python	MS Azure (placená služba - |cloud)
Pro vizualizaci	Ghepi (grafy) DataWrapper Tableau Public TilleMil (mapy) GeoDu (geostatistika) Google Fusion Tables (grafy mapy, atd) - experimentální verze	ArcGIS – geostatistika CartoDB (mapy)

Příloha č. 17: Osobní rozhovor s Nickem Evershedem

An interview with Nick Evershed, conducted in June 2014 in Sydney, Australia.

Nick Evershed is a data journalist for Guardian Australia. He worked briefly as a scientist then he switched careers and now specialises in data-based investigations and multimedia. He studied Science at the University of Sydney. He lives in Sydney.[55]

1. **What is data journalism?**

Data journalism is numbers and statistics and sources. Generally instead of speaking to someone you do your research and analysis but you will often speak to someone as well.

2. **How did you get started as a data journalist?**

I am originally a scientist. I have always been interested in new forms of visual journalism and programming. When The Guardian AUstralia was establishing its branch in Sydney and was recruiting new stuff, I already had my own portfolio with visualizations to present. I also knew programming, math, and statistics thanks to my scientific background. From the beginning I was able to work on projects on my own.

3. **What did you study?**

I studied science at the University of Sydney. Not journalism but Molecular biology. I taught myself programming and other things.

4. **How many people are in the data team and what are their roles?**

It is only me. I do everything by myself.

[55] LinkedIn: Nick Evershed Profile [online]. [Cit. 2014-07-18]. Available from: http://www.linkedin.com/profile/view?

5. **What are the skills someone needs to practice data journalism?**

You must be capable to write and ask questions as all journalists do. Plus you need to know math, statistics, programming, cleaning, analysing data and visualizing.

6. **Where and how do you keep your skills updated?**

I learn by myself from project to project. Also I learn from programmers or graphic designers that I get to work with.

7. **Do you cooperate with other data journalists from The Guardian in the United Kingdom?**

When I started with data journalism in The Guardian AUstralia I was often in touch with the office based in the UK. Nowadays I am rarely in contact with other data journalists from the UK newsroom but I quite frequently ask for help developing stuff from there because I usually don´t have any problems with writing but I may have some issues with programming.

8. **Do you cooperate with other journalists in the newsroom here?**

Our office is generally very small. We all sit in the same room. However, I am the only data journalist in the Australian office. There are other journalists, one editor, one graphic designer but no developer. We might hire one on some occasions.

We collaborate in the newsroom. I may discuss some topics with other journalists in the office to obtain deeper understanding of certain topics. As I said, the office is generally very small, which means there is not much team work like there would be in a larger team. To give you an idea, only 10 people were in the Australian Guardian team when it started last year.

9. How long does it usually take to create a story?

I need half a day at minimum to create a story but I work on many stories at one time. I also work on several long term projects.

10. How many stories do you create in a week? Are there any requirements on how many stories must be created in a certain period of time?

I usually create about 3 stories a week. There is no pressure regarding the number of topics I have to write. I can write as many articles as I want.

11. Who decides on the topics?

As regards topics, I come up with an idea then I ask our editor whether I can write about it or not. Or some topics may be proposed to me.

12. Is there much public interest in these articles?

I think readers are interested in data journalism. The outcomes are always journalistic articles with more data and numbers, which may be supported by graphics. A good example of an interesting data journalistic article is "Australian political donations: search the data". The database shows all declared donations received by Australian political parties up to the 2012 -2013 financial year. Readers could search the data or sort it by a simple click.

13. What's the one thing people always get wrong when they talk about data journalism?

People often think that data journalism is only about programming and reading numbers.

14. What are the biggest challenges regarding the profession?

The biggest challenge I cope with is my lack of knowledge or experience. I must be asking experts about topics. It is the same as in traditional journalism. You can create a data set, analyse data but you must talk to someone about it to understand the problem

and interpret it correctly afterwards. You do some science, programming but it is not only about that. It is a complex process.

Also getting the data is most of the time quite challenging. You must do some scraping. You might make a Freedom of information request. Or you might try to get in contact with an academic source and ask for access to the information.

Once you have the data you may struggle with the format you get the data in.

15. What is good about data journalism?

The communication is much more two way, more interactive. The audience can feel they are part of the story.

Data journalism brings more transparency, which leads to bigger trust.

16. Does the public contribute to the creation of stories?

Yes, sometimes through crowdsourcing.

17. Should data journalistic skills be taught at schools?

In my opinion schools should teach at least a subject broadening other skills related to the work with data as a part of the journalism studies.

18. What does the future hold for data journalism?

I think data journalism will remain divided from traditional journalism. It is more specialized. In newsrooms there should be data journalists that work with data.

United States are moving forward. They are establishing new media organizations, recently “VOX”, “FiveThirtyEight” and “The Upshot”. I am interested in seeing how they do.

Příloha č. 18: Osobní rozhovor s Craigem Buttem

An interview with Craig Butt, conducted on the 2nd of September 2014 at The Age building in Melbourne, Australia.

Mr Craig Butt is a Data Journalist in The Age's data team and Data Journalism Lecturer at the University of Melbourne's "Centre for Advancing Journalism"**.** He became The Age's first data journalism trainee in 2011. After completing the traineeship, he was taken on full-time as a data journalist in The Age's data team. As a Lecturer he taught a course in data journalism in January 2013, which was the first of its kind taught in Australia, covering data literacy, maths for journalists, statistics, data cleaning, data analysis and data visualization. He was a Co-founder of "News Hit", a student Journalism website and a Journalist at "Dangerous Ground", an investigative journalism website. He studied Journalism at the Monash University in Melbourne. He lives in Melbourne.[56]

1. **What made you to start with DJ?**

I had an experience with online journalism working in "News Hits Website", where I was already creating interactive maps. Also, I was an Investigative Journalist at "The Dangerous Ground". When The Age advertised that they were seeking for new journalists I simply signed up and showed them my project, which I did for "EPA" in the beginning of 2011. I was working on this project in my free time, learning new skills by myself while accomplishing this.

[56] LinkedIn: Craig Butt Profile [online]. [Cit. 2014-09-18]. Available from: http://www.linkedin.com/profile/view?

2. **What are the skills someone needs to practice data journalism?**

It is basic data journalistic skills such as; math for journalists, statistics, data cleaning, data analysing and data visualisation. But traditional journalistic skills like writing and approaching people to comment remains fundamental.

For bigger projects someone might use programming languages such as “Python”, “JavaScript” or “Ruby”. For example, my colleague Marc Moncrief uses Python.

3. **How many people are in the data team and what are their roles?**

It is just me and my colleague Marc Moncrieff. My role involves data scraping, building interactives for online and iPad, working with the graphics team on for print infographics, analysing data sets, writing blog posts. I and my colleague Marc work on stories independently, together and with the publication's reporters. Marc Moncrief has been working as a Journalist for The Age for about 10 years. As I have mentioned, he uses python a lot.

4. **What is your academic background?**

I previously studied science – chemistry and physics but I have always been interested in journalism and so I changed my major to journalism.

5. **What does the organization of work look like within the data team? Do you cooperate with other journalists in the newsroom?**

In The Age we have an open office space, based on the 2nd floor. We have no fixed desks. Every day, we pick a different desk. We sit together with people we need to work with. For instance, other journalists in the newsroom might ask me for help regarding some data work. But I might need help from other journalists who can explain the context of a story.

A story using traffic volumes data called “Threat to road link funding” is a good example of a story where it is necessary to be provided with a context. The traffic data were covering 10 years period. We needed the knowledge of the background to give the data a meaning. Here the journalist’s knowledge of the story was important together with numbers for data journalistic work.

Because, we are the only two data journalists in the newsroom we do most of the work ourselves but we have graphic artists in the newsroom. I usually make a sketch and they create a visualisation.

6. **Where and how do you keep your skills updated?**

I am learning with every story I am working on. I try new things and always challenge myself.

7. **Is it difficult to meet the demand for people with data journalistic skills?**

I think journalists should be data literate but not many are. In the future it will become a necessity. At present it might be difficult to find graduate journalists with such skills.

8. **Are schools training students properly?**

I am a Data Journalism Lecturer at the University of Melbourne's Centre for Advancing Journalism**.** The course was the 1st course dealing with data journalism in Australia. It was a 4-day course covering the basic data journalistic skills; such as maths for journalists, statistics, data cleaning, data analysis and data visualisation. The course teaches the basics but the university may offer also advanced option of the course in the future.

9. **Where do you get the data?**

We ask the government departments and other public bodies and committees, local councils, schools, hospitals, companies, the police etc.

10. **Are there any difficulties regarding the format of supplied data or requesting data?**

Sometimes we obtain the information in suitable formats, other times; we must scrap the data like from the Victorian Police website, where they share crime statistics.

If we have any problems getting the data, we often make a "Freedom of Information request". In other words, we contact an organization in writing to make a Freedom of Information request by email or letter, where we describe what information we would like to have. We may also request a specific format.

11. **How long does it usually take to create a story?**

There are stories that must be finished by the next day. The minimum time required for these is half a day. There are also long term stories that take a month or longer to create.

An example of a long term story is a featured story that has about 2500 words- "Missing money mystery: federal government seizes inactive bank accounts". It was derived from data; otherwise I wouldn't be able to see the story. We had to conduct interviews and do some farther research. It was more like traditional journalism work plus detective work, than just analysing the data.
Sometimes when data are released it is quick. We may only need to update some information, which can be done in minutes. An example of a story that has been updated regularly is about High school performance. I spent ages on it at first but to update the data takes literally minutes.

12. **How many stories do you create in a week? Are there any requirements on how many stories must be created in a certain period of time?**

There aren´t specific requirements on a number of stories. The number varies from week to week.

13. **Do you have to ask the editor whether you can write about a certain topic? If so, do you always put the topic across?**

There is an editor I report to at The Age. The Age has a society, state and investigations editor but I don't report to them. Instead, I report to the news director, who has some

idea of what is happening in each of those areas. Often I won't go to my editor with a story as such, rather a data set that I would like to spend some time analysing or scraping. Often I have an idea of potential stories that could come out of the data set but I won't know what the data will tell me until I've spent some time looking at it. I find this is a good way to work out whether I should be spending time on a topic.

There have been times when my editor has told me not to look into a data set because he didn't believe any interesting stories would come out of it, while other times I have spent a while analysing data only to find there isn't really much of a story in the figures after all. When this happens we will often post our analysis on "The Crunch" blog in case anyone else notices anything interesting and so that our hard work doesn't all go to waste!

14. Where do you get the inspiration for the stories?

We inspire each other in the newsroom. Somebody can point out that this is an interesting topic with much supporting data. Or I might come up with a dataset that is hiding one or more stories. Or there might be some data released that the editorial office wants us to write about. Or some unexpected event.

The newsroom sometimes determines what topic we must work on. But most of the time we are free to decide. It might also be the case that somebody from the public points out some data that we might be interested in.

15. What are the most common tools applied to DJ at the Age? Or what tools could you not live without?

For scraping data non-programmatically I like to use "Outwit Hub", I use "Excel" and database software for most of my analysis and I like to turn around quick data visualizations using "Tableau Public", "Google Fusion Tables", "CartoDB" and "DataWrapper".

16. Why is data journalism important in the context of the contemporary digital environment for information?

Data journalism is important because nowadays more and more data are available from digital sources. On one hand that is why journalists must be data literate. On the other hand data literacy is not enough because journalists still need to have traditional journalistic skills - the ability to look at data critically and formulate questions.

Data journalism helps the public to make decisions. It explains the data in a way that can be easily understood. “Vox” is an example of a news site that includes explanatory information.

Data journalism is a type of journalism that brings stories into focus. It can provide insights into some problematic areas and lead to a public discussion. For example, we were working on family violence data analysis in Australia and release number of stories regarding the topic. We pointed out the importance of this topic by presenting relevant data. We encouraged people to be engaged and awareness campaigns have increased.

Also, frequently we discover stories that would otherwise be hidden if not for data journalism.

Often, a data journalistic story includes the original data, which is available to the public for further research. This is important because it builds trust among readers.

17. Is data journalism within The Age published only online?

No. Our stories appear in the printed version as well. We try to publish as many stories as possible in the newspaper.

18. Is the public interested in reading data journalistic stories?

In my opinion people are interested in data journalistic stories. But the stories we create must not be only interesting but also simple and easy to understand. As I have previously mentioned, data journalism is not only about numbers but it is about a story. Data statistics can often be impersonal but these stories are about people based on data

not only about data. We must look at numbers in context and take into consideration population changes, inflation etc.

19. In which topics are people most interested?

Personally, I am not aware of any particular stories that people demand. But there are some expected stories like school performance releases, some annual data, data from hospitals, stories regarding budget etc.

20. Does the public contribute to the creation of stories?

Under a story we put our email address for people's ideas, comments that they want to share with us. We are happy if readers write to us regarding additional information or some interesting suggestions for a story.

An example of a project where people were contributing to the creation of a story is a crowdsourced bicycle accidents map "Cycle data riding high on reader feedback". People were entering marks on a map; they could also write comments and upload a picture. About 600 people participated. It was an amazing experience. And each entry was actually creating a new story. It is an example of many stories hidden in one dataset. The reason why we asked people to help us was that it is easy to get data on deaths but difficult to find data regarding accidents that were not fatal. The story also gave an attention to those who were injured.

21. What data journalism project are you the most proud of working on or creating?

I am proud of cycling map I was talking about, the featured story and stories about family violence.

22. How do you see data journalism in Australia, globally?

In Australia the trend tends to small teams that do a bit of everything.

Globally there are small and big teams. As an example of a big team is "The Upshot NYT", which is a New York Times website that provides analysis and data

visualizations. Their team is big and people within the team are specialized. Everyone concentrates on his task within a certain field.

23. What data journalism project created by someone else do you most admire?

I liked “The Global Mail”, an independent news site, which doesn’t exist anymore. They used to work on long term projects.

I really liked The Guardian’s work about riot incidents in London in 2011 where data on deprivation and the riot incidents happening in that time were mashed up together to find a possible correlation between them. It was a complex story including information of the most affected areas, providing the explanation of appearing problems, optionally how it looked after some time etc.

24. What does the future hold for data journalism?

I believe journalists will need to have the knowledge of basic data journalistic skills. The audience tends to be more fragmented, which leads to a bigger fragmentation of media. There are new news sites like “Vox” or “FiveThirtyEight” that do purely data journalism.

25. What are the biggest challenges for the data team in the newsroom?

Personally, I think the biggest challenge we face is that data is not a round in the same way that transport or health is, so the fact we can dip in and out of these areas is quite daunting. It is also a bit daunting when we don't necessarily have the subject knowledge or contacts within an area to really get to grips with a story. That's why it is important that we work with other reporters.

The other issue is that data journalism encompasses such a huge range of skills and the entire profession is moving so fast that keeping up to date is important. There are new tools, visualisation packages and software appearing all the time so keeping up with them can be hard. However, it is impossible to learn all the skills or follow everything so we have to be strategic in what we choose to do training in.

Příloha č. 19: Osobní rozhovor s Marcem Moncriefem

An interview conducted with Marc Moncrief at the workshop called "Best Practice in Data Journalism" in Melbourne on the 30th of September.

Marc Moncrieff is a Data Journalist within The Age. He has been working in this newspaper for more than 10 years. He studied Journalism and Public Policy at Murdoch University v Perth. He lives in Melbourne.[57]

1. **What is your definition of data journalism?**

Using technologies old and new to tell stories that can be hidden by the deluge of data we live with.

2. **Why should data journalism be attractive for the media outlets?**

The biggest sell for data journalism is engagement. Readers spend long time playing with the tool, more if it looks cool, is interactive or shows something interesting. Engagement is important metric for websites – how much time people spend on the website is important information for advertisers. The more time readers spend reading a story the more interesting it gets for advertisers. Also, to show a dataset with the story what we often do builds credibility. When you make a statement you need to have something backing it.

The last but not least, the newsrooms are not competing only in their area but with online news the competition becomes global. With new tools available come new ways of journalistic practice. Basically anyone, not only newsrooms, can do interactives on the web because the tools are so easy to access. Data journalism, interactives are trendy and if we didn´t follow the trend we would lose the track.

If the rest of the world is innovating and changing the way people expect to receive information you need to not just keep up with that but be ahead of that otherwise you won´t be successful.

[57] LinkedIn: Marc Moncrieff Profile [online]. [Cit.2014-11-12]. Available from: http://www.linkedin.com/profile/view?

Příloha č. 20: Osobní rozhovor s Margaret Simons

An interview with Margaret Simons, conducted on the 4th of September 2014 at the University of Melbourne in Melbourne, Australia

Margaret Simons is an award-winning freelance journalist and author. She is also the Director of the Centre for Advancing Journalism and coordinator of the Masters in Journalism at the University of Melbourne. In 2010 she was a co-founder of the news site "YouComm News", which is an Australian non-profit organisation that brings journalists, publishers and the public together for crowdsourcing of ideas and resources for high quality, community driven journalism. Simons has a doctorate from the University of Technology, Sydney. She lives in Melbourne.[58]

1. **Why is data journalism important in the context of the contemporary digital environment for information?**

More and more information is available online. Plus we have never had better tools for gathering and publishing information. Data journalism is responding to the changes. Nowadays, the journalistic practice involves work with data, posting raw data for the audience. The Inverted pyramid paragon is no longer relevant. Nevertheless, it raises the question of the role of a gatekeeper.

There are great threats and possibilities. And these changes are not happening only in media but also other businesses.

2. **What are the main advantages and disadvantages of data journalism?**

At the moment mainstream media organizations are in charge. Thanks to the Internet making distribution cheap and nearly universal, the media market of the future will be about the rise of the niche, or small but more deeply interested audience. A good model of an organization that focuses on data journalism is the American "Vox". It is a niche media model with expertise in Big Data.

[58] Margaret Simons [online]. [Cit. 2014-09-19]. Available from: http://margaretsimons.com.au/?page_id=2

3. **What is the relationship between data journalism and traditional journalism?**

Data journalism brings a new stream of journalistic practice. But the core remains the same. Journalists are still dealing with people, not only with data. And it is crucial to write in context.

4. **Are journalists data literate?**

No, unfortunately they are not data literate. A journalist rarely knows how to program or even how to work with excel.

5. **What skills should data journalists have?**

Besides skills that traditional journalists have a data journalist needs to efficiently work with big amount of data and be able to present them in a simple way that everyone can understand.

6. **Is the education system responding to the changes in journalism?**

The University of Melbourne opened a course in data journalism in January 2013 and it was the first and I believe still the only one of its kind taught in Australia.

7. **Can one course on data journalism cover all the basics data journalists should know?**

No. It is a good start but we have bigger ambitions, which are long term. I would be happy if the core subjects included data content and if we had some elective subjects covering all the skills, like inscription.

8. **Do readers contribute in creating stories?**

I think there is a potential for collaboration because data journalism is much more two way. Also in the time of the World Wide Web the gap between composing content and distributing it has radically narrowed.

9. **What are the challenges and problems that new-media entrepreneurs are facing?**

There are more readers of newspaper content than ever before. However, the mainstream media's decline is about business models and not the lack of appetite for news. The business model of mainstream media has depended on mass audiences in a single place and selling their attention to advertisers. Lately, the classified advert has been replaced by more efficient online advertising sites, many of which are free to use. And at the same time, the audience is fragmenting. Advertising will remain as a source but it must be more targeted. The future predictions are that media outlets will offer in-depth materials for selected audiences, which is relevant to Data Journalism. Targeted audience advertisers will pay the premium to reach the audience. It will be a different kind of advertising.

Also, information will be more a commodity as it is in other businesses like health care etc. where data has sell value. Another income will be from subscriptions. There can be higher subscriptions for high quality information. Foundations are another way of earning money. Media outlets are also holding public events, providing training and setting content as a way of influencing communication interest. For example; "The New Daily" that provides national news free of charge. It is a non-profit organization fund to draw an audience and educate them. Another example is "The Conversation" - a media outlet, which is funded by universities and finds connection between news and universities. Its main purpose is to influence the public life. With these two examples and many others, providing news is only part of the intention.

And this all raises the question. Is it independent?

10. **How do you see data journalism in Australia, globally?**

These trends, taken together, make it likely that the most interesting things in news media in the future will not be the big industrial complexes of mass media, but rather many smaller, specialist media enterprises. So the production of news-media content is likely to become a cottage industry, as well as something done by Big Media. In Australia and worldwide there is a new "cottage industry" media site almost every day. Many don't last, but all learn lessons in their brief lives.

11. What are the most significant examples of start-up news companies?

I must mention Craig Thomler who is a key advocate of open data and is setting up a website dealing with data. He has founded or supported several start-ups. And “The Open Knowledge Foundation Network in Australia”, which mainly provides information about government and focuses on sharing data and information with open communities in Australia.

Příloha č. 21: Osobní rozhovor s Janem Cibulkou

Rozhovor s Janem Cibulkou, datovým novinářem v Českém rozhlase, z 23. 10. 2014 v budově Českého rozhlasu v Praze.

1. **Co to je datová žurnalistika?**

Datová žurnalistika je primárně hledání a zpracovávání témat na základě tvrdých dat. My především pracujeme s daty státní správy a se soukromým sektorem pouze výjimečně. Datová žurnalistika se pro nás stává hodnotícím nástrojem. Je to možnost, jak pokrývat témata, o kterých máme nějaký poznatek a z toho poznatku zjistit a ověřit, zda jde o globální trend. Někdo řekne nějaký výrok na nějakém místě, ale informace k tomu získá jen z jednoho zdroje. Má jednu zkušenost z jednoho místa, nebo mu to někdo řekl, a my se snažíme ověřit, jestli se to něco dá říct i o jiných místech, jestli jde o globální trend, nebo jestli se to děje pouze na tomto jednom místě. My jsme to díky dostupným datům schopni ověřit a dát do souvislostí. Nebo typickým příkladem je interpretace výsledků voleb. Víme, že někde vyhrála komunistická strana a my se snažíme zjistit, proč to tak je. V jaké sociální třídě tam voliči žijí, jaké mají příjmy, v jaké jsou věkové kategorii atd. Můžeme dát do kontextu nějaký lokální poznatek. Díky spolupráci s aparátem Českého rozhlasu také pokrýváme lokální zpravodajství.

2. **Jaký je vztah mezi datovou žurnalistikou a tradiční žurnalistikou?**

Já bych datovou žurnalistiku od tradiční žurnalistiky neoddělоval. Datová žurnalistika je v podstatě jen další nástroj tradiční žurnalistiky. Nejde o dvě oddělené věci, ač v redakcích to tak funguje, protože obzvlášť české redakce jsou už zajeté v nějakém rytmu a nejsou přístupni změnám. Ve chvíli, kdy žurnalistům někdo řekne, pojďme se podívat na nějaké datasety, většinou tam je nedůvěra a spolupráce delší dobu nabíhá. Ale obecně by podle mého názoru měl každý novinář umět zacházet s Excelem a sečíst si nějaká čísla. Naším cílem je co nejvíce datovou žurnalistiku integrovat a ne oddělovat jako samostatný žánr.

3. **Jaká je organizace práce ve vašem čtyřčlenném týmu?**

Všichni se snaží programovat a všichni se snaží psát. Máme každý témata, kterými se zabýváme jak dlouhodobě, tak krátkodobě. Máme společné projekty, které si rozdělujeme mezi sebou, a pak jsou věci, na kterých pracuje každý vyloženě sám. Většinou záleží, jestli jde o nějaké programování, především vizualizace na web, což většinou skončí u Marcela. Pokud jde o mapy, tak to skončí u mě. Ale často spolu spolupracujeme. Zatímco jeden pracuje na vizualizaci, tak ten druhý například ověřuje správnost dat a tak dále.

4. **Stává se, že si s něčím nevíte rady a musíte požádat o externí pomoc?**

My především spolupracujeme s aparátem rozhlasu. Rozhlasoví redaktoři za námi přijdou se slovy, k tomuhle tématu by bylo vhodné mít nějaká data a pak už sami jedou na místo získat více informací. My také vyrážíme do terénu, ale jen výjimečně. Takhle by to podle mého mělo ideálně vypadat. Výhodou spolupráce s rozhlasem je, že má regionální síť, což v Hospodářských novinách není.

5. **To je důvod, proč jste odešli z Ekonomie do Českého rozhlasu?**

Je to jeden z důvodů.

6. **Kdo rozhoduje o tématech článků?**

Některá témata navrhujeme my, některá redakce. Máme týdenní porady, kde sdílíme nápady. Redakce přijde s požadavky, my řekneme, zda je to reálné, jak to může vypadat a za jak dlouho se to může vytvořit. Tady i v Hospodářských novinách se držíme témat, o kterých chceme psát. Máme určitou svobodu.

7. **Jsou témata, o která čtenáři stojí, kterých se dožadují?**

Držíme se aktuálního zpravodajského dění.

8. **Podílejí se čtenáři na tvorbě článků?**

Nemáme s tím žádné valné zkušenosti. Médium musí mít velkou podporu, to znamená, že potřebujeme hodně lidí, aby výsledek měl nějakou relevanci. Také to musí být téma,

na kterém se lidé chtějí podílet. V zahraničí ale určitě existují inspirativní příklady, jako je práce The Guardian.

9. Je na Vás vyvíjený nějaký tlak, co se počtu článků týče?

Hospodářské noviny jsme opustili v době, kdy se otevíralo téma, kolik článků se vyplatí za jaký čas. V Českém rozhlase se to neprobírá, protože rozhlas žije z koncesionářských poplatků. Jsme tady jako externí pracovníci a jsme odměňováni od projektu, to znamená, že píšeme, kolik článků chceme, ale nějaký počet napsat musíme, abychom měli nějaký plat.

10. Jak dlouho trvá tvorba článků?

Nejkratší doba napsání článku může trvat 3 hodiny, ale jiný článek se může vytvářet třeba i půl roku.

11. Jaký je Váš a Vašich kolegů background?

Petr Kočí je v novinách dlouho. Jan Boček také působil v novinách. Já jsem také dělal v novinách, pak v reklamní agentuře, pak jsem dělal reklamního analytika, pak v Hospodářských novinách a nyní tady. Marcel začal s novinařinou až v Hospodářských novinách.

12. Co vás vedlo k tomu stát se datovými žurnalisty?

Petr Kočí s Honzou Bočkem to sledovali jako trend v zahraničí a napadlo je, že by to mohli dělat v Hospodářských novinách. Já jsem se o to nikdy jako o trend nezajímal, ale začal jsem s tím pracovat, protože mě zajímá kriminalita, takže jsem začal zpracovávat policejní kriminální data. S Marcelem se známe z reklamky, kde jsme spolu pracovali, kde programoval, a my jsme v Hospodářských novinách tehdy potřebovali programátora.

13. Kde a jak jste získal dovednosti potřebné pro práci datového žurnalisty?

Programoval jsem od střední školy. V novinách jsem se vypsal. Celý čas, co jsme s týmem pracovali v Hospodářských novinách, jsme se u toho i učili. Ale žádný

zaměřený trénink jsme neměli, maximálně kolegové se zúčastnili nějaké přednášky blízké oboru. Já s Marcelem jsme si znalosti webových technologií přinesly z reklamky, kde jsme spolu pracovali. Další nástroj, který používáme a který jsem se naučil za pochodu vyloženě čtením judikátu a konzultací s lidmi, je „Zákon o svobodném přístupu k informacím“, samozřejmě v kombinaci se správním řádem a jinou legislativou. Když pominu seminář „Právo na informace“ vyučovaný Františkem Korbelem, spoluautorem zákona, na Právnické fakultě, tak o jiném místě, kde se to probírá, nevím.

14. Jak často používáte zákon o svobodném přístupu k informacím?

Používáme ho poměrně často a celkem úspěšně. Soudních sporů vedeme minimum. V současné době vedeme dva, což považuji za úspěch. Samozřejmě některá témata nejsou tak nosná, abychom šli do sporu. Většinou se je snažíme použít až ve chvíli, kdy se nemůžeme s úřady dohodnout. Ale ve většině případů, jakmile novinář chce nějaké informace v jiném formátu, než je obvyklé, automatická odpověď úřadů je ne.

15. Jak dlouho trvá celý proces získání informací, pokud použijete zákon 106 o svobodném přístupu k informacím?

Noviny pracují na týdenním cyklu. Základní termín na první kolo žádosti je 15 dní a úřady to většinou vyřizují až k poslednímu dni, protože to pro ně často není prioritní agenda. Takže pro noviny to není úplně použitelný nástroj. Ve chvíli, kdy jde úřad proti, pak tam je odvolání, a to je dalších 15 dní plus dalších 15 dní na rozhodnutí, tedy měsíc. A tam to může mít zase několik konců. Ve chvíli, kdy nadřízený úřad svolí, stejně to většinou pošle zpět k novému řízení, takže dalších 15 dní. A to se ještě může stát, že podřízený úřad nám to i tak nebude chtít dát, použije nové argumenty a kolečko se může opakovat. To už potom naskakuje do měsíců. Ve chvíli, kdy nadřízený úřad potvrdí stanovisko úřadu pod ním a my i tak s tím nesouhlasíme, tak správní žaloba v České republice trvá okolo 2 let, než dojde k jejímu vyřízení. V takovém případě se téma odloží, nebo se snažíme data získat jinou cestou.

16. V jakém formátu data dostáváte? Říkáte si o konkrétní formát, a pokud ano, s jakou úspěšností jej dostanete?

Když jsme začínali a neuměli jsme v tom chodit, dostávali jsme data v méně použitelných formátech, jako je pdf. Ale dnes si bez problému vyžadujeme formát, který chceme, a nejlepším formátem, aby to bylo vhodné pro nás i pro druhou stranu, je excel. Protože s excelem úředníci pracují denně. Pro nás by samozřejmě bylo lepší získat výjezd z databáze, ale než to vysvětlovat informačnímu oddělení a tiskové mluvčí, je jednodušší si vyžádat informace ve formátu Excel.

17. Jaké nástroje používáte na zpracování dat?

Na tabulková data používáme Excel, OpenOffice, konkrétně „LiberOffice“, „OpenRefine“ nebo „Adobe Acrobat“ na dolování dat z formátu pdf. Z programovacích jazyků používáme na zpracovávání dat programovací jazyk „R“ nebo například „Python“. Ze serverových technologií používáme „Microsoft Azure“, který si platíme. Na mapy používáme například aplikaci „TileMill“. Na grafové věci používáme „Gephi“. Na geostatistiku používáme „GeoDu“. Co se týče geostatistiky, budeme přecházet na technologii „ArcGIS“, která je komerční.

18. Jaká je finanční náročnost práce s daty?

Nástroje, které používáme, jsou většinou zdarma, tedy nenáročné.

19. Je datová žurnalistika pro noviny komerčně atraktivní?

Ve chvíli, kdy noviny budou prodávat výhradně na webu a jejich komerční návratnost se bude měřit pouze počtem prokliků na reklamu, tak to nemá nejmenší šanci. Opačný případ by byl, pokud noviny budou prodávat obsah exkluzivně v rámci předplatného. Samozřejmě otázka je, jestli je v České republice dostatečná předplatitelská základna, v Německu například ano, u nás to je sporné. Nebo to redakce může dotovat jako jakýsi symbol prestiže.

Další obchodní model, který existuje v zahraničí, v České republice jsou k tomu noviny opatrné, je, že některé projekty jsou sponzorované z venku. Timesy si investigativu

nechávají sponzorovat z venku. V České republice existují dva fondy pro podporu investigativní žurnalistiky.

20. Často u aplikací, map nebo jiných vizualizací čtenáři stráví více času. Má to efekt na inzerci?

Absolutně ne. Reklama se prodává za počet kliků. Novinám jde o to, aby čtenář klikl na co nejvíce krátkých článků, ne aby byl u jednoho článku dlouho. Nejvíc reklamy protočí články s bulvárními tituly. Ve chvíli, kdy prodáváte práci jen za reklamní prostor, tak musíte tvořit velké množství výplňové hmoty. Máte několik článků, které vám lidi přitahují, sdílí se na sociálních sítích. Vy ale stejně potřebujete co nejvíc lidí, co přijdou, povodit po webu. Proto jsou pod každým druhým českým článkem diskuze, protože i diskuze je pro čtení reklamy. Diskuze je dlouhá proto, aby to tvořilo dalších 50% stránky podle matrix spinu, což se počítá jako nové page view, tzn. nové načtení reklam. Některé servery v České republice používají obsahové farmy typu „Helena Vondráčková.com“, obsahu nejnižší kvality, který má za cíl lidi přitáhnout a povodit je po webu s co nejvíce zobrazení reklam. Problém je, že page view je stejný jak u dobrého, tak u špatného článku. Dobrý článek jich třeba udělá dvakrát víc, než průměrný článek, na druhou stranu průměrných článků vychrlí novinář za den třeba dva až tři. Ne množství čtenosti, ale množství vydaných průměrných článků je pro noviny stále finančně atraktivnější, než jeden obsáhlý článek, který sice může přinést určitou prestiž, třeba i přitáhne předplatitele, ale to těch kvalitních článků už musí být více.
Prodej obsahu za reklamní prostor se obecně nevyplácí, nemluvě o tom, že ceny za reklamní plochy padají dolů. Nehledě na to, že většina obsahu se prodává přes mediální agentury, takže ještě z těch peněz ukrajuje zprostředkovatel. Zadavateli je v zásadě jedno, kde to bude. Jde o page views a stlačení ceny co nejníže za page view. Samozřejmě některé firmy si hlídají, aby jejich reklama nebyla vedle článku, který o nich špatně referuje. Ale na webu je to povětšinou dosti každému jedno. Často se nakoupí reklamní prostor, třeba na několika portálech najednou v různých balíčcích pro různou cílovou skupinu.

21. A co informace jako komodita?

Datasety, s kterými pracujeme pod otevřenou licencí, veřejně zdarma publikujeme na webu. Výsledky naší práce, zdroje, všechno. My často vyžadujeme, že je to naše dílo,

ale nebráníme komerčnímu použití výsledku. K dílům máme stále plná práva, ale někdo jiný z toho také může vytěžit. V Economii, protože šlo o komerční společnost, jsme svou práci publikovali pod podmínkou nekomerčního použití jinými.

22. Kde se vzděláváte?

Praxí a internetem, možná konference. Vzděláváme se každý ve svých specializacích.

23. O koho byste měl zájem jako posilu týmu - absolventa žurnalistiky nebo programátora?

Raději programátora, který se naučí prezentovat data. I v novinách je lepší, když vám přijde specialista na právo, než novinář, který chce pochopit právní systém. Obecně je lepší, když novinář přijde z nějakého oboru, který je jiný než žurnalistika. Gramotně umíme všichni psát ze školy. Styl psaní se dá naučit. Zvlášť pokud má kvalitního editora, který ho nějak vede. Ale opačným procesem, kdy přijde novinář, který je nevyhraněný, píše o nevyhraněných tématech, o kterých se dočte na BBC, vznikají dezinterpretace.

24. Jaká jsou omezení v datové žurnalistice?

Musíme si dávat pozor, abychom nezabředli do čísel moc a nezačali produkovat pseudovědecké práce. U datové žurnalistiky je problém, že se tváří jako věda, ale není to věda. Věda využívá spoustu ochranných mechanismů, aby nevznikaly hlouposti. Jsou tam principy kritického uvažování, ověřování údajů a jiné. Tendence je v novinách těchto vědeckých mechanismů zanedbávat. Věda má status něčeho ověřeného. Pokud něco prezentujete ve stylu, že napíšete: „*K závěru jsme došli ověřením statistických údajů.*“, lidé mají tendenci tomu věřit mnohem více. V novinách tyto mechanismy chybí, že by se ověřovaly výsledky, opakovaly se výsledky. My vlastně děláme experimenty, ale tím, že publikujeme data pod článek, dáváme možnost ostatním ověřit si výsledek.

25. Stalo se vám, že jste někdy publikovali informace nepřesně a byli jste na to upozorněni čtenáři?

Určitě jsme se dopustili několika nepřesností, na které jsme byli upozorněni. V novinách chybí mechanismus kontroly a alespoň takto jsme otevřeni kontrole ze strany veřejnosti. Lidé nejsou odkázáni na to, že musí věřit autorovi článku, ale mohou si to ověřit podle zdrojových dat.
Obecně téma otevřenosti v metodách a zdrojích dat je kritický. Datová žurnalistika vychází z vědeckých metod. Ve společenských vědách se některé výzkumy nedají opakovat, ale dá se zkontrolovat metodologie. Je férové ukázat zdrojová data u článku, nebo odkázat na článek bokem, který o tématu referuje, nebo kde je popsána metoda. To posiluje důvěryhodnost.

26. Spolupracujete na tématech s odborníky v oboru?

Snažíme se spolupracovat se specialisty. Snažíme se neinterpretovat poznatky jen na základě dat.

27. Setkal jste se s problémy získat informace od akademiků nebo jiných odborníků?

Nikdy se mi nestalo, že mi nějaká vědecká instituce odmítla poskytnout informace. Často i zadarmo, když je požádám.
Problém je to spíš na ministerstvech, třeba Ministerstvo vnitra. Tam chybí veřejná kontrola nebo kontrola odborná. To, co dělá výzkum kvalitní, je veřejná a odborná kontrola, kontrola co nejširší.

28. Ocenil byste, kdyby existovala nějaká centrální databanka, odkud byste mohl informace čerpat?

Bylo by krásné, kdyby něco takového existovalo. Pro tableární data by to bylo možné, ale pak tu jsou 3D data, ND data a úplně jiná data a celé to centralizovat do jednoho bodu, je technicky nepředstavitelné a navíc by to bylo strašně drahé. Šíře dat je obrovská. Co jsou data? Jakákoliv informace. Vytvářet databanku, která pojme vědění celé státní správy, je scestné.

V určitých oblastech to existuje a je to užitečné. Máme „Kompetenční zákon“ v České republice, který stanovuje, čím se jaký úřad zabývá. Prací s touto informací můžeme zjistit, kdo tu informaci má a poptávat ji. Z webových katalogů je to „Caslin“, který sdružuje všechny knihovny v České republice nebo katalogy statistického úřadu. Existuje seznam všech státních útvarů, což je také svým způsobem katalog informací, protože jsem zase schopen zjistit, kdo se zabývá čím a poptávat data.

29. Na který článek, který jste zpracoval nebo jste se na jeho tvorbě podílel, jste nejvíce pyšný?

Nejvíce jsem pyšný na článek referující o fungování městské hromadné dopravy v závislosti na změnách jízdních řádů. V článku jsme ukázali, jakým způsobem je Praha dostupná. Také vizuálně je to zajímavé.

Příloha č. 22: Osobní rozhovor s Marcelem Šulkem

Rozhovor s Marcelem Šulkem, datovým novinářem v Českém rozhlase, z 23. 10. 2014 v budově Českého rozhlasu v Praze.

1. **Jak jste se dostal do týmu datových žurnalistů?**

Dostal jsem se k tomu náhodou. Petr Kočí a Honza Boček si stěžovali tak dlouho, že nemají programátora, až přemluvili tehdejší šéfredaktorku Lucii Tvarůžkovou, aby sehnala programátora, respektive, aby založila oddělení redakčního vývoje. Martin Malý byl přijat jako první a poté mi nabídl práci v Economii. Znali jsme se ze startupu, na kterém jsme pracovali společně a který nevyšel.

Jsem samouk. Nemám v oboru žádné vzdělání. Školy na to nemám, ale myslím, že se to nikde ani neučí.

2. **Váš kolega, Jan Cibulka, mluvil o vašem společném projektu, o městské hromadné dopravě. Jak vypadala práce na tomto projektu?**

Nejsložitější na tom bylo získat od Dopravního podniku data. Hodně se bránili. Věděli jsme, že dávají data Googlu, ale pravděpodobně je nechtěli dát někomu dál, protože považovali za nějaké obchodní tajemství. Nicméně po několika různých urgencích a vyhazovech v jejich řadách, nám nakonec data dali. Získání dat trvalo 3 až 4 měsíce. Potom jsme to zpracovávali cca 3 týdny. Byl to nejdelší projekt, jaký jsme kdy dělali. Vždy je boj, abychom to za co nejkratší dobu co nejefektivněji dali dohromady tak, aby to dávalo smysl.

Co se týče technické stránky, potřebovali jsme spočítat spoje mezi každými pražskými zastávkami. Zastávek je kolem 3000. Přeložili jsme to na „Amazon Cloud Drive“, kde jsme si pronajali asi stovku serverů a nechali to přelouskat vždy přes noc, čímž jsme snížili výpočetní náročnost. Nikdo jsme na to neměli znalosti. Místo toho, abych to dva týdny programoval, jsme to zaplatili. Celkem jsme zaplatili asi 2000 Kč a já si ušetřil 2 týdny práce. Vlastně to takto vyšlo levněji, protože má práce by byla dražší.

Nemělo to však úplně přesné výsledky. Původně jsme chtěli dokázat, jestli někde jsou dojezdové časy posunuty se změnou jízdních řádů, jestli se někde prodloužily nebo zkrátily. Každý nadával, že se doba prodloužila a že se změnou řádů ty časy nesedí. Spočítali jsme dva datasety a v podstatě jsme dokázali, že chyba, kterou my máme v našem výpočtu, je vyšší než změny dojezdu. Takže to, co jsme původně chtěli dokázat, jsme tím vlastně nedokázali. Jen ty nejvýraznější případy se trošku projevily. Nejužitečnější na tom bylo to, že člověk mohl zjistit, jak daleko dojede za jaký čas, a to fungovalo.

3. **Jaké máte technické zázemí v newsroomu?**

Co si najdeme, to máme. Naše výtvory nejsou dlouhodobě udržované. V tom máme výhodu. Nikoho nezajímají včerejší noviny a programy. Takže my, když si něco vymyslíme, nějak to udržíme dva dny v chodu a pak to zahodíme. Hodně používáme cloudové složky. Když jsme byli v Hospodářkých novinách, používali jsme Amazon. Pak nás Microsoft přemluvil na „Azar Cloud“, což je v podstatě to samé, akorát že to poskytuje Microsoft. Člověk si zde může pronajmout x serverů, nechá si něco spočítat a zase to vypne. Platí jen za to, co skutečně použije.

Snažili jsme se vždy držet naše výmysly mimo architekturu publishera, ať už to byla Economie nebo Český rozhlas. Protože nemáme čas testovat, děláme často experimenty. Snažíme se být vždy co nejvíc odděleni, abychom, když u nás něco padne, nepoložili například web Českého rozhlasu.

4. **Vy se soustředíte zejména na vizualizaci dat. Jaké nástroje používáte?**

Na webové vizualizace používám „D3 - data driven document“ a java scriptovou knihovnu. Je to zdarma. Vymysleli to v New York Times. Je to poměrně nepříjemné uživatelsky. Je to nízkoúrovňové, že tomu pořád musí člověk říkat, co to musí udělat. Ale tou nízkou úrovností to na druhou stranu nemá skoro žádné limity, to znamená, že to udělá téměř cokoliv.

Existuje celé spektrum nástrojů. Například klikací nástroje, jak my je nazýváme, jako je „Data wrapper“, „Google spreadsheets“, „Tablo“, „Carto dp“ a jiné. Tady se jen

nasypou data, klikne se, jak se to má zobrazit, a vyjede hotová vizualizace. Také Excel je dobrý příklad.

5. Jaká je nejkratší doba vytvoření vizualizace?

Nejkratší doba je 5 minut, ale vizualizace projektu o MHD trvala asi týden. Týden trvala i sčítačka komunálních voleb.

Příloha č. 23: On-line korespondence s Filipem Rožánkem

Korespondence s Filipem Rožánkem, žurnalistou v Hospodářských novinách, ze dne 16. 11. 2014.

1. **Jaký je vztah mezi tradiční žurnalistikou a datovou žurnalistikou?**

Žádný zvláštní, obojí je prostě žurnalistika. Je to stejné, jako se ptát, jaký je vztah mezi tradiční novinařinou a investigativní novinařinou. Datová žurnalistika je žánr, podskupina novinářské práce. V určité formě existovala už před rozšířením internetu a osobních počítačů, její kořeny můžeme najít v rešerších. Bývalo zvykem, že každá velká mediální organizace měla své rešeršní oddělení, které vyhledávalo informace a souvislosti. Datová žurnalistika v podstatě nedělá nic jiného, jen k tomu má teď databáze, přístupy v reálném čase a automatizované zpracování dat.

2. **Proč by měla být datová žurnalistika pro noviny atraktivní?**

Datová žurnalistika sama o sobě je minoritní věc spíše pro nadšence, kteří chtějí proklikávat grafy, strukturované sady dat a vizualizace. Moderní žurnalistika staví na příběhovém zpracování událostí, pokud tedy mají být výstupy datové žurnalistiky použitelné pro širší čtenářskou obec, musí nad dohledanými daty vzniknout „story“, se kterou se čtenář může identifikovat.

3. **Hospodářské noviny částečně zpoplatnily web. Jak se na tuto změnu díváte?**

Je to první pokus zavést na českém trhu model, který je obvyklý na západ i na východ od našich hranic. Čeští čtenáři mají silný návyk na informace zdarma, ale takový přístup podráží vydavatelům v době klesajících příjmů z inzerce nohy. Je logické, že i vydavatel HN hledá nový zdroj příjmů. Není to navíc „tvrdé“ zpoplatnění, spíš zdvořilé připomenutí, že novinářská práce také něco stojí.

4. **Často u vizualizací nebo interaktivních aplikací čtenáři stráví více času. Má to efekt na inzerci?**

Taková vazba tam podle mého názoru není. Prodejní model inzerce na internetu obecně není postaven na času, který uživatel stráví na stránce, ale na tom, kolika uživatelům je reklamní sdělení doručeno.

5. **Kdo rozhoduje o tématech a počtu článků?**

Model je stejný jako u jakékoliv jiné redakční práce. Autor navrhne téma, které je editory nebo šéfredaktorem buď přijato, nebo doporučeno k dopracování, případně odmítnuto. Počet článků je daný jednoduše kapacitou – personální, technickou aj.

6. **Jaké dovednosti musí mít datový žurnalista? Kde a jak se může učit?**

Datový žurnalista je kříženec programátora, grafika a novináře. Lepší předpoklady tedy mají techničtí nadšenci se stylistickým nadáním, tak jako u ekonomické novinařiny bude mít výhodu absolvent VŠE oproti absolventovi přírodovědných oborů. Datový žurnalista musí mít už během získávání a analýzy dat určitou představu o tom, jak bude vypadat konečná prezentace informací. To se dá jistě časem naučit, ale především to člověk musí mít v sobě, musí mít „hlad“ po nových technologiích, zkoumat je a osvojovat si je. Třeba se sám naučit vytvářet mapy na Google Maps, na to žádný školní obor neexistuje a akademická sféra bude nutně vždy trochu pozadu za prudkým technologickým rozvojem.

7. **Jak kooperujete v redakci?**

Datová žurnalistika se prolíná napříč redakcí. Redaktoři mi průběžně nabízejí různé sumy dat, které mají k dispozici, já hodnotím, jestli existuje nějaký klíč k jejich atraktivnější prezentaci a redakční vývojář zhotoví nejvhodnější technické řešení kompatibilní s redakčním systémem.

8. **Jaká jsou největší omezení v práci datových žurnalistů?**

Čas. Získání, uspořádání, vyčištění, interpretace a prezentace dat zkrátka nějakou dobu trvá. Podklady pro jeden článek mohou vznikat dny i týdny. U veřejných rejstříků je

tradičně největším problémem technická odlišnost, případně uzavřenost formátů a systémů daná nekoncepční IT politikou českého státu.

9. Jak vidíte datovou žurnalistiku u nás a ve světě?

Velké světové trhy mají oproti tomu českému zásadní výhodu v tom, že existuje nejen víc subjektů, které mají o datovou žurnalistiku zájem, ale také víc čtenářů, kteří její výstupy budou číst. A i když vydavatelé po celém světě řeší obdobné problémy se snižováním nákladů a hledáním nových příjmů, západní země si přece jenom mohou stále dovolit investovat do datové žurnalistiky řádově vyšší objem prostředků.

Příloha č. 24: On-line korespondence s Petrem Kočím

Korespondence s Petrem Kočím, datovým novinářem v Českém rozhlase, z 21. 11. 2014.

1. **Co bylo důvodem vašeho odchodu z Hospodářských novin?**

Důvody byly ekonomické. Jeden člen vedení redakce nám řekl, že jsme jako luxusní auto, které si už nemůže dovolit. Naše produkce prý sice posilovala renomé značky, ale to se těžko vyčísluje. Z účetního hlediska jsme byli proděleč né oddělení. Proto se ho management rozhodl zredukovat a zbytek rozpustit do jiných redakcí. Nám bylo líto zahodit všechno, co jsme se za dva roky spolupráce naučili. Zdálo se nám, že v navrhovaném novém pracovním zařazení bychom nemohli naplno využít svou kvalifikaci. Měli jsme pocit, že se svými dovednostmi dobře doplňujeme a bylo nám líto takový tým rozbít. Proto jsme raději odešli všichni. Problém byl v komunikaci na obou stranách: Nám se nepodařilo dostatečně přesvědčivě vysvětlit redakci, k čemu všemu jí můžeme být užiteční.

Noviny jsou pod silným ekonomickým tlakem, novináři často jen s vypětím všech sil plní stanovený objem produkce. Na zkoušení nových postupů, které by umožnily hlubší pochopení problémů, o nichž referují, a nových způsobů komunikace se čtenáři, obecně na jakékoli inovace zbývá v komerčních médiích minimální prostor. Proto jsme se rozhodli nabídnout své služby veřejnoprávnímu Českému rozhlasu.

Příloha č. 25: Osobní rozhovory se studenty žurnalistiky

Interview with Michelle See-Tho - a University of Melbourne student- conducted at the "Best Practice in Data Journalism Workshop" at the University of Melbourne on the 29th of September 2014.
Contact: michelleseetho@gmail.com

1. **What is your major and year of study?**

I'm studying a Master of Journalism. I am a part time student and I am in the 1st year of my studies.

2. **Why are you interested in data journalism?**

I think data journalism is important because not many people do it. And people need to learn the importance of it. If we had discussions like these it will increase awareness.

3. **Why do you think data journalism is important?**

In my opinion nowadays it is important to look at data visually. It makes it easier to understand big amounts of data. Also new technology is emerging. And importantly newspapers are less popular now. More people tend to read the news on the internet. Data journalism is something that uses the medium well.

4. **Are you looking for a job in this area?**

I would like to become a journalist and would like to do both- writing and doing infographics.

5. **As a student of journalism at the University of Melbourne can you consider yourself data literate?**

No I cannot. The university is offering a course on data journalism. I haven't taken the course yet but other students who have said it was really good. But besides this course there are no other subjects specifically about data journalism.
I make infographics at work. I work for a magazine and I learned it there.

6. **Is there anything you wish was in the curriculum of your studies?**

No. I think the skills come with practice.

Interview with Andrew Hazel - a University of Melbourne student- conducted at the "Best Practice in Data Journalism Workshop" at the University of Melbourne on the 29th of September 2014.
Contact: onepolarbear@gmail.com

1. **What is your major and year of study?**

I am studying a Master of Journalism and this is my 1st year.

2. **Why are you interested in data journalism?**

It feels like it is new. It is kind of a cutting edge. Data journalism is tied into the industry. It is a more scientific way of doing journalism. I have a background in natural medicine. But I wanted to go into research. I have a Master Degree in health and social science. And Journalism is my second master.

3. **Why do you think data journalism is important?**

There is so much data and it is important that many people can access the information and understand it. Data journalism is a way to show data in a form people can understand. It can put many important topics into a conversation.

4. **Do you work in the field or do you want to become a journalist?**

I work as a journalist for The Saturday Paper, a weekly newspaper.

5. **As a student of journalism at the University of Melbourne can you consider yourself data literate?**

No. There is so much data out there that I certainly cannot understand. I am good at health data. And considering the data journalistic practice, I do not have the skills needed.

6. **Is there anything you wish was in the curriculum of your studies?**

I would like to learn some programming.

7. **What do you think will happen with data journalism in the future?**

I wouldn't be surprised if the industry stays separated but in the future I think journalists will get better at dealing with data. Everyone will be able to do the basic data journalistic work.

Interview with Daryl Holland- a University of Melbourne student- conducted at the "Best Practice in Data Journalism Workshop" at the University of Melbourne on the 29th of September 2014.
Contact: darylholland@gmail.com

1. **What is your major and year of study?**

My major is Master of Journalism. I am studying part-time and I am in my 2nd year. My background is in science. I have a Master in biology.

2. **Do you want to become a journalist or do you already work in the field?**

Yes. I am already working as a journalist and I would love to be able to master data journalism.

3. **Can you consider yourself- as a student of journalism- data literate?**

Partly yes thanks to my scientific background where I was using excel every day. And also thanks to the Data Journalism Course. It was a 4 day intensive course, which was a good basic introduction to Data Journalism. It was very practical. We learned how to use "Tablo", "Data Wrapper "etc. We were given a fictional dataset and we had to find a story in the data.

I also took the MOOC course, which was very similar to the one I did at the uni.

4. **Can you compare those two courses?**

The one at the university was more intense, more cooperative with other students. But online we were given more time to learn the programs on our own.

5. **Does completing the course Data Journalism give you the confidence and skills needed to become a data journalist?**

It was a good basis. But I think it is more about training on my own. It provided a good basic overview.

6. **Are there any other subjects at the university you would consider as helpful for the profession?**

There is a course called "Advanced Multimedia Course". But besides this I cannot think of any others.

7. **Would you recommend any other related subjects you wish to learn at the university?**

I would appreciate learning some math, statistics and programming.