

UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut komunikačních studií a žurnalistiky

Katedra žurnalistiky

Diplomová práce

2022

Nela Krawiecová

UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut komunikačních studií a žurnalistiky

Katedra žurnalistiky

**Otevřená data a jejich využití v tuzemské datové
žurnalistice**

Diplomová práce

Autor práce: Nela Krawiecová

Studijní program: Žurnalistika

Vedoucí práce: PhDr. Václav Moravec, Ph.D., Ph.D.

Rok obhajoby: 2022

Prohlášení

1. Prohlašuji, že jsem předkládanou práci zpracovala samostatně a použila jen uvedené prameny a literaturu.
2. Prohlašuji, že práce nebyla využita k získání jiného titulu.
3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze dne 1.1.2022

Nela Krawiecová

Bibliografický záznam

KRAWIECOVÁ, Nela. *Otevřená data a jejich využití v tuzemské datové žurnalistice*. Praha, 2022. 131 s. Diplomová práce (Mgr.). Univerzita Karlova, Fakulta sociálních věd, Institut komunikačních studií a žurnalistiky, Katedra žurnalistiky. Vedoucí diplomové práce PhDr. Václav Moravec, Ph.D., Ph.D.

Rozsah práce: 109 718 znaků s mezerami, bez poznámek pod čarou

Abstrakt

Základním pramenem datové žurnalistiky, která se v posledních letech etablovala v českém mediálním prostoru, jsou datové sady a přístup k informacím. Akademická obec v rámci žurnalistického diskurzu téma otevřenosti dat a jejich důležitost doposud opomíjela. Datoví žurnalisté jsou přitom mnohdy prvními, kdo data široké veřejnosti zpřístupní. Nedílnou součástí diplomové práce je představení historických předpokladů pro vývoj datové žurnalistiky, fenoménu otevírání dat státní správou, a ukotvení termínu otevřená data v české a evropské legislativě. Práce dále zkoumá využití Zákona č. 106/1999 Sb. o svobodném přístupu k informacím v žurnalistické praxi. Cílem práce je na základě hloubkových rozhovorů s českými datovými novináři identifikovat zdroje dat, které nejčastěji využívají ve svých redakčních rutinách, zhodnotit kvalitu těchto zdrojů a představit hlavní limity, na které novináři narážejí. V praktické části teoretické poznatky autorka práce přenesla do analýzy dostupnosti datových zdrojů poskytovaných Ministerstvem zdravotnictví a Ústavem zdravotnických informací a statistiky (ÚZIS) během pandemie COVID-19 v letech 2020 a 2021. Na konkrétních případech nastíní proces získávání, čištění, analýzy a vizualizace získaných informací novináři.

Abstract

The fundamental source of data journalism, which has established itself in the Czech media landscape in recent years, is the availability of usable data sets. However, academicians have so far neglected the topic of data openness and its importance in journalistic discourse. Simultaneously, data journalists are often the first to make data available to the public. An integral part of the thesis is mapping the historical development of data journalism, including the phenomenon of open government with a focus on Czech and European legislation and obtaining information with the help of the Freedom of Information Act. The aim of the thesis is, based on in-depth interviews with Czech data journalists, to identify the data sources which they frequently use in their editorial routines, to evaluate the quality of these sources and to present the main limits that journalists encounter. In the conclusive part of the thesis the author transfers the focus to analysis of the availability of data sources provided by the Ministry of Health and the Institute of Health Information and Statistics during the COVID-19 pandemic in 2020 and 2021 and outlines the process of acquisition, cleaning, analysis, and visualization of information by journalists.

Klíčová slova

Datová žurnalistika, otevřená data, Zákon č. 106/1999 Sb. o svobodném přístupu k informacím, data státní správy, pandemie COVID-19

Keywords

Data journalism, open data, Freedom of Information Act, government data, COVID-19 pandemic

Title

Open Data and their use in Czech Data Journalism

Poděkování

Na tomto místě bych ráda poděkovala vedoucímu práce PhDr. Václavu Moravci, Ph.D., Ph.D. a respondentům, kteří diplomové práci přispěli letitými zkušenostmi a nekončícími snahami otevírat data. Nemalé poděkování si zaslouží také nejbližší, kteří mě po celou dobu studia podporovali.

Obsah

Úvod	3
1. Teoretická část	5
1.1. Historické předpoklady vývoje datové žurnalistiky	6
1.1.1. Počítačově asistované zpravodajství (CAR)	6
1.1.2. Vznik online novináře	7
1.2. Dopady mediální konvergence na datovou žurnalistiku	8
1.2.1. Interaktivita	8
1.2.2. Lokálnost	9
1.3. Limity datové žurnalistiky	9
1.4. Datová žurnalistika v českém mediálním prostředí	11
1.5. Otevřená data	12
1.5.1. Otevřená data státní správy	13
1.6. Ukotvení otevřených dat v českém právním řádu	15
1.6.1. Národní katalog otevřených dat	15
1.6.2. Neziskové organizace	17
1.7. Zákon č. 106/1999 Sb. o svobodném přístupu k informacím	17
1.7.1. Znění zákona	18
1.7.2. Formální náležitosti podání žádosti a stanovené lhůty	20
1.7.3. Příklady využití informačního zákona v českém mediálním prostředí	21
1.7.4. Vizualizace dat získaných podle informačního zákona	23
1.8. Evropská legislativa	23
1.8.1. Srovnání ČR v rámci unie	25
1.9. Kvantitativní analýza textů datové žurnalistiky v českých médiích	25
1.9.1. Výsledky kvantitativní analýzy	26
2. Metodologie	27
2.1. Kvalitativní výzkum	28
2.1.1. Hlubkový rozhovor	29
2.1.1.1. Referenční studie – hloubkové rozhovory	29
2.1.1.2. Formulace otázek	30
2.1.2. Případová studie	32
2.2. Výběr vzorku	32
2.2.1. Redakce SAMIZDAT	33
2.2.2. Redakce Deník	34
2.2.3. Redakce Seznam Zprávy.cz	36

2.3.	Analýza rozhovorů	37
2.3.1.	Úvodní otázky	38
2.3.2.	Otázky zaměřené na práci s daty	39
2.3.3.	Hlavní otázky	42
2.4.	Závěry metodologické části	48
3.	Praktický příklad práce s daty – případová studie	49
3.1.	Časová osa	49
3.2.	Závěry případové studie	63
3.2.1.	Univerzální dostupnost dat ve strojově čitelném formátu	64
3.2.2.	Přístup k otevřeným datům není nikomu limitován	64
3.2.3.	Data sety jsou opatřeny dokumentací a meta daty dostatečně vysvětlující jejich hodnoty	65
	Závěr	66
	Summary	67
	Použitá literatura	69
	Tištěné zdroje	69
	Online zdroje	71
	Zpravodajské texty	74
	Závěrečné práce	75
	Zákony, novely, usnesení a nálezy	76
	Podklady případové studie	76
	Teze Diplomové práce	80
	Seznam příloh	83
	Přílohy	83

Úvod

„Nejzajímavější způsob využití vašich dat vymyslí někdo jiný.“¹

Žurnalistika 21. století klade důraz na rychlost a srozumitelnost. Práce s daty je mnohdy zdoluhavá a komplikovaná. Jak tedy pracují datoví novináři, kteří z komplexních informačních bodů produkují jasná sdělení?² Formuje dostupnost otevřených zdrojů jejich činnost a pokud ano, do jaké míry? V rámci žurnalistického diskurzu neexistuje ucelený pohled na možnosti využití otevřených dat a postupy získávání těch neveřejných. Důležitost otevřeného přístupu k informacím v posledním roce silně akcentovala pandemie COVID-19, která státní správě dala do rukou obrovské množství nových dat a rozhodnutí, jak s nimi naložit. Situace poukázala na limity ze strany novinářské obce v případech, kdy data přístupná nejsou a zároveň vyzdvihla důležitost využití dat v žurnalistice: přijímání informovanějších rozhodnutí.

Závěrečná práce je členěna do třech tematických celků. Jejím cílem je poskytnout vhled do problematiky otevřených dat, nastínit redakční rutiny práce tuzemských datových novinářů a na konkrétních případech demonstrovat limity jejich profese.

V teoretické části diplomantka definuje základní pojmy spojené s historickým vývojem datové žurnalistiky, mapuje otevírání dat státní správou, s ním spojenou legislativu, a nabízí srovnání České republiky se zahraničím. Podstatná část práce je věnována získávání neveřejných data setů, a to za pomoci Zákona č. 106/1999 Sb. o svobodném přístupu k informacím. Rozšířenost a relevanci datové žurnalistiky v rámci českého mediálního prostředí bude zkoumat kvantitativní analýza.

Pro analyzování dostupnosti datových zdrojů jsou však zásadní metody kvalitativní: hloubkový rozhovor a případová studie. Od tezí práce, které zahrnovaly metody srovnání a dotazování, se postupy výzkumu odlišují. Důvodem odklonu od původních metod je turbulentní vývoj situace ve zkoumané oblasti, který lze nejlépe postihnout případovou studií. Na základě hloubkových rozhovorů s vybranými českými datovými novináři se práce

¹ POLLOCK, Rufus. *The Open Revolution: Rewriting the Rules of the Information Age*. Open Revolution, 2018, 126 s. ISBN 9781983033223.

² STUHLÍKOVÁ, Lucie. Clarifying the complex: Czech Radio's data journalism team. *European Journalism Observatory* [online]. 19.12.2019 [cit. 2021-12-06]. Dostupné z: <https://en.ejo.ch/specialist-journalism/clarifying-the-complex-czech-radios-data-journalism-team>

pokusí vymezit limity, na které novináři narážejí. V praktické části práce, případové studii, budou tyto poznatky využity k analýze dostupnosti dat během pandemie COVID-19. Použitým metodám práce a vybranému vzorku se diplomátka blíže věnuje v metodologické části.

Výběr tématu závěrečné práce autorka nezvolila náhodně, využití dat v žurnalistice je tématem více než aktuálním, jak již bylo zmíněno pandemie COVID-19 důležitost kvalitních a dostupných dat nebývale zdůraznila. Je nutné zmínit, že autorka práce má s datovou žurnalistikou osobní zkušenosti. Na akademické půdě se datové analýze a vizualizaci věnovala během studia žurnalistiky ve Spojených státech v letech 2017 a 2018. O rok později se připojila k datové redakci Českého rozhlasu, ve které doposud působí. Podněty pro zpracování tématu tedy vzešly z metody zúčastněného pozorování, jež má výhody v otevřenosti vůči zkoumaným konceptům a schopnosti zachytit rutinní situace.³ Autorka práce však zaručuje objektivitu při analýze získaných výsledků a pokusí se téma uchopit nestranně. Odborná literatura mediálních studií datové žurnalistice v posledních letech věnuje okrajovou pozornost, proto autorka práce odstoupila od více obecných definic k jedné části této problematiky: práce s daty.

³ ŠVAŘÍČEK, Roman, Klára ŠEĐOVÁ. Kvalitativní výzkum v pedagogických vědách. Vyd. 2. Praha: Portál, 2014. ISBN 978-80-262-0644-6.

1. Teoretická část

Bounegru, Chambers a Gray (2012) data definují jako: „malé informační body, které jsou často v jednotlivých případech nedůležité, ale ze správného úhlu pohledu nesmírně podstatné.“⁴

Tvrzení potvrdila první analýzou dat Florence Nightingale již v roce 1858. V grafice známé pod názvem Mortality of British Army komunikovala široké veřejnosti informaci o tom, že většina vojáků britské armády nezemřela v boji, nýbrž na nemoci, kterým se dalo předcházet. Byl to kontext, který tato první předchůdkyně moderní datové vědkyně malým informačním bodům přidělila a vizualizace, která z nich učinila pozornost poutající zjištění. Pro své analýzy používala menší množství dat a manuální výpočty hodné 19. století.⁵ Novodobý novinář má s postupným zvyšujícím se počtem datových souborů v kombinaci s přístupnějšími nástroji pro jejich zpracování téměř neomezené možnosti informovat o problémech či příbězích, které dříve zůstaly skryty.⁶

Nástup digitální a online žurnalistiky dokonce institucionalizoval spojení žurnalistické práce a datové analýzy pod pojem datová žurnalistika. Využití dat ve zpravodajství či publicistice ale není zdaleka nic nového. Novináři o cifrách informovali vždy, ať už se jednalo o nová zjištění statistických úřadů či témata z ekonomické sféry.⁷

Datoví novináři, kteří zkoumají důležité oblasti veřejného zájmu ve větším detailu, se ale odlišují inklinací k výzkumným metodám společenských věd: kvantitativní či statistické analýze. Datová žurnalistika ze své přirozenosti zasahuje do odvětví investigativní žurnalistiky, může být nositelem silného veřejného zájmu, kontrolovat a vytvářet tlak na vládnoucí garnitury.⁸

Felle (2016) v této souvislosti vyzdvihuje témata datové žurnalistiky s přesahem do veřejných výdajů, zakázek či dotací a dodává, že publikování nezpracovaných, původních, dat zvyšuje důvěryhodnost zjištění (...). V médiích použití dat přispělo k většímu zapojení

⁴ BOUNEGRU, Liliana, Lucy CHAMBERS a Jonathan GRAY. *The Data Journalism Handbook*. European Journalism Centre, 2012. ISBN 9781449330064.

⁵ BRADSHAW, Noel-Ann. *Florence Nightingale (1820–1910): An Unexpected Master of Data* [online]. 2020 [cit. 2021-12-02]. Dostupné z: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7660360/>. School of Computing and Digital Media, London Metropolitan University, London, UK.

⁶ FELLE, Tom. Digital Watchdogs? Data Reporting and the News Media's Traditional 'Fourth Estate' Function. *Journalism* 17, no. 1, 2016, str. 85–96. <https://doi.org/10.1177/1464884915593246>.

⁷ Tamtéž, str. 92

⁸ Tamtéž, str. 88

publika, navýšilo kredibilitu zpráv a čtenářům nabídlo možnost ověřit si informace u původního zdroje, což může vést až k vyvracení dezinformací.⁹

1.1. Historické předpoklady vývoje datové žurnalistiky

Rozvoji datové žurnalistiky předcházela technologická mediální konvergence, která dle Granta (2009) započala přechodem od analogového k digitálnímu vysílání. Vyústit měla přijetím digitálních technologií jako primárních formátů pro přenos informací. S odstupem více než dekády lze dodat, že k této transformaci skutečně došlo, digitální technologie se staly běžnou součástí novinářské práce. Mediální technologická konvergence stále probíhá a dle Deuze (2007) dále sblížuje data směrem k jedné platformě.

1.1.1. Počítačově asistované zpravodajství (CAR)

Bounegru, Chambers a Gray (2012) za klíčové pro rozvoj datové žurnalistiky, kterou nazývají žurnalistikou „veřejné služby“, považují rozšíření počítačově asistovaného zpravodajství,¹⁰ které poprvé využila americká komerční televizní síť CBS již v 50. letech 20. století, a to během amerických prezidentských voleb.

K masivnímu využití počítačových technologií v redakcích po celém světě došlo o zhruba 40 let později, a to s rozšířením internetové sítě. Internet novinářům nabídl možnosti vyhledávání zdrojů online, nové technologie umožnily rychlejší ověřování a analýzu informací, mezi které patřila také statistická data.¹¹ Od přejímání oficiálních dat se novináři s postupnou adaptací nových nástrojů přesunuli k sestavování vlastních databází, a to ve snaze nezávisle ověřovat data veřejných institucí.¹² Pro tento způsob novinářské práce se v 70. letech 20. století prosadil termín „precizní žurnalistika“ (z anglického „*precision journalism*“).¹³

Jak dodávají Nguyen a Lugo-Ocando (2016) sdílení statistik k žurnalistice vždy patřilo a je nedílnou součástí denního zpravodajství, od „*hard news*“ neboli zpráv o politice, ekonomice, podnikání, financích, vědě a vzdělávání až po „*soft news*“ kategorie, jako jsou

⁹ FELLE, Tom. Digital Watchdogs? Data Reporting and the News Media's Traditional 'Fourth Estate' Function. *Journalism* 17, no. 1, 2016, str. 85–96. <https://doi.org/10.1177/1464884915593246>.

¹⁰ Z anglického „Computer assisted reporting“ (CAR)

¹¹ ČURÍK, Jaroslav, Rudolf BURGR, Jiřina SALAQUARDOVÁ, Martin ŠKOP a Lenka WASCHKOVÁ-CÍSAŘOVÁ. *Nové trendy v médiích I.* [online]. 2. Brno: Masarykova univerzita, 2012 [cit. 2021-12-02]. ISBN 978-80-210-5826-2. Dostupné z: https://is.muni.cz/publication/987880/Nove_trendy_v_mediich_I.pdf

¹² FELLE, 2016, str. 88

¹³ BOUNEGRU, CHAMBERS a GRAY, 2012, str. 23

zdraví, kriminalita, sport a zábava. Až nové technologie a masmédiá ale tento trend rozvinuly. Vyhličky do budoucna přitom naznačují, že datová žurnalistika nezanikne, ale bude se „více rozvíjet, jelikož se společnost založená na velkém množství dat známé jako „big data“ postupně normalizuje.“¹⁴

Z hlediska technologií tedy mediální konvergence nyní již nepostupuje s takovou rychlostí, kvantita sbíraných dat a jejich potenciální využití v žurnalistice však nadále exponenciálně roste.

1.1.2. Vznik online novináře

Mediální konvergence proměnila role tradičních žurnalistů, na které je kladen silnější požadavek ke kombinaci způsobů jejich práce. Novodobé žurnalisty, kteří pracují multimediálně, Waschková Čísařová (2012) označuje pojmem „online novinář.“ Definuje jej jako nadšeného technologického inovátora, kterého v začátcích tradiční média nebrala příliš vážně. V praxi se jedná o člověka, který vyjma reportérskou činnost zvládá pořizování fotografií, videí či audiozáznamů a jejich následnou post produkci. Pro setrvání v médiích dnešní doby je osvojení praktik online novináře již standardem.

V případě datových novinářů přibyl požadavek na ovládání základů programovacích jazyků či statistiky. Lewis a Usher (2013) popisují také okrajový fenomén fúze informatiky a žurnalistiky, která vygenerovala nový typ novináře-programátora, tedy „hackerského novináře“ (Betancourt 2009, Minkoff 2011). Programátoři přebírají důležitější role ve zpravodajských organizacích a přispívají k výstupům vycházejících či doplněných o data, zpravodajské algoritmy či jiné projekty založené na kódování.¹⁵ Využívají přitom

¹⁴ NGUYEN, An a Jairo LUGO-OCANDO. The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism* [online]. 2016, 17, 3-17 [cit. 2021-10-02]. Dostupné z: doi:10.1177/1464884915593234

¹⁵ LEWIS, Seth C. a USHER, Nikki. Open Source and Journalism: Toward New Frameworks for Imagining News Innovation. *Media, Culture & Society* 35, no. 5, 2013, str. 2–109. <https://doi.org/10.1177/0163443713485494>.

softwarových nástrojů pro vytěžování velkého objemu dat (z anglického „*data mining*“) či stahování dat z webových stránek (z anglického „*data scraping*“).¹⁶

V metodologické části práce se autorka pokusí identifikovat, do jaké míry jsou technologické znalosti podstatné pro práci tuzemských datových novinářů a zda se vnímají více jako novináři či programátoři.

Ačkoliv se může zdát, že nové technologie novináře absolutně pohltily, teoretik médií McQuail (2009) dodává, že „*vzestup nové technologie obvykle zcela nezatlačí do pozadí stará média, nýbrž je nutí se přizpůsobit novým tržním podmínkám,*“¹⁷ a proto ani autorka práce nepředpokládá, že nové požadavky na novináře upozadí jejich primární funkci „*autorizovaných vypravěčů pravdy*“¹⁸ a základními požadavky zůstanou přesnost a faktičnost komunikovaných sdělení.

V neposlední řadě k rozvoji datové žurnalistiky přispělo stále větší množství elektronicky uchovávaných informací. Felle (2016) s odkazem na Frey či Maude (2014, 2012) v této souvislosti vyzdvihuje odklon mnoha vlád a úřadů od papírových formulářů k trendu „*otevřeného vládnutí*“.

1.2. Dopady mediální konvergence na datovou žurnalistiku

Přechod od tradičních médií do online prostředí nabídl širokou škálu nových, vizuálních možností komunikace. Atraktivita těchto nových nástrojů, a především pak jejich interaktivní povaha dle Felleho (2016) přináší zcela nové způsoby vyprávění příběhů.

1.2.1. Interaktivita

Opakovaným rysem datových výstupů je jejich interaktivní povaha. Příkladem mohou být interaktivní mapy, ve kterých čtenář v několika málo krocích může zkoumat a porovnávat hodnoty na odlišných místech, ať už známých či vzdálených. Interaktivita má potenciál zvýšit atraktivitu žurnalistiky, vést k vyšší informovanosti a dosáhnout většího zapojení veřejnosti do společenského dění.¹⁹

¹⁶ BOUNEGRU, Liliana, Lucy CHAMBERS a Jonathan GRAY. *The Data Journalism Handbook*. European Journalism Centre, 2012. ISBN 9781449330064.

¹⁷ MCQUAIL, Denis. Úvod do teorie masové komunikace. 4., rozš. a přeprac. vyd. Praha: Portál, 2009. ISBN 978-80-7367-574-5.

¹⁸ MCNAIR, Brian. *Sociologie žurnalistiky*. Přeložil Hana ANTONÍNOVÁ. Praha: Portál, 2004. ISBN 80-7178-840-6.

¹⁹ FELLE, 2016, str. 92

1.2.2. Lokálnost

Zatímco tradiční média v minulosti reportovala spíše o celonárodních trendech, či neměla možnosti věnovat se lokálním příběhům více do hloubky, interaktivní datové analýzy a vizualizace mohou čtenářům nabídnout unikátní, lokální příběhy.²⁰ Tento trend autorka práce vyzorovala také u tuzemské datové produkce. Například regionální médium Deník, který má v České republice více než 70 mutací, využívá data až na úroveň obcí. Datoví novináři veřejnoprávního Českého rozhlasu interaktivitu a lokálnost využívají během voleb, kdy má čtenář přístup k výsledkům na úroveň jednotlivých volebních místností a může se tak ztotožnit s místem události.

Goggin a Dwyer (2015) dodávají, že interaktivita a lokálnost v menší míře byly vždy součástí tradiční žurnalistiky, ale až s rozšířením mobilních technologií došlo k jejich rozvoji do trendu lokalizovaných zpráv (z anglického „*locative news*“).²¹

Jak tvrdí Lorenz (2012) díky těmto prvkům je žurnalistika osobnější a datová novinářina má potenciál obohatit či vzdělat publikum vybízením k uzpůsobování příběhu na personalizovaný narativ.²²

1.3. Limity datové žurnalistiky

Datová žurnalistika je silně determinovaná dostupností a otevřeností datových souborů a její důležitost tkví v dopadech na společenské dění (Porlezza, Splendore 2019). Datoví žurnalisté umožňují dalším aktérům, čtenářům či nevládním organizacím, dále zkoumat a šířit data nebo závěry z nich vyplývající (Lewis, Usher, 2013) a přivádět držitele moci k odpovědnosti.

Spoléhání se pouze na data státní správy, a ochotu úřadů data poskytnout, ale může být problematické pro naplnění role novináře jako „hlídače demokracie“. Datoví novináři si nemožnost data ověřit či je získat uvědomují a podle Porlezza a Splendore (2019) rozlišují své výstupy na ambicióznější datové projekty, které zahrnují sběr dat či jejich získávání, a projekty, které se spoléhají pouze na dostupná data veřejné správy.²³

Žurnalistika, která se silně spoléhá na data a jejich statistické hodnocení může vést také k potenciálnímu riziku, které Felle (2016) s odkazem na Dickinsona (2014) popisuje jako riziko exkluzivity a vymezení se jen pro technologicky zdatné či gramotné publikum.

²⁰ Tamtéž, str. 94

²¹ GOGGIN, G., MARTIN, F., a DWYER, T., 2015. *Locative News*. *Journalism Studies*, 16, str. 4-59.

²² LEWIS, Seth C. a USHER, Nikki, 2013, str. 88

²³ Tamtéž

Jinými slovy, aby datová žurnalistika měla možnost zasáhnout masové publikum a kontrolovat demokratické procesy, musí být přístupná a uchopitelná širokou veřejností.²⁴ Dickinson (2014) se v této souvislosti ptá:

„Opravdu zpřístupnění dat novinářům v tabulce Excel přispívá k demokratizaci informací? Může možnost vyhledávání informací podle poštovního směrovacího čísla skutečně ulehčit práci novinářů v terénu? Nevzniká jen malá žurnalistická elita?“²⁵

Limity datové žurnalistiky tedy mohou vnímat sami zástupci novinářské obce, ve které přetrvává stereotyp novináře jako člověka, který se rozhodl pro práci se slovy právě proto, že nechce nebo neumí pracovat s čísly (Yarnall 2008). Tento stereotyp se snaží vyvrátit Nguyen a Lugo-Ocando (2016), kteří dodávají, že se novináři nemusí nutně učit složité technologie či matematické operace, ale mnohdy pracují s předem připravenými datovými balíčky:

„Co datoví novináři potřebují nejvíce není perfektní znalost programování či zpracování dat, ale aplikování logických, platných, novinářských postupů a skepticismu k (a) vyhledávání a získávání dat, (b) zkoumání a hodnocení jejich skutečného významu a zasazení dat do kontextu, (c) zkoumání souvislostí, které data spoluvytvářejí a (d) zpracovávat data vyváženým, spravedlivým, přesným, přístupným či pochopitelným a poutavým způsobem.“²⁶

Limity datové žurnalistiky lze nalézt také na straně publika. Lidé mají tendenci věřit více číslům než slovům a způsob, jakým jsou data prezentována ve zpravodajství, hraje klíčovou roli v posilování nebo zpochybňování této víry.²⁷ Nepřesná či chybná prezentace dat proto může nést tristní dopady ve společnosti, zejména jedná-li se o data medicínská.

²⁴ FELLE, 2016, str. 88

²⁵ HERMIDA, Alfred a Mary Lynn YOUNG. *Data Journalism and the Regeneration of News* [online]. Routledge, 2019 [cit. 2021-10-02]. ISBN 9781138058934. Dostupné z: <https://www.routledge.com/Data-Journalism-and-the-Regeneration-of-News/Hermida-Young/p/book/9781138058934>

²⁶ NGUYEN, An a Jairo LUGO-OCANDO, 2016, str. 90

²⁷ Tamtéž, str. 93

Takové nesprávné interpretace mohou dle Reyna a kol. (2009) vyvolávat „zbytečné obavy, falešné naděje, nesmyslné změny životního stylu, utrácení za léky či nepřiměřený odpor k lékařským doporučením.“²⁸

Během pandemie COVID-19 s daty musely začít pracovat téměř všechny tuzemské redakce. Obavy ze špatné interpretace dat ze strany novinářů a jejich nepochopení na straně publika vedoucí k vyvolání paniky, ač oprávněné, přispěly k utajování informací ze strany úřadů, kterému se autorka práce více věnuje v metodologické a praktické části.

1.4. Datová žurnalistika v českém mediálním prostředí

V českém mediálním prostoru se datová žurnalistika jako samostatné odvětví začala etablovat okolo roku 2012 v Hospodářských novinách.²⁹ Nerozšířila se masově, v současnosti se datové analýze a vizualizaci nad míru oficiálních výstupů úřadů věnují žurnalisté v Českém rozhlasu, Seznam Zprávách či regionálním Deníku.³⁰

V zahraničí mají silné zastoupení redakce britských Financial Times a The Economist, či americké redakce serverů The New York Times, The Wall Street Journal, veřejnoprávní rozhlasová stanice National Public Radio (NPR). Známy pro analýzu volebních dat je analytik a bývalý novinář Nate Silver publikující na webu FiveThirtyEight.³¹

Tuzemská datová žurnalistika se od zahraniční ale v mnohém odlišuje. Z rozhovorů s českými datovými novináři, které autorka práce oslovila vyplývá, že čeští kolegové pracují s menšími finančními rozpočty, datová práce má menší tradici a příliš se neklade důraz na komerční aktivity. Česká datová žurnalistika je tedy méně specializovaná a novináři pokrývají větší množství témat. Datový novinář Českého rozhlasu Jan Cibulka v rozhovoru pro účely diplomové práce uvedl, že tyto limity vedou k nutnosti publiku opakovaně vysvětlovat postupy datové analýzy a vizualizace, což brání přechodu ke složitějším tématům. Nabízí se zamyšlení, zda zatímco v počátcích datové novinářské práce omezovaly

²⁸ Tamtéž, str. 103

²⁹ Datová žurnalistika v Česku v roce jedna. *Nadace OSF* [online]. 28.06.2012 [cit. 2021-11-12]. Dostupné z: <https://osf.cz/2012/06/28/datova-zurnalistika-v-cesku-v-roce-jedna/>

³⁰ Blíže se redakcím datové žurnalistiky autorka práce věnuje v metodologické části

³¹ Ve Spojených státech je datová žurnalistika vyjma hlavní vydavatelské domy rozšířena okrajově také do lokálních redakcí, mezi příklady dobré praxe lze uvést noviny The Tampa Bay Times, The Detroit News, Iowskou The Gazette, či stanici ABC v Severní Karolině.

technologické znalosti, nyní se pozornost přesouvá směrem k formě komunikovaných informací a k získávání dat, která nejsou veřejná.

1.5. Otevřená data

Získávání a práce s neveřejnými daty je sice součástí datové žurnalistiky, ale většinová část stále zahrnuje zpracování dat otevřených. Lewis a Usher (2013) s odkazem na Pollocka (2006) data definují jako otevřená, pokud je kdokoliv může volně, opakovaně používat a redistribuovat nanejvýš s požadavkem data zveřejňovat ve stejné podobě, v jaké je získal.

Tato definice však podle Kitchina (2014) naráží na různorodost toho, co lze chápat pod pojmem „otevřený“ a jak toto chápání ovlivňuje právo na duševní vlastnictví dat. Je nutné mít na paměti, že právo otevřená data použít, přepracovat a předat, nutně nemusí znamenat právo data zveřejňovat za účelem zisku. Přístup k datům tedy může být otevřený, co lze s daty dále dělat nikoliv.³²

Základní požadavky, kterých data musí dosáhnout, aby se dala považovat za otevřená, do třech hlavních bodů shrnuje organizace Open Knowledge Foundation³³:

- 1) *Dostupnost a přístup: data musí být dostupná jako celek, nejlépe stažením přes internet, a to volně či za nízký poplatek. Data musí být dostupná ve vhodném, upravitelném a strojově čitelném formátu*
- 2) *Opětovné použití a redistribuce: data musí být poskytnuta za podmínek, které umožňují jejich opětovné použití*
- 3) *Univerzální využití: každý musí být schopen data obstarat a použít. Neměla by existovat žádná diskriminace. Data nesmí být limitována na určité použití, například pouze ve vzdělávání*

Otevřená data nejsou specifikem žádného oboru. Mohou zasahovat do sfér socio-ekonomických, kulturních, environmentálních nebo vědeckých. Důraz je ale v obecné rovině „*kladen na zveřejňování dat, která mají vysoký veřejný zájem.*“³⁴

³² KITCHIN, Rob. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences* [online]. California: Sage Publication, 2014 [cit. 2021-11-02]. ISBN 978-1446287484. Dostupné z: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jors.12293>

³³ What is open? *Open Knowledge Foundation* [online]. [cit. 2021-11-05]. Dostupné z: <https://okfn.org/opendata/>

³⁴ Tamtéž, str. 102

Především proto, aby veřejnost mohla v oblastech, která mají přímý dopad na jejich život, získat erudovanější informace a učinit tak potenciálně informovanější rozhodnutí. Z tohoto důvodu podle Janssen a Yiu (2012) otevřená data podporují „*aktivní účast na veřejném životě, sociální inovace, posílení vztahů v komunitě a zvyšují úroveň veřejné debaty. V důsledku toho otevřená data umožňují a podporují participativní demokracii.*“³⁵

1.5.1. Otevřená data státní správy

Porter (1996) již ke konci 20. století označil státní aparát a menší samosprávné celky za hlavní sběratele a zpracovatele dat. Přednostmi státní správy v této oblasti jsou vysoká organizovanost a tradice data shromažďovat.³⁶ Vytvoření a údržba datových sad může být časově i finančně náročná, a v důsledku toho jsou mnohdy data a z nich odvozené informace uzavřeny v archívech jednotlivých institucí. Limitace přístupu k vybraným datovým sadám státní správy či zpoplatnění jejich využití tedy není podle Kitchina (2014) zcela neoprávněné.

Státní data mají pro odborníky, novináře, či širokou veřejnost, obrovský potenciál. Přínosná mohou být ale také pro samotný státní aparát, který jejich zpřístupněním může zvýšit efektivitu státní správy³⁷ a vytvářet tlak na řešení problémů.³⁸

Hlavní benefity otevírání dat státní správy³⁹:

- *Zlepšování kvality veřejných služeb*
- *Vyšší transparentnost a otevřenost veřejné správy*
- *Nový zdroj podnikatelských příležitostí*
- *Zlepšení komunikace úřadů mezi sebou*

Jako první si benefity otevírání dat státní správy uvědomily Spojné státy⁴⁰ a vytvořily úřad Nejvyššího informačního úředníka, což mělo vést k institucionalizaci otevírání datových sad.

³⁵ Tamtéž, str. 105

³⁶ LOWREY, Wilson, Ryan BROUSSARD a Lindsey A. SHERRILL. *Data journalism and black-boxed data sets* [online]., 2019, str. 62-89 [cit. 2021-11-02]. Dostupné z: <https://journals.sagepub.com/doi/pdf/10.1177/0739532918814451>

³⁷ Why open data? *Open Knowledge Foundation* [online]. [cit. 2022-11-07]. Dostupné z: <https://okfn.org/opendata/why-open-data/>

³⁸ BOČEK, Jan, Jakub MRÁČEK a Jindřich MYNARZ. *Otevřená data: Příležitost pro Českou republiku* [online]. Nadace Open Society Fund Praha: Fond Otakara Motejla [cit. 2021-11-02]. ISBN 978-80-87725-03-0. Dostupné z: https://osf.cz/wp-content/uploads/2012/08/Opendata2_web_komplet.pdf

³⁹ Tamtéž, str. 12-14

⁴⁰ BOČEK, Jan, Jakub MRÁČEK a Jindřich MYNARZ, str. 5

V roce 2009 se postu ujal Vivek Kundra. Poskytování informací podle Kundry mělo přinést mimo jiné „*nové formy žurnalistiky, které dokážou oddělit signál od šumu a poskytnout smysluplné informace.*“⁴¹

Prvních 47 datových sad Američané publikovali na webu Data.gov, který se v současné době rozrostl o více než 337 tisíc data setů federálních úřadů i vlád jednotlivých států.⁴² Spojené státy v roce 2010 následovalo Spojené království.⁴³

Česká republika své první závazky v oblasti otevřených dat představila v roce 2012, a to závazkem k vytvoření katalogu otevřených dat státní správy. V Akčním plánu České republiky 2011-2012 se pod vedoucí institucí Ministerstvem vnitra k příspěvní do katalogu zavázaly Ministerstvo spravedlnosti, Ministerstvo pro místní rozvoj, Ministerstvo financí, Ministerstvo dopravy a Český statistický úřad.⁴⁴

Je diskutabilní, zda absolutní počty zveřejněných data setů objektivně vypovídají o situaci v dané zemi. I přesto, autorka práce poskytne základní srovnání organizace Global Open Data Index, podle které mezi nejčastější okruhy poskytovaných dat státní správy patří data⁴⁵:

- *Státní rozpočet: data o plánovaných výdajích a příjmech na nadcházející rok*
- *Národní statistiky: demografické a ekonomické ukazatele, například hrubý domácí produkt (HDP), nezaměstnanost*
- *Zadávání veřejných zakázek: monitoring výběrových řízení, všechny zakázky seskupeny podle úřadů*
- *Kvalita ovzduší a vody: Údaje o průměrné denní koncentraci látek znečišťujících ovzduší, zejména těch, které jsou potenciálně škodlivé pro lidské zdraví*
- *Firemní rejstřík*
- *Vlastnictví půdy*

⁴¹ KUNDRA, Vivek. Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect. *Harvard Kennedy School* [online]. 2012, 25.01.2012 [cit. 2021-10-08]. Dostupné z: <https://shorensteincenter.org/digital-fuel-of-the-21st-century-innovation-through-open-data-and-the-network-effect/>

⁴² Dostupné z: <https://www.data.gov/>

⁴³ BOČEK, Jan, Jakub MRÁČEK a Jindřich MYNARZ, str. 6

⁴⁴ Czech Republic. *Open Government Partnership* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.opengovpartnership.org/members/czech-republic/#commitments>

⁴⁵ Tracking the state of open government data. *Global Open Data Index* [online]. [cit. 2021-11-02]. Dostupné z: <https://index.okfn.org/>

1.6. Ukotvení otevřených dat v českém právním řádu

Otevřená data jsou definována v českém právním řádu, a to dle § 3 odst. 11 Zákona o svobodném přístupu k informacím č. 106/1999 Sb. jako:

„(...) informace zveřejňované způsobem umožňujícím dálkový přístup v otevřeném a strojově čitelném formátu, jejichž způsob ani účel následného využití není omezen a které jsou evidovány v Národním katalogu otevřených dat.“⁴⁶

Informace o životním prostředí, zejména ty, které vypovídají o jeho stavu a vývoji lze požadovat dle samostatného Zákona č. 123/1998 Sb. o právu na informace o životním prostředí.

Přístup k informacím lze jistě vnímat také jaké součást práva na svobodu projevu, kterou dle Míška (2021) není možné vykonávat bez možnosti být dostatečně informován. V českém právním řádu svobodu projevu zaručuje Ústavní zákon č. 2/1993 Sb., Listina základních práv a svobod. Článek 17 dále stanovuje, že *„cenzura je nepřípustná a státní orgány jsou přiměřeným způsobem povinny poskytovat informace o své činnosti.“⁴⁷*

Svobodná média jsou také jednou z podmínek demokracie podle definice Evropského soudu pro lidská práva, který zdůrazňuje, že pokud má žurnalistika přispívat k otevřené diskusi a zastávat veřejný zájem, zásahy do jejího fungování mají být co nejmenší.⁴⁸

1.6.1. Národní katalog otevřených dat

Legislativně je zveřejňování informací státní správy podmíněno existencí informačního systému: Národního katalogu otevřených dat (NKOD), který spravuje Ministerstvo vnitra na Portálu otevřených dat České republiky.⁴⁹ Národní katalog je každoročně analyzován

⁴⁶ Celé znění zákona je součástí příloh diplomové práce, online verze je dostupná z:

<https://www.zakonyprolidi.cz/cs/1999-106?text=svobodn%C3%A9m+p%C5%99%C3%ADstupu>

⁴⁷ Usnesení č. 2/1993 Sb.: Usnesení předsednictva České národní rady o vyhlášení LISTINY ZÁKLADNÍCH PRÁV A SVOBOD jako součástí ústavního pořádku České republiky. *Zákony pro lidi* [online]. [cit. 2021-11-22]. Dostupné z: <https://www.zakonyprolidi.cz/cs/1993-2?text=%C3%9Astavn%C3%AD+z%C3%A1kon+%C4%8D.+2%2F1993+Sb>

⁴⁸ ČURÍK, Jaroslav, Rudolf BURGR, Jiřina SALAQUARDOVÁ, Martin ŠKOP a Lenka WASCHKOVÁ-CÍSAŘOVÁ, 2012, str. 185

⁴⁹ Dostupné z: <https://data.gov.cz/>

Hlavním architektem eGovernmentu Ministerstva vnitra České republiky za účelem implementace nových technologických norem a funkcí.

Ministerstvo vnitra dále upřesňuje technologické požadavky, kterých by měla otevřená data docílit. Otevřená data by měla být:

- *Přístupná jako datové soubory ve strojově čitelném a otevřeném formátu s úplným a aktuálním obsahem databáze nebo agregovanou statistikou*
- *Opatřená neomezujícími podmínkami užití*
- *Evidovaná v Národním katalogu otevřených dat (NKOD) jako přímé odkazy na datové soubory*
- *Opatřená dokumentací*
- *Dostupná ke stažení bez technických překážek (registrace, omezení počtu přístupů apod.)*
- *Připravena s cílem co nejsnazšího strojového zpracování programátory apod.*
- *Opatřená kontaktem na kurátora pro zpětnou vazbu (chyby, žádost o rozšíření apod.)⁵⁰*

Za otevřená naopak ministerstvo nepovažuje data, která jsou limitována vyhledávacím formulářem, registrací či smlouvou, nebo jsou poskytována v nedostatečné kvalitě, nevyhovujících formátech jako je PDF, XLS či chybná CSV.⁵¹ Podstatnou součástí katalogu je ukládání meta dat, tedy doplňujících informací, „ze kterých je možné zjistit, kdo data publikuje, zda a za jakých podmínek je možné data použít či kdy byla data publikována.“⁵²

Seznamy informací, které je nutné zveřejňovat jako otevřená data v Národním katalogu otevřených dat upravuje novela nařízení vlády č. 425/2016 Sb., o seznamu informací zveřejňovaných jako otevřená data. V platnost vstoupila na přelomu roku 2019 a

⁵⁰ MÍŠEK, Jakub. Co jsou otevřená data? *Otevřená data* [online]. 2021 [cit. 2021-11-02]. Dostupné z: <https://opendata.gov.cz/informace:start>

⁵¹ Tamtéž

⁵² Národní katalog otevřených dat na Portálu veřejné správy. *Ministerstvo vnitra České republiky* [online]. [cit. 2021-11-09]. Dostupné z: <https://www.mvcr.cz/clanek/zpravodajstvi-narodni-katalog-otevrenych-dat-na-portalu-verejne-spravy.aspx>

rozšířila počet stávajících povinně zveřejňovaných datových sad z 11 na 24.⁵³ Ke konci roku 2021 je možné vyhledat více než 175 různých tematických okruhů datových sad od 52 poskytovatelů. Mezi největší poskytovatele v objemu datových sad řadíme Český úřad zeměměřický a katastrální (128 639 datových sad), Ministerstvo vnitra (4008 sad) a Český statistický úřad (701 sad).⁵⁴

1.6.2. Neziskové organizace

Vyjma státní správy v České republice snahu data otevírat projevují také neziskové organizace. Mezi nejvýraznějšími Nadace OSF a její aktivity v rámci projektu Náš stát, naše data, který podporuje aktivity občanů v oblasti digitalizace státní správy. Propojuje také odborníky na konferencích Open Data Expo či Civic Tech Open a každoročně pořádá soutěž Společně otvíráme data. Soutěž oceňuje inovativní projekty založené na otevřených datech, včetně výstupů datové žurnalistiky.

V oblasti získávání dat za pomoci Zákona č. 106/1999 Sb. o svobodném přístupu k informacím poskytuje poradenství nezisková organizace Otevřená společnost, ve které působí spoluautor daného zákona Oldřich Kužílek. Přístup k informacím je dle Kužílka (2004) pouze jednou částí otevřenosti státní správy. Otevřená státní správa by měla svým občanům nabízet otevřenost rozhodování, tedy umožnit jim „*vidět, jak se dospělo k rozhodnutím, včetně postupů, podkladů, hlasování, účasti na rozhodování (...) a právo šířit informace.*“⁵⁵

1.7. Zákon č. 106/1999 Sb. o svobodném přístupu k informacím

Cenzura je v demokracii nepřijatelná, a to včetně zatajování či neposkytnutí informací. I přesto se datoví novináři mnohdy potýkají s neochotou úřadů informace poskytnout. Způsobem obrany je již zmíněný Zákon č. 106/1999 Sb. o svobodném přístupu k informacím, pro účely této práce dále odkazován jako „informační zákon“. Informační

⁵³ ULMANOVÁ, Jiřina. *Souhrnná výzkumná zpráva o problematice otevřených dat* [online]. In: . Praha, 2021, s. 82 [cit. 2021-11-02]. Dostupné z: <https://vubp.cz/soubory/vyzkum/projekty/V03-S4/Vsouhrn-Souhrna-vyzkumna-zprava-o-problematice-otevrenych-dat.pdf>

⁵⁴ Tamtéž

⁵⁵ KUŽÍLEK, Oldřich. *Dalších 106 odpovědí na Vaše dotazy (s rozsudky): příručka pro občany o svobodném přístupu k informacím a otevřenosti veřejné správy*. Praha: Otevřená společnost, 2004. ISBN 80-903331-1-7.

zákon poskytuje právně vymahatelné prostředky k získání otevřených dat, podstatné je ale znát možnosti a limity daného zákona, aby mohl být efektivně využit v novinářské praxi. Termín „otevřená data“ byl do samotného informačního zákona ukotven novelou z roku 2016, která s účinností od prosince 2023 veřejným orgánům ukládá povinnost „*automaticky zveřejňovat informace ze svých veřejně přístupných registrů, evidencí, seznamů a rejstříků jako otevřená data.*“⁵⁶

1.7.1. Znění zákona

Informační zákon má 22 paragrafů. O informace veřejnost může dle § 2 zákona žádat pouze státní orgány, územní samosprávné celky, jejich orgány a veřejné instituce. Na fyzické osoby zákon nedopadá.

Orgány státní správy blíže definuje Zákon č. 2/1969 Sb., neboli kompetenční zákon, a to jako ústřední orgány státní správy, v jejichž čele je člen vlády. Do tohoto výkladu dle § 2 kompetenčního zákona spadají všechna ministerstva. Dále také: „*Český statistický úřad, Český úřad zeměměřický a katastrální, Český báňský úřad, Úřad průmyslového vlastnictví, Úřad pro ochranu hospodářské soutěže, Správa státních hmotných rezerv, Státní úřad pro jadernou bezpečnost, Národní bezpečnostní úřad, Energetický regulační úřad, Úřad vlády České republiky, Český telekomunikační úřad, Úřad pro ochranu osobních údajů, Rada pro rozhlasové a televizní vysílání, Úřad pro dohled nad hospodařením politických stran a politických hnutí, Úřad pro přístup k dopravní infrastruktuře, Národní úřad pro kybernetickou a informační bezpečnost a Národní sportovní agentura.*“⁵⁷

Seznam všech veřejných institucí ale žádný ze zákonů nedefinuje a může tedy představovat šedou zónu žádostí o informace. Znaky veřejné instituce nesou například Český rozhlas a Česká televize, měla by tedy být založena veřejnoprávním způsobem, plnit veřejný

⁵⁶ Viz. celé znění zákona v příloze diplomové práce

⁵⁷ Zákon č. 2/1969 Sb.: Zákon České národní rady o zřízení ministerstev a jiných ústředních orgánů státní správy České socialistické republiky. *Zákony pro lidi* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.zakonyprolidi.cz/cs/1969-2>

zájem a pracovat s veřejnými penězi. Ústavní soud nálezem IV. ÚS 1146/16 rozlišuje další ze znaků veřejných a soukromých institucí následovně⁵⁸:

„a) způsob vzniku či zániku instituce (z pohledu přítomnosti či nepřítomnosti soukromoprávního úkonu), b) hledisko osoby zřizovatele (z pohledu toho, zda je zřizovatelem instituce jako takové stát, či nikoliv; pokud ano, jedná se o znak vlastní veřejné instituce), c) subjekt vytvářející jednotlivé orgány instituce (z toho pohledu, zda dochází ke kreaci orgánů státem, či nikoliv; jestliže ano, jde o charakteristický rys pro veřejnou instituci), d) existence či neexistence státního dohledu nad činností instituce (existence státního dohledu je typická pro veřejnou instituci), a e) veřejný nebo soukromý účel instituce (veřejný účel je typickým znakem veřejné instituce).“⁵⁹

Důležitým bodem informačního zákona je § 2 odst. 4, který stanovuje, že má žadatel právo pouze na informace, které má instituce k dispozici. Pokud instituce danou informaci ke své činnosti sbírat nepotřebuje, není nucena kvůli žádosti informace zjišťovat a nevníká nárok veřejnosti na tuto informaci.

Zákon rovněž poskytuje výčet informací, které institucí nemusí být poskytnuty. Zejména *„utajované informace (§ 7), obchodní tajemství (§ 9), důvěrnosti majetkových poměrů (§ 10) a další omezení (§ 11) jako neposkytnutí informací o probíhajícím trestním řízení (§ 11, 4 (a)) či informaci, která je přípravě, průběhu a projednávání výsledků kontrol v orgánech Nejvyššího kontrolního úřadu (§ 11, 4 (d)).“⁶⁰*

Problematickým je pro novináře § 8b, který zohledňuje příjemce veřejných prostředků. Dle paragrafu § 8b *„(1) Povinný subjekt poskytne základní osobní údaje o osobě, které poskytl veřejné prostředky. (2) Ustanovení (...) se nevztahuje na poskytování veřejných prostředků podle zákonů v oblasti sociální, poskytování zdravotních služeb, hmotného zabezpečení v nezaměstnanosti (...) (3) Základní osobní údaje (...) se poskytnou pouze v*

⁵⁸ Přičemž platí, že zařazení subjektu do jedné z kategorií vyplývá z převahy znaků typických pro danou kategorii

⁵⁹ IV.ÚS 1146/16 ze dne 20. 6. 2017: N 101/85 SbNU 679. *Ústavní soud* [online]. 2017 [cit. 2021-11-02]. Dostupné z: https://nalus.usoud.cz/Search/GetText.aspx?sz=4-1146-16_2

⁶⁰ Viz. celé znění zákona v příloze diplomové práce

rozsahu: jméno, příjmení, rok narození, obec, kde má příjemce trvalý pobyt, výše, účel a podmínky poskytnutých veřejných prostředků.“⁶¹

Datový novinář Českého rozhlasu Jan Cibulka v tomto ohledu poznamenává,⁶² že není výjimkou zveřejňovat platy zaměstnanců výše vyčtených orgánů státní správy, hraniční problematika je ale s informacemi o úřednících veřejných institucí. Ústavní soud v tomto ohledu rozhodl, že k získání informace je nutné ze strany žadatele doložit zdůvodnění, proč danou informaci požaduje a zda sleduje veřejný zájem.⁶³ Tento požadavek dle Cibulky zamezuje sběru informací ze strany novinářů, neboť může vést k utajování informací ze strany úřadů v případě, kdy novinář získáním informace sleduje konkrétní kauzy.

1.7.2. Formální náležitosti podání žádosti a stanovené lhůty

Pro podání žádosti o informace není nutné doložit žádné oficiální doklady, lze ji podat písemně, ústně i elektronicky. Žádost lze podat prostřednictvím datové schránky nebo elektronickou podatelnou, a to bez poplatku. Oslovené úřady ale mohou po žadateli vyžadovat úhradu nákladů, které vznikly v souvislosti s vyhledáváním a poskytováním informací, a to v úměrné výši.

Žádost podle Ministerstva vnitra musí obsahovat:⁶⁴

- Identifikaci instituce, které je určena
- Identifikaci žadatele: jméno, příjmení, adresa bydliště, datum narození
- Prohlášení, že se žadatel domáhá poskytnutí informace podle zákona č. 106/1999 Sb.
- Znění požadované informace

Žádost ve valné většině případů naopak nemusí obsahovat vysvětlení, proč požadují danou informaci, a za jakým účelem ji chci získat.⁶⁵ Instituce má povinnost na žádost odpovědět

⁶¹ Viz. celé znění zákona v příloze diplomové práce

⁶² Dostupné z rozhovoru s Janem Cibulkou, příloha č. 9

⁶³ Další nález Ústavního soudu k poskytování informací o platech zaměstnanců. *Ústavní soud* [online]. 2018, 2018 [cit. 2021-11-02]. Dostupné z: <https://www.usoud.cz/aktualne/dalsi-nalez-ustavniho-soudu-k-poskytovani-informaci-o-platech-zamestnancu>

⁶⁴ Jak žádat o informace. *Ministerstvo vnitra České republiky* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.mvcr.cz/clanek/jak-zadat-o-informace-453908.aspx>

⁶⁵KUŽÍLEK, Oldřich, 2004

do patnácti dní, lhůty lze o deset dní prodloužit. V případě nevyhovění žádosti má úřad povinnost vystavit rozhodnutí. Pokud úřad na žádost neodpoví, lze podat stížnost na postup při vyřizování žádosti o informace podle § 16a informačního zákona. V případě odmítnutí poskytnutí informace lze podat odvolání. V krajních případech lze rozhodnutí napadnout u soudu, a to v prvním stupni u krajského soudu, pod jehož obvod spadá sídlo úřadu, který žádost zamítnul.⁶⁶

1.7.3. Příklady využití informačního zákona v českém mediálním prostředí

Praktických příkladů využití informačního zákona v tuzemských redakcích v posledních letech můžeme nalézt mnoho. Měříme-li úspěšnost žádosti dosažením cíle, tedy poskytnutím informace, mezi pravidelně získávané informace spadají již zmíněné odměny státních úředníků. Úřady ale mnohdy poskytují data, ač s konkrétními částkami, ale v anonymizované podobě. Novináři tedy musí dovozovat, o kterého úředníka se jedná.⁶⁷ Odlišná je také ochota jednotlivých ministerstev údaje na žádost poskytovat.⁶⁸

Využití informačního zákona v žurnalistice se blíže věnovala Koktová (2015), která oslovila novináře Tomáše Němečka, Vojtěcha Blažka a Radka Kedroně, kteří s podáváním žádostí měli nejvíce zkušeností. Z hloubkových rozhovorů s novináři identifikovala nejčastější důvody pro podání žádostí o informace:⁶⁹

- Neochota úřadu data poskytnout neoficiální cestou
- Předchozí zkušenost s úřady, které požadují formální podání žádosti o informace
- Předjímání neposkytnutí informace
- Snaha informovat o neochotě úřadu data poskytnout, spíše než získat konkrétní informace

⁶⁶ KUŽÍLEK, Oldřich, 2004

⁶⁷ KUBANT, Vít. Náměstci dostali za podzimní boj s covidem i čtvrt milionu. Lékaři a zdravotníci z první linie zatím nic. *IROZHLAS.cz* [online]. 2021, 08.02.2021 [cit. 2021-11-02]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/odmeny-ministerstvo-zdravotnictvi-penize-lekari-hygienici-uzis_2102080500_vtk

⁶⁸ BOČEK, Jan a Hana MAZANCOVÁ. VELKÝ PŘEHLED: Ministerské odměny za rok 2016. Kdo data tají a kdo bez problémů zveřejňuje? *IROZHLAS.cz* [online]. 2017, 07.02.2017 [cit. 2021-11-02]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/velky-prehled-ministerske-odmeny-za-rok-2016-kdo-data-taji-a-kdo-bez-problemu-zverejnuje-_201702070600_hmazancova

⁶⁹ KOKTOVÁ, Alena. *Svobodný přístup k informacím z pohledu novinářů* [online]. Praha, 2015 [cit. 2021-11-92]. Dostupné z:

https://dspace.cuni.cz/bitstream/handle/20.500.11956/69815/BPTX_2013_2_11230_0_388677_0_150987.pdf?sequence=1&isAllowed=y. Bakalářská práce. Univerzita Karlova. Vedoucí práce Mgr. Peter Hanák.

Kedroň, který měl s informačním zákonem hojně zkušenosti jako novinář, nastoupil v roce 2016 na post šéfredaktora zpravodajského serveru Českého rozhlasu iROZHLAS.cz. Ačkoliv získávání dat skrz informační zákon není výsadou datových novinářů, právě veřejnoprávní rozhlas disponuje samostatným datovým oddělením a v oblasti získávání dat skrz informační zákon se v současnosti v mediálním prostoru řadí k nejvíce aktivním.

Přelomovým pro novinářskou obec je právní spor datového novináře iROZHLAS.cz Jana Cibulky, který v roce 2016 žádal Policii České republiky o data sledující kriminalitu. Policie poskytnutí dat oddalovala a požadovala za jejich zpracování 26 milionů korun. Cibulka se pod záštitou Českého rozhlasu rozhodl o data soudit. Spor Cibulka vyhrál po pěti letech, a to rozhodnutím Ústavního soudu z roku 2021, který rozhodl, že poplatek je v neúnosné výši a brání zveřejnění údajů.⁷⁰ Během sporu Policie České republiky vytvořila online databázi kriminality na serveru kriminalita.policie.cz, a tudíž zveřejnila volně k použití údaje, které Cibulka požadoval.⁷¹

Podstatný je také spor Cibulky s Českým hydrometeorologickým ústavem, který začal v roce 2015 podáním žádosti o informace naměřených na jednotlivých meteorologických stanicích. V roce 2020 ústav data začal zveřejňovat, a to po rozsudku Nejvyššího správního soudu.⁷² Cibulka důvody toho, proč se o data rozhodl soudit i přesto, že s pokračujícím sporem nemusí zůstat relevantní, komentoval: „*Věděli jsme, že klimatická změna bude probíhat a rozhodli jsme se, že o informace budeme mít zájem i v budoucnu. Prošli jsme soudními spory, Nejvyšší správní soud nám nakonec dal za pravdu a data nyní mohou využívat také jiní novináři.*“⁷³

Novináři se k podávání žádostí o informace neubírají příliš často, a to především z důvodu značné časové investice. V metodologické části práce autorka nabídne detailnější pohled na využitelnost zákona v praxi.

⁷⁰ KUBIŠTOVÁ, Dominika. Ústavní soud dal datovému novináři za pravdu. Policie nemůže chtít miliony za informace o kriminalitě. *IROZHLAS.cz* [online]. 2021, 08.07.2021 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/policie-data-o-kriminalite-ustavni-soud-novinar-nalez_2107080924_dok

⁷¹ Dostupné z: <https://kriminalita.policie.cz/>

⁷² VINKLOVÁ, Julie. Počasí, věc veřejná. Meteorologové po letech sporů začali poskytovat historické záznamy. *IROZHLAS.cz* [online]. 2020, 26.03.2020 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/chmu-pocasi-data-historicke-zaznamy-o-pocasi-pravo-na-informace_2003260832_elev

⁷³ Viz. příloha č.9

1.7.4. Vizualizace dat získaných podle informačního zákona

Získání informací může vést nejen k výstupům investigativní povahy, praktickým využitím dat je vytvoření interaktivních aplikací či map, které mohou čtenáři využít v běžném životě. Takovým příkladem je vizualizace dat získaná datovými novináři Českého rozhlasu, která na mapě České republiky vyznačuje místa s výskytem ilegálních poutačů u silnic a dálnic. Data novináři získali na základě informačního zákona od Ředitelství silnic a dálnic.

Příloha č.1: Interaktivní mapa: vizualizace ilegálních poutačů u silnic (obrázek)



Autoři: Cibulka a Krawiecová (2020), zdroj: https://www.irozhlas.cz/zpravy-domov/billboardy-reklama-poutace-dalnice-dl-ministerstvo-dopravy-reditelstvi-silnic-a_2003020600_nkr

1.8. Evropská legislativa

Otevřená data jsou zakotvena také v evropské legislativě, a to „směrnici o otevřených datech a opakovaném použití informací veřejného sektoru v Úředním věstníku Evropské unie pod číslem 2019/2021.“ S účinností od roku 2021 směrnice předkládá „změny související s rychlým vývojem technologií, reaguje na požadavek poskytování dat v

reálném čase a současně reaguje na překážky při opakovaném použití informací veřejného sektoru.“⁷⁴

Evropská unie rovněž představila seznam šesti okruhů „datových sad s vysokou hodnotou“ (z anglického „*High Value Dataset*“, HVD), které se propisáním do evropské legislativy bude snažit zveřejnit unifikovaně napříč členskými státy unie. Ve fázi identifikace datových sad, které mají z hlediska otevřenosti společnosti a ekonomického rozvoje vysokou hodnotu, se ale názory představitelů členských států mohou lišit. Mezi důvody pro zveřejnění těchto konkrétních okruhů dat unie uvádí jejich „*potenciál pro využití evropskou společností, rozvoj hospodářství a výzkumu, či informovanější rozhodování a vývoj nových datových aplikací.*“⁷⁵

Okruhy data setů s vysokou hodnotou (HVD) podle evropské legislativy zahrnují:⁷⁶

- 1) Data geoprostorová
- 2) Data environmentální
- 3) Data meteorologická
- 4) Data statistická
- 5) Data o vlastnictví firem
- 6) Data o mobilitě, názvy ulic

Snahy o identifikaci a výběr data setů s vysokou hodnotou (HVD) v České republice již započaly, a to na úrovni jednotlivých ministerstev. Uživatelé, kteří opakovaně s otevřenými daty pracují, mají od Ministerstva vnitra k dispozici tzv. Open Data wish-list, tedy otevřený seznam přání, kam mohou vepisovat požadavky na data, která chtějí otevřít. Ministerstvo seznam systematicky prochází a poskytuje žadatelům zpětnou vazbu.

⁷⁴ Výroční zpráva o stavu otevřených dat v České republice za rok 2019. In: *Otevřená data: Odbor Hlavního architekta eGovernmentu* [online]. Praha: Ministerstvo vnitra, 2020, 16.06.2020, s. 48 [cit. 2021-11-02]. Dostupné z: <https://data.gov.cz/attachments/2020-07-14/V%C3%BDro%C4%8Dn%C3%AD%20zpr%C3%A1va%20o%20stavu%20otev%C5%99en%C3%BDch%20dat%20v%20%C4%8Cesk%C3%A9%20republice%20za%20rok%202019.pdf>

⁷⁵ HUYER, Ester, BLANK, Marit, European Data Portal Publications Office. Analytical Report 15: High-value datasets: understanding the perspective of data providers [online]. Luxembourg: Publications Office of the European Union, 2020, 2020, 21 [cit. 2021-11-29]. ISSN 2600-0601. Dostupné z: doi:10.2830/363773

⁷⁶ Tamtéž

1.8.1. Srovnání ČR v rámci unie

Evropská komise sleduje vývoj v oblasti otevřených dat napříč členskými státy v každoročním reportu.⁷⁷ V roce 2020 se Česká republika umístila na 21.místě z 28 členských států a se skóre 72 procent se zařadila mezi „followers“ tedy následovatele trendů v oblasti otevírání dat.⁷⁸ Skóre reflektuje četnost, ale také kvalitu poskytovaných dat státní správou a nabývá hodnot do 100 procent. V čele tabulky s hodnotami nad 90 procent se umístily státy: Dánsko, Španělsko, Francie, Irsko, Estonsko, a sousední Polsko s Rakouskem a řadí se tedy k „trend-setters“ tj. nastavují v oblasti otevírání dat trendy. V kategorii „beginners“, tj. začátečníků, můžeme nalézt osm členských států.⁷⁹

Report obsahuje sadu doporučení pro jednotlivé kategorie států, pro „následovatele trendů“ mezi deseti body zmiňuje například „zodpovědné analyzování chování učitelů na datových portálech, včetně zajištění jejich soukromí.“ Zaobírá se ale také identifikováním komunit opakovaných uživatelů dat, mezi něž se řadí datoví žurnalisté či výzkumníci, a vyzývá státy k aktivní komunikaci a jiným aktivitám zvyšující povědomí o otevřených datech v rámci těchto komunit.

Rok 2020 byl dle reportu klíčovým pro další rozvoj otevřeného vládnutí, ačkoliv pouze 44 procent členských států uvedlo „vysoký dopad“ otevřených dat na rozvoj ekonomiky v zemi. Státy Rakousko, Německo a Česká republika deklarovaly, že v souvislosti s pandemií COVID-19 zaznamenaly nárůst dostupnosti a opětovného využití otevřených dat.⁸⁰ Tato situace jistě mohla vést také k nárůstu výstupů datové žurnalistiky.

1.9. Kvantitativní analýza textů datové žurnalistiky v českých médiích

Pro účely práce je nezbytné za použití kvalitativních metod získat ucelený přehled o dostupnosti datových sad a práce s daty v českém mediálním prostoru. Relevanci zkoumanému tématu však může přinést kvantitativní analýza textů datové žurnalistiky, která zodpoví, zda je tento směr v českých médiích rozšířený. Kvantitativní analýza nebude zaměřena na obsahy textů a použitých datových zdrojů, nýbrž jejich četnost, a proto není pro účely této práce klíčová.

⁷⁷ *Open Data Maturity: Report 2020* [online]. Luxembourg: Publications Office of the European Union, 2020 [cit. 2021-11-20]. ISSN 2600-0512. Dostupné z:

https://data.europa.eu/sites/default/files/edp_landscaping_insight_report_n6_2020.pdf

⁷⁸ Do stejné kategorie se řadí: Rumunsko, Lucembursko, Belgie, Velká Británie a Švýcarsko

⁷⁹ Slovensko, Portugalsko, Maďarsko, Malta a státy, které oficiálně nezapadají do kategorie členských států EU: Norsko, Ázerbájdžán. Lichtenštejnsko a Gruzie

⁸⁰ Tamtéž, str. 2-92

Data získaná za pomoci kvantitativní analýzy podle Sedlákové (2014) mají standardizovaný charakter a reflektují frekvenci výskytu určitých témat, pro účely diplomové práce frekvenci výskytu textů nesoucích znaky datové žurnalistiky. Autorka práce texty posuzovala individuálně, a pro zařazení do analýzy musely nést následující znaky: novinář autorsky pracuje s daty, ať veřejnými či získanými za pomoci informačního zákona, a zároveň je graficky zpracovává, a to nad rámec oficiálních výstupů poskytovatele daných dat. Do kvantitativní analýzy byly zařazeny texty stálých autorů redakcí SAMIZDAT pro server iROZHLAS.cz, regionálního Deníku a Seznam Zpráv.cz. Blíže se redakcím autorka věnuje v metodologické části práce. Zkoumané období publikovaných textů osciluje na základě délky fungování jednotlivých redakcí.

1.9.1. Výsledky kvantitativní analýzy

Redakce SAMIZDAT má mezi porovnávanými nejdelší tradici, nejvíce publikovaných textů a nejpočetnější sestavu autorů. V rubrice Datová žurnalistika⁸¹ byl na serveru iROZHLAS.cz první text publikován 06.10.2014 a ve zkoumaném období, tj. do 20.12.2021 jich v rubrice nalezneme **784**. Mezi opakujícími se autory článků je Jan Boček, Jan Cibulka, Petr Kočí a Nela Krawiecová. V minulosti v rubrice publikoval texty také Marcel Šulek a Michal Zlatkovský.

Na zpravodajském serveru Deník.cz mají datové analýzy samostatnou rubriku pod názvem Deník data.⁸² Jediným autorem publikovaných textů je Aleš Vojíš, který je zároveň jediným stálým členem datové redakce. V časovém rozmezí od 12.09.2016 do 20.12.2021 publikoval v rubrice **433** článků, které nesou rysy datové žurnalistiky.

Server Seznam Zprávy.cz většinu textů spadajících do kategorie datové žurnalistiky publikuje v rubrice Fakta.⁸³ První text byl v rubrice publikován 22.11.2019 autory Jiří Burýšek a Petr Juna, kteří se řadí k nejčastějším přispěvatelům rubriky. Je nutné zmínit, že do rubriky Fakta patří také analytické texty věnující se vyvracení dezinformací. Objevují se také analytické texty autorů z redakce SZ Byznys,⁸⁴ ekonomické části redakce Seznam Zpráv.cz.

⁸¹ Dostupné z: <https://www.irozhlas.cz/zpravy-tag/datova-zurnalistika>

⁸² Dostupné z: <https://data.denik.cz/data/>.

⁸³ Dostupné z: <https://www.seznamzpravy.cz/sekce/fakta-205>

⁸⁴ Tyto texty reportují o ekonomických tématech za použití analýz, autoři: Eva Moniová, Michal Strnad, Karolína Štuková

Nejedná se však o analýzy za použití dat, a proto byl každý uveřejněný článek v rubrice zkoumán individuálně. V rubrice Fakta bylo identifikováno **368** textů nesoucí rysy datové žurnalistiky, a to v časovém rozmezí od 22.11.2019 do 20.12.2021. S příchodem datové novinářky Kateřiny Mahdalové v září 2020 došlo v redakci k výraznějšímu vyhranění datové žurnalistiky. Zatímco autoři Juna a Burýšek přispívají i nedatovými texty, Mahdalová se specializuje na datovou žurnalistiku a během let 2020 až 2021 publikovala na webu 106 datových textů.

Datová žurnalistika je v českých médiích okrajovým směrem, v posledních letech se nicméně dostává od původní sestavy SAMIZDAT i do jiných redakcí. Z kvantitativní analýzy je patrné, že ačkoliv je redakce SAMIZDAT nejpočetnější, v absolutních počtech výstupů ji kolegové z jiných médií konkurují. Na místě je tedy porovnat rutiny práce jednotlivých redaktorů. Analýza publikovaných textů autorce práce poskytla mnohé další podklady pro kvalitativní šetření, kterému se věnuje v následující metodologické části.

2. Metodologie

Jak již bylo zmíněno, autorka práce přistoupila k triangulaci metod kvantitativních i kvalitativních. Zda je dostupnost, kvalita či kvantita zveřejňovaných dat pro praktické využití v české žurnalistice dostatečná mohou nejlépe hodnotit aktéři, kteří data zpracovávají. Během pandemie COVID-19 statistiky a vizualizace pronikly do médií, která s datovou žurnalistikou žádné předchozí zkušenosti nemají. Je tedy nutné identifikovat novináře, kteří data zpracovávají systematicky a v rámci redakcí zaujímají post datových žurnalistů.

Jelikož jsou hlavními cíli práce poskytnout ucelený přehled o využitelnosti dat a nastínit postupy získávání dat neveřejných, a to na konkrétních případech, byly zvoleny metody kvalitativní: hloubkové rozhovory a případová studie doplněná o analýzu dostupnosti vybraného datového zdroje. Zahraniční výzkumy zaměřují svou pozornost na

limity datové žurnalistiky, proto se i tato práce pokusí v praktické části představit konkrétní překážky při získávání a zpracovávání dat pro účely zpravodajství.

2.1. Kvalitativní výzkum

Hlavním benefitem kvalitativního výzkumu je schopnost přinést „*kulturně specifická a kontextově bohatá data*.“⁸⁵ Plán kvalitativního výzkumu má pružný charakter, podle Hendla (2005) se v čase rozvíjí, přizpůsobuje okolnostem a získaným výsledkům.⁸⁶ Na začátku výzkumu ale stojí téma a základní výzkumné otázky. Klíčový je delší a „*intenzivnější kontakt s terénem nebo situací jedince nebo skupiny. Situace jsou obvykle banální, reflektující každodennost*.“ Cílem je pochopit chování respondentů výzkumu a okolnosti, které jejich běžné interakce doprovázejí.⁸⁷

Sedláková (2014) zdůrazňuje, že v rámci kvalitativního výzkumu je nutné „*nahlížet na sociální realitu z pohledu samotných aktérů, v kontextu jejich života, jejich vlastním jazykem a v předem nestrukturovaných interpretacích*.“ Vhodné je zvolit uvolněnější metody sběru dat, mezi které se řadí například hloubkový rozhovor.⁸⁸

Autorka diplomové práce považuje za nutné zopakovat, že má s prací v redakci datové žurnalistiky letité zkušenosti, a proto ve fázi přípravy materiálů, formulace výzkumných otázek a výběru zkoumaného vzorku využila poznatků získaných metodou zúčastněného pozorování. Švaříček a Šedřová (2014) zúčastněné pozorování vymezují jako interakci mezi výzkumníkem a pozorovanými účastníky, při které badatel nezasahuje do práce účastníků a přináší jiný typ informací než hloubkový rozhovor. Kvalitativní metody sběru dat jsou oproti kvantitativním metodám benevolentnější, i přesto se autorka pokusila

⁸⁵ ŠVAŘÍČEK, Roman a Klára ŠEĎOVÁ. Kvalitativní výzkum v pedagogických vědách. Vyd. 2. Praha: Portál, 2014. ISBN 978-80-262-0644-6.

⁸⁶ HENDL, Jan. *Kvalitativní výzkum: Základy metody a aplikace* [online]. Praha: Portál, 2005, str.63 [cit. 2021-11-12]. ISBN 80-7367-040-2. Dostupné z: https://dl1.cuni.cz/pluginfile.php/882806/mod_resource/content/1/06%20Hendl%2C%20J.%20-%20Kvalitativn%C3%AD%20v%C3%BDzkum%3B%2024-63.pdf

⁸⁷ Tamtéž, str.50

⁸⁸ SEDLÁKOVÁ, Renáta. *Výzkum médií: Nejužívanější metody a techniky* [online]. Praha: Grada Publishing, 2014, str.51 [cit. 2021-11-02]. ISBN 978-80-247-3568-9. Dostupné z: <https://books.google.cz/books?id=710nBgAAQBAJ&printsec=frontcover&key=AIzaSyDIPfI89JdFhWBVsMVsavVo6aNh057xITc#v=onepage&q&f=false>

výzkum a hloubkové rozhovory vést tak, aby její zkušenosti neovlivňovaly výsledné závěry a zaručuje objektivitu při analýze získaných výsledků.

2.1.1. Hloubkový rozhovor

Primární výzkumnou metodou práce je hloubkový rozhovor, který dle Sedlákové (2014) generuje data, ne sbírá. Směřování hloubkového rozhovoru, oproti dotazníkům kvantitativního výzkumu, je možné v jeho průběhu měnit a otázky modifikovat na základě znalostního zaměření respondentů, čehož autorka práce využila.

Hloubkový rozhovor Švaříček a Šed'ová (2014) dělí na dva hlavní typy: polo strukturovaný rozhovor, který je veden dle předem stanoveného seznamu otázek, a rozhovor nestrukturovaný. Pro účely této práce byla zvolena první varianta, otázky autorka práce rozdělila podle Sedlákové (2014) na základní výzkumné otázky a otázky od nich odvozené, které „*dále rozvíjejí a specifikují oblast zájmu.*“

Otázky podle Kerlingera (1972) splňují požadavky: „*relevance, cílenosti na žádanou informaci, etické a emocionální přijatelnosti, nesugestivnosti a tázání se pouze na informace, které respondent zná.*“⁸⁹

Při vedení hloubkových rozhovorů se autorka práce snažila vyvarovat chyb, na které upozorňuje Sedláková (2014): přehršel otázek, nedostatečný čas na odpovědi respondenta, zaměření se na okrajové drobnosti, příliš velká aktivita tazatele: „*informátor se nesmí v rozhovoru cítit ani jako u výslechu, ani jako pod palbou otázek, ani mít dojem, že jej tazatel chytá za slovo nebo jej kontroluje a hodnotí, případně obrací již řečené proti němu.*“

Klíčové je tedy zvolit otázky, které jsou přímé, ale zároveň nechávají respondentovi dostatečný prostor odpověď rozvést do dalších okruhů.

2.1.1.1. Referenční studie – hloubkové rozhovory

Metodu hloubkových rozhovorů při zkoumání fenoménu datové žurnalistiky použili například Felle (2016) či Porlezza a Splendore (2019). Felle se v obecnější rovině zkoumání zabíral otázkami dopadů digitalizace a využití dat na tradiční roli žurnalistiky

⁸⁹ Tamtéž str. 221

jako hlídacího psa demokracie.⁹⁰ Autorka práce tuto myšlenku rozvádí a směřuje k tomu, zda je kontrola státního aparátu v éře datové žurnalistiky efektivnější.

Důležitost dostupnosti datových sad a jejich získávání mapují Porlezza a Splendore (2019) s důrazem, do jaké míry je datová žurnalistika určována těmito faktory a zda státní správa formuje podobu novodobé novinářiny. V hloubkových rozhovorech s patnácti italskými žurnalisty zkoumají, jaké strategie využívají ke získání nedostupných dat a jak jejich práci limituje tato nedostupnost.⁹¹

Rutiny práce zkoumají Lowrey, Broussard, Sherrill (2019), kteří se věnují otázkám: *Do jaké míry datoví novináři zpochybňují kvalitu a relevanci datových sad, které jim státní správa poskytuje? Jak novináři ověřují relevanci a kvalitu datových souborů, a pokud v nich nalézají chyby, reportují je poskytovateli dat? Jaké jsou hlavní výzvy, kterým čelí datoví novináři při získávání a používání datových souborů z vládních zdrojů a jak se odlišují přístupy jednotlivých úřadů?*⁹²

Autorka práce využila poznatků výše zmíněných studií a odborné literatury při formulaci otázek pro hloubkové rozhovory, a to s ohledem na menší rozšířenost datové žurnalistiky v českém mediálním prostředí oproti zahraničí.

2.1.1.2. Formulace otázek

Hloubkový rozhovor autorka práce rozdělila do třech sekcí. V první části respondenty vyzvala ke shrnutí jejich dosavadního žurnalistického působení a znalostí, které se s jejich prací pojí. Otázky tedy spadají do kategorie úvodních otázek, kterým předchází „*představení výzkumníka a jeho projektu, požádání o souhlas k participaci na výzkumu a požádání o souhlas rozhovor nahrávat.*“⁹³

Úvodní otázky

- 1) Jaká je délka Vaší novinářské praxe?
 - Studoval/a jste žurnalistiku, programování či jiný obor na vysoké škole?
- 2) Jaká je Vaše role v rámci redakce, ve které působíte?
 - Kolik novinářů v redakci působí?

⁹⁰ FELLE, 2016, str. 99

⁹¹ PORLEZZA a SPLENDORE, 2019, str. 14

⁹² LOWREY, Wilson, Ryan BROUSSARD a Lindsey A. SHERRILL. *Data journalism and black-boxed data sets* [online]. 2019 [cit. 2021-11-02]. Dostupné z: <https://journals.sagepub.com/doi/pdf/10.1177/0739532918814451>

⁹³ ŠVARÍČEK a ŠEĎOVÁ, 2014, str. 87

- Kolik novinářů je vyčleněno pro práci s daty?
 - Máte v týmu grafika/grafičku či programátora?
- 3) Jsou pro Vás pro zpracování dat důležité technologie?
- Umíte programovat?
 - Jak často a jakým způsobem adaptujete nové nástroje pro zpracování a vizualizaci dat?
- 4) Spolupracujete napříč redakcemi s ostatními datovými novináři?
- Sdílíte získané datové sady či chyby v datech?

Druhá část hloubkových rozhovorů byla zaměřena na práci s daty, především k vymezení frekventovaně využívaných datových zdrojů veřejné správy a k poskytnutí prostoru popsat zkušenosti se získáváním dat neveřejných.

Otázky zaměřené na práci s daty

- 1) Jaké datové zdroje pro svou práci nejčastěji používáte a jak je získáváte?
- Jakým tématům se nejčastěji věnujete?
- 2) Máte zkušenosti se žádostí o informace podle zákona č. 106/1999 Sb. o svobodném přístupu k informacím?
- 3) Znáte a používáte nástroje ministerstva vnitra Národní katalog otevřených dat (NKOD) a seznam Open data wish-list?

Třetí část rozhovoru datovým novinářům poskytuje prostor pro subjektivní hodnocení, zkoumá redakční rutiny a hledá limity jejich práce. V odborné literatuře hovoříme o hlavních otázkách, jejichž cílem je *„povzbudit jedince, aby hovořili o svých zážitcích, dojmech a přesvědčeních, spíše, než aby podávali „správné odpovědi“ v souladu s názorem vedení organizace nebo s odbornými publikacemi.“*⁹⁴

Hlavní otázky

- 1) Jak hodnotíte dostupnost a kvalitu otevřených dat?
- Vnímáte ze strany státní správy ochotu či neochotu zveřejňovat datové sady?
- 2) Ověřujete či zpochybňujete kvalitu dat, která pro svou práci využíváte?

⁹⁴ ŠVARÍČEK a ŠEĎOVÁ, 2014

- 3) Reportujete nedostatečnou kvalitu či chyby v datech?
 - Pokud ano, jaká je podle Vaší zkušenosti komunikace s poskytovateli dat?
- 4) Formuje podle Vás výběr a kvalita zveřejňovaných dat státní správnou datovou žurnalistiku?
 - Jsou témata, která chcete zpracovat, ale nejsou pro něj dostupná otevřená data?
- 5) Jaké jsou hlavní limity Vaší práce?
- 6) Vysvětlujete čtenářům metody práce použité při zpracování dat? Pokud ano, uveďte konkrétní příklad

2.1.2. Případová studie

Kvalitativní výzkum ze své podstaty nemá široký záběr, věnuje se pouze „jednomu objektu či výskytu jevu, případu (jedinci, sdělení, skupině, médiu, události).“⁹⁵ Sedláková (2014) dodává, že při jeho realizaci získané poznatky výzkumník musí vztahovat k základnímu vymezenému problému a ověřovat, zda jsou pro něj relevantní. „Vždy je přínosnější detailní analýza jednoho jevu, například případová studie než obecný, a proto často povrchní projekt se širokým záběrem.“⁹⁶ Cílem práce je na konkrétních případech ukázat získávání a práci s daty, v průběhu přípravy diplomové práce přitom došlo v oblasti otevírání dat k přelomovým událostem. Autorka práce se rozhodla identifikovat texty, které se procesu získávání dat během pandemie COVID-19 věnují, a formou případové studie v nich validovat poznatky získané v hloubkových rozhovorech. Případová studia byla zvolena neb se „nejedná o metodu zkoumání, ale o typ výzkumu, formu jeho provedení, jehož cílem je porozumění zkoumanému případu a jeho individuálnímu průběhu a specifickým.“⁹⁷

2.2. Výběr vzorku

Za účelem výzkumu bylo osloveno pět datových novinářů napříč třemi tuzemskými redakcemi. Jak již bylo zmíněno, za datové novináře autorka práce považuje redaktory, kteří se práci s daty věnují systematicky, primární data sami zpracovávají, vizualizují a dodávají datům kontext nad míru oficiálních výstupů. Práce s daty není výsadou žurnalistiky, nýbrž

⁹⁵ SEDLÁKOVÁ, 2014, str. 52

⁹⁶ SEDLÁKOVÁ, 2014, str. 51

⁹⁷ SEDLÁKOVÁ, 2014, str. 52

se objevuje také v oborech vědeckých. Česká žurnalistika disponuje erudovanými vědeckými redaktory, mezi něž se jistě řadí například Daniel Stach (Česká televize) či Petr Koubský (Deník N). Ačkoliv metody analýzy dat ve své práci využívají, autorka práce je nepovažuje za ryze datové novináře, a proto nebyli v rámci této práce osloveni.

Autorka práce si je vědoma limitů velikosti vybraného vzorku respondentů, v rámci tuzemského mediálního diskurzu ale oslovila všechny žurnalisty, kteří definici datového novináře naplňují. Respondenti rovněž potvrdili, že se datovou žurnalistikou nezabývají jiné redakce než ty oslovené. Je však nutné znovu připomenout, že práce s daty je součástí tradiční žurnalistiky, a proto se může nárazově objevovat napříč mediální produkcí.

2.2.1. Redakce SAMIZDAT

Průkopníky datové žurnalistiky v českých médiích je skupina SAMIZDAT ve složení novinářů Jan Boček, Petr Kočí a Jan Cibulka. Po prvních zkušenost v Hospodářských novinách, kde v roce 2012 založili první formální redakci datové žurnalistiky, přešel tým v roce 2014 do Českého rozhlasu.⁹⁸ Na zpravodajském serveru iROZHLAS.cz, doplněn o novinářku Nelu Krawiecovou, editora Michala Kašpárka a grafičku Kristínu Zákopčánovou, působí do současnosti. Datový tým je pravidelně nominován k získání Novinářských cen Nadace Open Society Fund. V letech 2013, 2015 a 2019 nominace proměnil v ocenění za první místa v kategorii Inovativní a online žurnalistika.⁹⁹ Datové výstupy a vizualizace novináři prezentují na serveru iROZHLAS.cz, který se umístil mezi deseti nejvíce

⁹⁸ ČTK. Tým datové žurnalistiky z Economie přešel k rozhlasu. *Mediář* [online]. 2014, 10.10.2014 [cit. 2021-11-07]. Dostupné z: <https://www.mediar.cz/tym-datove-zurnalistiky-z-economie-presel-k-rozhlasu/>

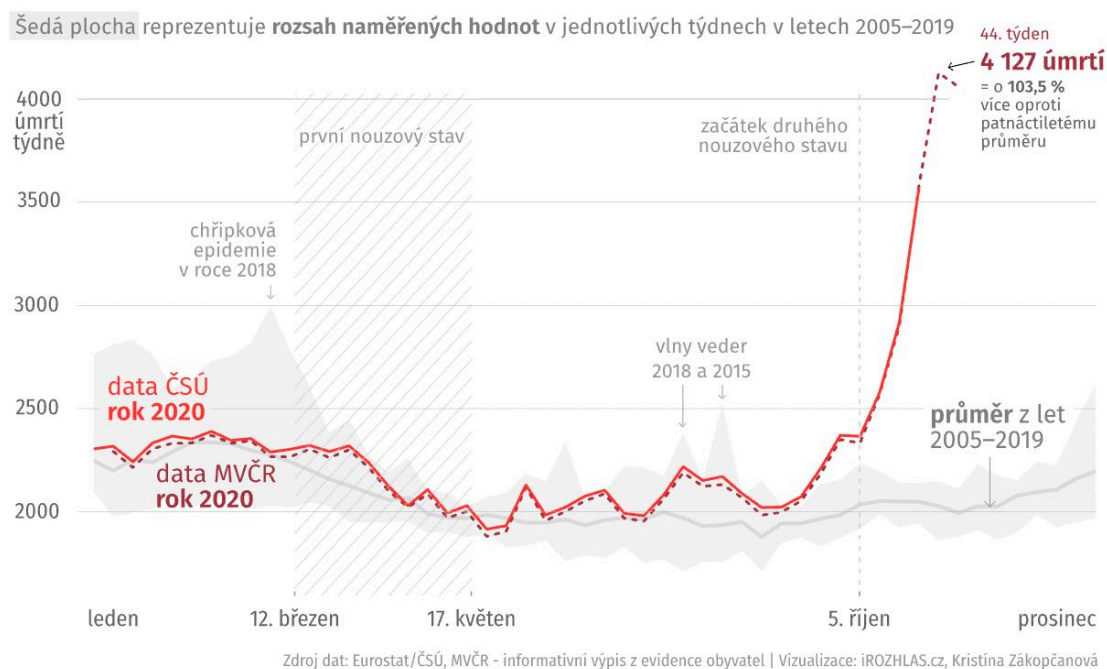
⁹⁹ IROZHLAS.CZ. Novinářská cena 2019: Dataři zvítězili s projektem Rozdělení svobodou, za Český rozhlas uspěl i Lukáš Houdek. *IROZHLAS.cz* [online]. 2020, 17.06.2020 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravodomov/novinarska-cena-vyhra-datova-zurnalistika-irozhlas_2006172015_ada

navštěvovanými zpravodajskými servery českých médií,¹⁰⁰ a také ve vysílání rozhlasových stanic Radiožurnál a Plus.

Příloha č.2: Ukázka datové žurnalistiky na iROZHLAS.cz (obrázek)

Koronavirus počtem obětí překonal velké chřipkové epidemie i vlny veder

Počty zemřelých po týdnech v České republice mezi lety 2005–2020



Autor: Kočí (2020), zdroj: https://www.irozhlas.cz/zpravy-domov/celkova-umrti-listopad_2012070635_pek

2.2.2. Redakce Deník

Od roku 2016 datovou žurnalistiku zařadila do svého portfolia také redakce regionálního Deníku, ve které působí jediný člen, Aleš Vojíř. Žurnalistice se Vojíř věnuje od roku 2000, v minulosti působil v Hospodářských novinách, Ekonomu a časopisu Týden.¹⁰¹ Jeho práce v Deníku spočívá v identifikaci otevřených dat, jejich zpracování, interpretaci a vizualizaci. Výstupy datové žurnalistiky prezentuje v tištěné verzi Deníku, v jeho on-line mutaci

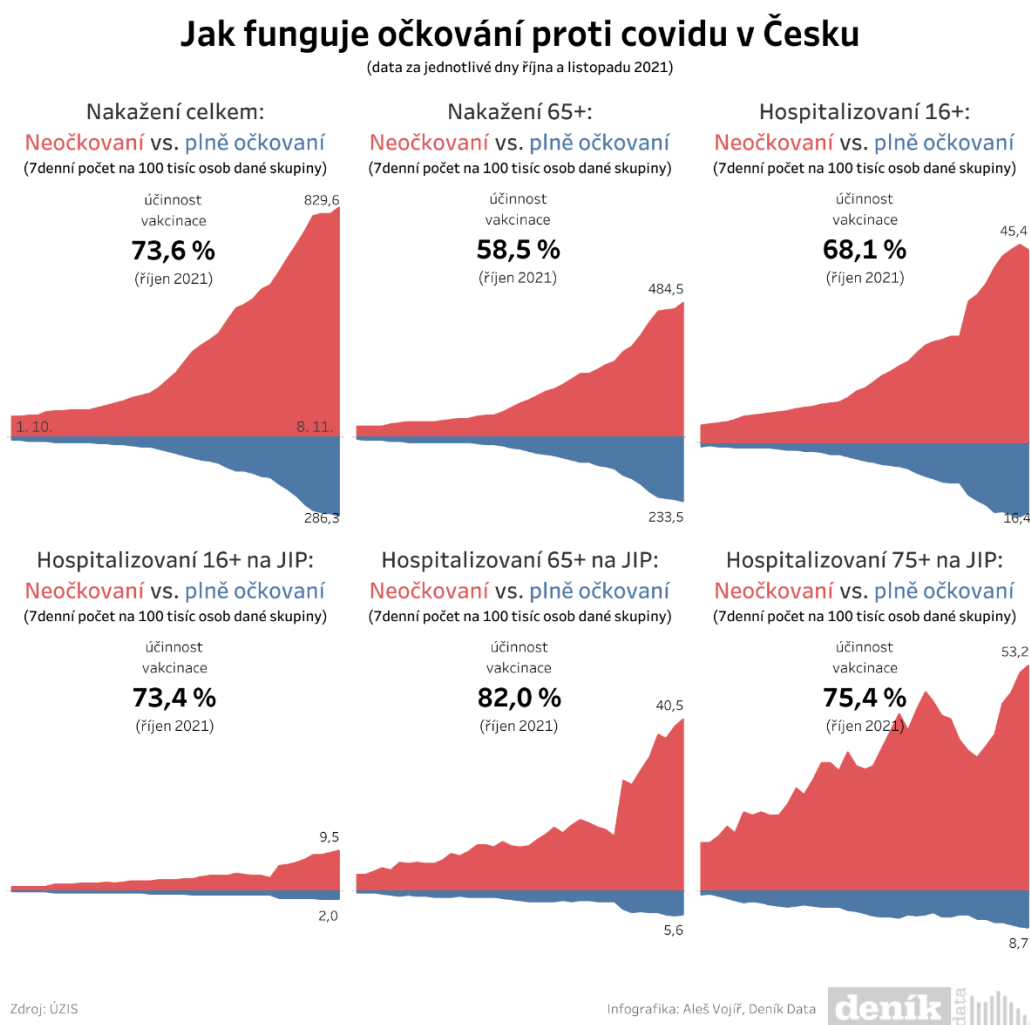
¹⁰⁰ IROZHLAS.cz se zařadil do první zpravodajské desítky: v březnu ho navštívily rekordní tři miliony čtenářů. *IROZHLAS.cz* [online]. 2020, 03.04.2020 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/irozhlas-server-zpravodajstvi-rekord-navstevnost-brezen_2004031320_tzr

¹⁰¹ Aleš Vojíř: vedoucí Deník Data. *Deník.cz: My jsme Deník* [online]. [cit. 2021-11-07]. Dostupné z: <https://www.denik.cz/my-jsme-denik/praha/ales-vojir-3115.html>

Deník.cz a ve vysílání zpravodajské televize CNN Prima News.¹⁰² Deník má 70 mutací na regionální i celostátní úrovni¹⁰³ a jeho vydavatelem je VLTAVA LABE MEDIA a.s.

Ve 2. a 3. čtvrtletí roku 2021 měl tištěný Deník podle dat výzkumu čtenosti tisku Unie vydavatelů MEDIA PROJEKT čtenost 432 000 a prodaný náklad činil 73 218 kusů.¹⁰⁴ Online platformu Deník.cz navštíví každý měsíc zhruba čtyři miliony reálných uživatelů.

Příloha č.3: Ukázka datové žurnalistiky na Deník.cz (obrázek)



Autor: Vojtěch (2020), zdroj: <https://data.denik.cz/cesi-v-cislech/ockovani-proti-koronaviru-funguje.html>

2.2.3. Redakce Seznam Zprávy.cz

Redakce Seznam Zpráv.cz je mezi zkoumaným vzorkem nejmladší, svůj provoz web internetového poskytovatele Seznam.cz zahájil v roce 2016 a poměrně rychle se zařadil mezi trojici nejnavštěvovanějších zpravodajských serverů českého internetu.¹⁰⁵ Ačkoliv redakce nemá specializovanou rubriku datové žurnalistiky, tento post od září 2020 vykonává Kateřina Mahdalová. Před angažmá v Seznam Zprávách.cz prošla redakcí Respektu, kde se věnovala investigativně analytické činnosti. Založila web datové a kontextové žurnalistiky, na kterém publikovala první vizualizace dat s přesahem do žurnalistiky.¹⁰⁶ Působila v České tiskové kanceláři, pro kterou navrhla univerzální šablony pro vizualizace informací. Na webu Seznam Zpráv.cz v rubrice Fakta s Mahdalovou působí také Petr Juna a v minulosti Jiří Burýšek. Externím spolupracovníkem redakce je Michal Škop, zakladatel občanského sdružení KohoVolit.eu. Mahdalová spolupracuje se zahraničními zpravodajskými servery a

¹⁰² Viz. příloha č. 11

¹⁰³ 4 důvody, proč číst Deník. *Vltava Labe Media: Deník* [online]. [cit. 2021-11-07]. Dostupné z: <https://www.vlmedia.cz/denik>

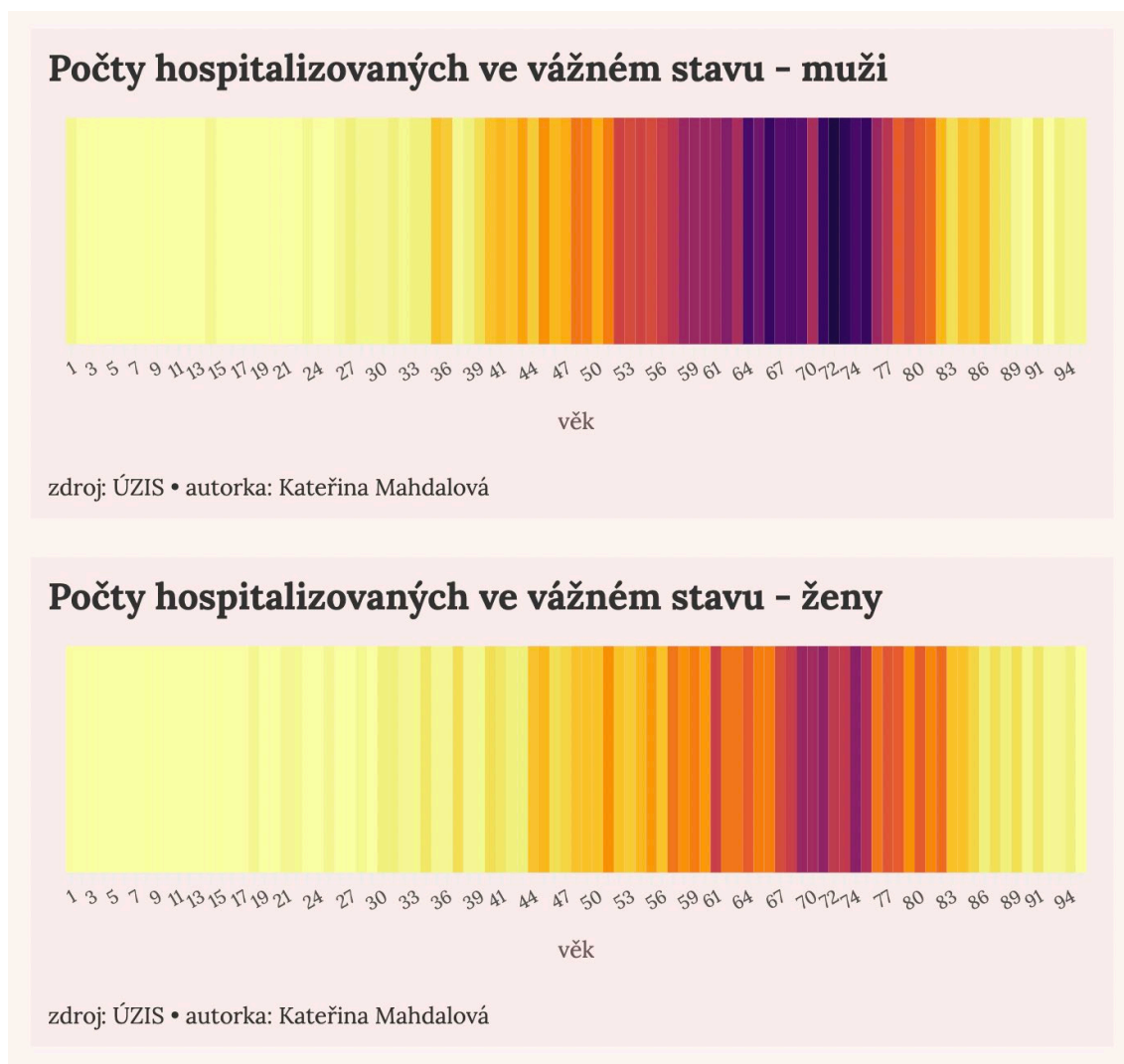
¹⁰⁴ VÝZKUM ČTENOSTI TISKU, MEDIA PROJEKT. *Unie vydavatelů* [online]. [cit. 2021-11-07]. Dostupné z: http://www.unievydavatelu.cz/cs/unie_vydavatelu/medialni_data/vyzkum_ctenosti

¹⁰⁵ Zpravodajské weby 2020: Na čele Novinky, v TOP 10 i nový web CNN Prima. *MediaGuru* [online]. 2021, 21.01.2021 [cit. 2021-11-07]. Dostupné z: <https://www.mediaguru.cz/clanky/2021/01/zpravodajske-weby-2020-na-cele-novinky-v-top-10-i-novy-web-cnn-prima/>

¹⁰⁶ Dostupné na webové adrese: www.datovazurnalisitka.cz

po dvou nominacích v letech 2014 a 2018 získala za rok 2020 Novinářskou cenu Nadace Open Society Fund v kategorii Inovativní a online žurnalistika.

Příloha č.4: Ukázka datové žurnalistiky na Seznam Zprávy.cz (obrázek)



Autor: Mahdalová (2020), zdroj: <https://www.seznamzpravy.cz/clanek/data-v-cesku-je-dvakrat-vice-pacientu-v-tezkem-stavu-nez-pocatkem-roku-146945>

2.3. Analýza rozhovorů

Rozhovory byly pořizeny v časovém rozmezí od 4.11.2021 do 26.11.2021 prostřednictvím online platformy Google Meet. Respondenti odpovídali samostatně, průměrná délka rozhovoru přesáhla 60 minut.

- Jan Boček, iROZHLAS.cz (datum pořízení rozhovoru: 4.11.2021)
- Jan Cibulka, iROZHLAS.cz (datum pořízení rozhovoru: 5.11.2021)
- Petr Kočí, iROZHLAS.cz (datum pořízení rozhovoru: 10.11.2021)
- Aleš Vojíš, Deník (datum pořízení rozhovoru: 11.11.2021)
- Kateřina Mahdalová, Seznam Zprávy.cz (datum pořízení rozhovoru: 26.11.2021)

2.3.1. Úvodní otázky

Studoval/a jste žurnalistiku, programování či jiný obor na vysoké škole?

Všichni oslovení respondenti deklarovali, že k technologiím měli vždy kladný vztah. Základy žurnalistiky si část z nich osvojila již během středoškolského studia, formální žurnalistické vzdělání získal Boček, Kočí a Vojíš. Dotazovaní Boček a Vojíš studovali programování jako obor na vysoké škole, zatímco Kočí studoval žurnalistiku souběžně s prací v Lidových novinách. „*Po deseti nebo dvanácti letech jsem začal mít pocit, že aby má práce měla smysl, měl bych se zaměřit na téma, kterému rozumím více než ostatní,*“ říká Kočí, který začátkem 21. století začal o technologiích psát články do týdeníku Respekt, navštěvoval konference zaměřené na IT, a v oboru se dále rozvíjel: „*Nejdříve jsem měl pocit, že mé téma je, jak technologie mění společnost. Postupem času mě ale čím dál více lákalo si webové technologie sám vyzkoušet,*“ říká.

Kateřina Mahdalová se věnovala studiu bohemistiky, v rámci výměnného pobytu ve Velké Británii prošla kurzem online žurnalistiky a zaujala ji data. „*V Anglii jsem před deseti lety zjistila, že existuje něco jako datová žurnalistika. Tehdy Guardian začínal s Datablogy. K datové žurnalistice jsem se tedy dostala oklikou,*“ uvedla Mahdalová.

Výjimkou je Cibulka, který na vysoké škole nevystudoval žurnalistiku ani programování, znalosti těchto oborů si ale osvojil již na gymnáziu: „*Programování jsem se začal učit na střední škole (...), k novinářské práci jsem se poprvé dostal také již na gymnáziu, a to v radničních novinách. Byl jsem na stáži v deníku Blesk nebo v Parlamentních listech a střídal jsem to s prací v IT firmě,*“ uvedl.

Jsou pro Vás pro zpracování dat důležité technologie?

Společným rysem pro celou skupinu respondentů je schopnost adaptovat nové technologie svépomocí. Od jednodušších programů, mezi které se řadí zpracování dat v nástroji Microsoft Office Excel či online platformy pro vizualizaci dat, se postupně

dostali k sofistikovanějším analytickým a vizualizačním nástrojům. Důvodem byly zpravidla limity, na které postupem času naráželi: „*Výrazný krok pro mě byly začátky datové žurnalistiky, kdy jsme využívali nástroj Google Sheet. Velmi rychle jsme narazili na různá omezení a já se rozhodl naučit programovací jazyk R. Základy jsem získal v online kurzu, spolu se základy statistiky,*“ uvedl Kočí, který společně s Bočkem programovací jazyk R určený pro statistickou analýzu využívá dodnes.

„*V poslední době se stává málo kdy, že bych potřeboval technologii nebo nástroj, který neznám. Nové nástroje se učím převážně tehdy, když chci něco ukázat svým studentům,*“ uvedl Boček, který vyučuje kurz Datové žurnalistiky na Masarykově univerzitě.

Cibulka tentýž předmět vyučuje na Univerzitě Karlově, mezi oslovenými je technologicky nejzdatnější, nástroje pro svou práci obměňuje pravidelně. Autorka práce má za to, že jako jediný naplňuje definici „hackerského novináře“. Cibulka o svých technologických znalostech říká: „*Je dobré sledovat, jak se vyvíjí nástroje, které se osvědčily. Kolega Marcel Šulek měl zásadu, že se učil jednu novou technologii za rok. Nemám to takto striktně, ale je pravda, že minimálně jednu věc by měl člověk pochopit a třeba ji posléze i opustit, pokud zjistí, že mu nevyhovuje,*“ dodal. Tým SAMIZDAT spoluvytvářel webové rozhraní serveru iROZHLAS.cz, který je proto schopen přizpůsobit se jejich vizualizačním požadavkům. Odlišují se tak od zbylých dvou redakcí, které spoléhají na veřejně dostupné vizualizační nástroje.

Mahdalovou toto nastavení limituje. Ačkoliv nové nástroje pro zpracování dat aktivně vyhledává, nemůže je použít. „*Ne všechny technologie mohu použít v redakčním systému. Proto i když některé nové technologie zkouším, nepovede se je použít,*“ dodává.

Obdobné zkušenosti deklaroval Vojíř, který se od volně dostupného nástroje Infogram přesunul k nástroji Tableau. „*Začínám narážet na limity, kdy se snažím, aby vše bylo precizní. Na doporučení Jana Cibulky jsem začal s programovacím jazykem Java Script, který by to mohl vyřešit. (...), Vidina je, že bych se od Tableau v budoucnu oprostil,*“ uvedl.

2.3.2. Otázky zaměřené na práci s daty

Jakým tématům se nejčastěji věnujete?

Kočí uvedl, že v posledních dvou letech silnou tematickou specializací neměl. Opakovala se u něj témata volební, pandemická a okrajově také ekonomické analýzy.

Silnou profilaci naopak deklaroval kolega Boček, který se dlouhodobě věnuje zdravotnictví a tématům s environmentálním přesahem: lesnictví, energetika a klimatická změna. „*Jsou to měkká témata, ale tvrdý, datový přístup. Témata jsou komplexní, nesnažím se je jen popisovat, ale pochopit, vyřešit a předat řešení dál,*“ popisuje. Boček vnímá podobnost své práce se žurnalistikou investigativní, okrajově zpracovává témata přerozdělování dotací či korupce. Autorka práce ve výstupech Bočka identifikuje také přesah do žurnalistiky řešení (z anglického „*solution journalism*“).

U Vojíře je výběr témat podmíněn regionálním charakterem Deníku a personálním omezením redakce, ve které je na zpracování dat sám. „*Věnuji se pouze tématům, která jsou přínosná pro naše čtenáře,*“ uvedl Vojíř. Publikum charakterizoval jako lidi v regionech, staršího či středního věku. Před pandemií koronaviru zpracovával data populační či socioekonomická: „*Vždy fungují důchody nebo inflace, tedy jak zdražují jednotlivé položky ve spotřebním koši,*“ popisuje. Dále se pravidelně věnuje tématům exekucí či dopravních nehod, z populačních dat sňatkům, rozvodům, či úmrtím. Společný ukazatel dat, které vyhledává, je regionální přesah. „*Jsou to témata, která se vždy dají najít v regionálním členění alespoň na úroveň krajů, pro nás optimálně na úroveň okresů, protože máme přes 70 mutací Deníku,*“ dodává Vojíř.

Cibulka připouští, že se jeho role v datovém týmu Českého rozhlasu v posledních letech odklonila od klasické analýzy dat k technologické žurnalistice a zajištění zpracování dat. „*Technologická znalost mě natlačila dělat technologickou žurnalistiku, která se v Česku redukovala na recenzi produktů, což je sice legitimní žánr, ale pro popis dopadu technologií na společnost nestačí,*“ popisuje Cibulka, který rovněž poskytuje edukaci v oblasti bezpečnostní žurnalistiky. I nadále ale zpracovává data o koronaviru a témata s přesahem do demografie. Podle svých slov majoritně využívá data veřejné správy, znám je pro právní spory o data, mimo jiné již zmíněná data kriminality Policie České republiky a data hydrometeorologická.

Mahdalová témata vyhledává v oblasti lidských práv a justice. „*Věnuji se také tématům školství, inflace, obsazenosti věznic, jsou to desítky a desítky věcí. Kdybych to měla zobecnit jsou to veřejná data státní správy,*“ dodává.

Jaké datové zdroje nejčastěji využíváte a jak data získáváte?

Práci všech respondentů silně ovlivnila pandemie koronaviru, mimo individuální tematická zaměření pracují s otevřenými daty na portále Ministerstva zdravotnictví

Onemocnění aktuálně.cz,¹⁰⁷ či jinými daty Ústavu zdravotnických informací a statistiky. Mezi frekventované zdroje dat dále uvádí Český statistický úřad, Českou správu sociálního zabezpečení, data ministerstev, či geoprostorová data Českého úřadu zeměměřického a katastrálního.

Máte zkušenosti se žádostí o informace podle zákona č. 106/1999 Sb. o svobodném přístupu k informacím?

Nejvíce zkušeností s podáním žádosti o informace má z dotazovaných Cibulka. Alespoň jednu žádost podali také Boček a Kočí. Naopak Mahdalová a Vojíř oficiální žádosti nepodávají. Dotazováni se shodli na značné časové investici, která je pro podání žádosti nutná. „*V rámci redakce je to tedy i investice finanční. Výsledek se může táhnout i měsíce,*“ uvedl Cibulka. Pokračuje: „*Vidíme opravdu velký odpor státní správy k tomuto zákonu, který nemizí. Představa byla, že časem bude odpor klesat, ale to se ani náznakem nestalo,*“ komentuje situaci ke sklonku roku 2021.

Limity využití informačního zákona v žurnalistice dále popisuje Boček, který vyzdvihuje potřebu data získat v reálném čase, nikoliv v nejasném časovém horizontu. „*Obzvláště v denním zpravodajství je obtížné vysvětlovat vedení redakce, proč zpracování jednoho tématu může trvat tak dlouho. Po podání žádosti o informace ale odpověď dostaneme za náhodnou dobu,*“ říká.

Kolega Kočí o informace žádal podle svých slov přibližně pětkrát, výsledky byly smíšené. „*Většinou se snažím vycházet z toho, co je k dispozici. Souvisí to také s tím, že jsem tematicky méně vyhraněný,*“ dodává.

Mahdalová do diskuse přidává zajímavou perspektivu, a to získávání dat pomocí informačního zákona jako práci vhodnou pro média veřejné služby. „*V komerčních médiích není na podávání žádostí prostor. Proto si myslím, že veřejnoprávní média jsou ta, která by primárně měla data získávat,*“ popisuje. Souhlasí také Vojíř, který zároveň oceňuje benefity otevírání a získávání dat touto cestou. „*Jsem „černý pasažér“ všech snah kolegů, například zmíněného Cibulky, protože dělá skvělou práci, na které vyděláváme my všichni.*“

¹⁰⁷ Dostupné na webové adrese: <https://onemocneni-aktualne.mzcr.cz/>

Znáte a používáte nástroje Ministerstva vnitra Národní katalog otevřených dat (NKOD) a seznam Open data wish-list?

Žádost o konkrétní data set prostřednictvím Open data wish-listu Ministerstva vnitra nevyužívá pravidelně žádný z dotázaných. Cibulka ale iniciativu hodnotí jako velmi potřebný způsob veřejné konzultace. „*Open data wish-list má tu výhodu, že ministerstvo seznam systematicky prochází a pokud něco zveřejnit nemůže, vysvětlí žadateli, z jakých důvodů,*“ říká Cibulka.

Spolupráci s Ministerstvem vnitra zmiňuje také Boček: „*Vím, že wish-list existuje. V minulosti jsem dával podněty na konkrétní datové sady přímo Michalu Kubáňovi, který působil jako koordinátor otevřených dat na Ministerstvu vnitra a spolupráce byla dobrá.*“ Boček za problematické v Národním katalogu otevřených dat i ve zmiňovaném „seznamu přání“ vnímá především chybějící užitečnost otevíraných dat. „*Pokud chcete zajímavá data otevřít, musíte mluvit s menšími úřady, organizacemi, vysvětlovat proč by měly data otvírat. Aby se do katalogu dostala užitečná data, vyžadovalo by to značné úsilí. Je k tomu zapotřebí lidí, kteří jsou aktivní, možná až aktivističtí, kteří se o jednotlivé data sety starají,*“ říká.

Vojír obě nabízené možnosti zná, ale nevyužívá je. „*V tématech se točíme, protože víme, co funguje. I když se jednou za čas podívám do národního katalogu, co je nového a říkám si, že je tam spousta zajímavých věcí, reálně je nevládám zpracovat.*“

2.3.3. Hlavní otázky

Vnímáte ze strany státní správy ochotu či neochotu zveřejňovat datové sady?

Ochota či neochota jednotlivých úřadu poskytovat data podle Kočího nelze zobecňovat, zatímco některé úřady otevírání dat pojaly jako užitečné PR, jiné zveřejňování informací nevnímají jako prioritu a vyřizování žádosti je pro ně komplikací.

Vyhraněnější postoj mají Cibulka s Bočkem, ti vnímají silnou neochotu, a především pak zhoršení transparentnosti v posledních letech, kdy podle nich došlo k výměně úředníků na nižších stupních administrativy a výrazné politizaci komunikace a poskytování informací novinářům. „*Zatímco předtím se vyměnila vláda, ministři, náměstci a občas šéfové odborů, tak teď došlo k obměně i na nižších stupních. Úředníci dostali jistou míru nezávislosti ale zároveň se „bojí“ svého politického pána a rozhodují ve prospěch politických linií,*“ říká Cibulka a pokračuje: „*Viděli jsme to během pandemie na*

Ministerstvu zdravotnictví, kdy o zveřejnění informací rozhodovali lidé, kteří by neměli mít rozhodovací pravomoc. (...) Zpolitizovaná státní správa a poplatnost vládnoucí garnituře je extrémní problém žurnalistiky,“ dodává.

Boček zmiňuje čím dál viditelnější byrokracii i na menších úřadech, kde jsou úředníci pod dohledem tiskových mluvčích, kteří ovlivňují, které informace blokovat. *„Nyní už to lze říct obecně: dříve se s úřady bavít a dostávat z nich data dalo, bylo to méně formální. V poslední době se ale bez informačního zákona člověk nedostane skoro k ničemu,“* říká. Roli podle něj hrají také osobní sympatie či antipatie úřadů k jednotlivým redaktorům. *„U spousty úřadů jsme v „nemilosti“ a to právě proto, že jsme se s nimi v minulosti o data soudili. To by ale nemělo vůbec hrát roli, úřad by neměl koukat na vaše články a podle toho, zda se mu líbí, co píšete, vám vycházet vstříc. Mám pocit, že za vlády Andreje Babiše se tento trend rozjel. Je to nepříjemné, všechny komunikace s jakýmkoliv úřadem jsou nepříjemné,“* uzavírá.

Mahdalová v tomto směru dodává, že pokud instituce vnímají užitečnost novinářů jejich data dále šířit, ochota je vysoká. *„Pokud je to něco, co má instituci konfrontovat, ochota je velmi nízká.“*

Jak hodnotíte dostupnost a kvalitu otevřených dat?

Kvalita a dostupnost dat se podle respondentů ve velké míře odlišuje napříč úřady veřejné správy. Nejvíce je s kvalitou a dostupností spokojen Vojtěch, který vychází ze zkušeností práce s daty slovenských úřadů, které hodnotí jako podstatně horší. *„Oproti Slovensku se máme úžasně. A to nejen proto, že máme ohromné množství data setů, se kterými lze pracovat. Překvapuje mě, že to není případ jen dat o pandemii, ale i u dat statistického úřadu, u kterých jsem si dříve myslel, že metodika bude jednotná,“* říká.

Kočí naopak veřejnou databázi Českého statistického úřadu považuje za nepřívětivou. *„Je dost náročné dokliktat se k tomu, co potřebuji, popřípadě to tam není. Na druhou stranu mají celé informační oddělení, které má v popisu práce lidem radit a pomáhat.“*

Mahdalová se podle svých slov setkává s daty velmi různého charakteru i spolehlivosti, jejich kvalita je kolísavá, protipóly až extrémní. *„Nemůžu ani říct, že je to například půl na půl. Setkávám se s mlčením, nepochopením toho, co chci, špatnými výstupy, anebo dostanu data v profesionální kvalitě, přesně vědí, co chci a nabízí něco*

navíc. Rozepře uvnitř některých úřadů navíc ovlivňovaly to, jestli mi poskytnou nebo neposkytnou data,“ popisuje.

V otázkách kvality je podle Cibulky podstatné zohledňovat meta data, tedy přesné popisy parametrů daných data setů a okolnosti jeho vzniku. Zároveň zmiňuje, že kvalitu dat ovlivňuje jejich oborový standart. Z tohoto důvodu například data geoprostorová hodnotí jako vysoce kvalitní.

Ověřujete či zpochybňujete kvalitu dat, která pro svou práci využíváte?

Ověřování kvality dat a hledání možných nedostatků jako součást zpracování informací praktikují všichni respondenti. Panovala shoda na explorační analýze, metodě známé ze sociologického výzkumu jako přístup vyhledávající vztahy v datech a formulování závěrů o datovém souboru. Explorační analýza podle Řeháka:

„Nevychází z předem striktně stanovených statistických hypotéz, ale je postavena na vhodné posloupnosti otázek, na něž jsou z dat hledány odpovědi. (...) Postup pokračuje tak dlouho, dokud není vytvořen konzistentní a pro cíle výzkumu zajímavý a důležitý závěr, nebo dokud není vyčerpána informační hodnota datové báze. (...) Výsledek má status hypotézy nebo podkladové informace.“¹⁰⁸

Kočí explorační analýzu popisuje jako rychlé vykreslení základních grafů, ve kterých hledá nepravděpodobné výkyvy nebo jevy, pro které nemá vysvětlení. *„Často dojdou k tomu, že je to například metodou sběru dat nebo se může jednat o technickou chybu.“* Shoduje se ale s Cibulkou, že jakmile data vypadají věrohodně, přesouvá pozornost k jejich dalšímu zpracování a vizualizaci. Nevylučuje, že ve velkém množství dat lze přehlédnout jednotlivosti. *„Samozřejmě, že když na první, druhý, třetí pohled data vypadají, že jsou v pořádku, tak může člověk přehlédnout jednotlivost (...), ale chce to alespoň zkontrolovat, zda sedí rozptýlení hodnot, vyjet si histogram,*“ říká Cibulka.

¹⁰⁸ ŘEHÁK, Jan. Data, statistika a sociologie: analýza dat explorační. *Sociologický časopis* [online]. 1981, 11.12.2017, XVII. [cit. 2021-11-02]. Dostupné z: https://encyklopedie.soc.cas.cz/w/Anal%C3%BDza_dat_explora%C4%8Dn%C3%AD

Mahdalová data ověřuje z více zdrojů, oficiální statistiky poskytnuté úřady porovnává s jejich výročními zprávami a hledá možné rozdíly. Chyby v datech takto reportovala několikrát. *„Stává se to. Před lety jsem dokonce reportovala chybu v datech Bílému domu. Mám pro to pochopení, většinou jsou poskytovatelé dat rádi, když je na chybu upozorním,“* dodává.

Kladné zkušenosti s reportováním chybných dat poskytovatelům má také Vojíš, který věří, že poučená veřejnost nebo novináři, kteří veřejná data zpracovávají, mohou pomáhat s jejich zpětným ověřováním, a tedy dosáhnout zlepšení jejich kvality. *„Častěji chyby nalézám u historických dat, která nikdo přede mnou nezpracovával,“* říká.

Shovívavější přístup ke zpracování otevřených dat má Boček, který se jako jediný v rozhovoru zabíral jejich pravdivostí. *„Pokud jsou data použitelná, považuji je za pravdivá. (...) Jakmile vidím, že nejsou v dostatečné kvalitě, беру je s rezervou. (...) Nemyslím si, že by v Česku někomu stálo za to si data vymýšlet. Problém je opravdu spíše v tom, že je v datech nepořádek, a ne v tom, že by lidé nechtěli říct pravdu.“*

Formuje podle Vás výběr a kvalita zveřejňovaných dat státní správnou datovou žurnalistiku?

Dat zveřejňovaných státní správou podle datových novinářů přibývá, paradoxně ale ne těch, která by úřadům mohla být nepříjemná. Kočí mluví o „datové mlze“, tedy publikování velkého množství nepřiliš užitečných data setů pro uspokojení poptávky a navození pocitu otevřenosti. *„Skutečně podstatné věci nebo ty, ze kterých by mohly vyplynout problémy, si úřady nechávají pro sebe,“* říká Kočí, který připouští, že úřady zveřejňováním či zatajováním dat mohou ovlivňovat, co bude v mediálním prostoru rezonovat.

Cibulka upozorňuje, že problém vzniká již tím, že se spousta informací v České republice vůbec nesbírá a nesleduje, což přináší problémy napříč obory: *„Zatímco ekonomové z těchto důvodů nemohou navrhnout politiky nebo zkoumat daný problém, my o něm nemůžeme psát. (...) Veřejná správa není schopna dělat analytickou činnost,“* říká Cibulka.

Boček dodává, že i u dat, která jsou veřejně dostupná, mnohdy potřebuje specifický úhel pohledu. Dosažení hypotetického „ideálního“ stavu, kdy by státní správa poskytovala všechny potřebné údaje, proto podle něj ani není možné. Lepší kvalita a dostupnost dat by ale podle něj vedla k rozšíření řad datových novinářů: *„Měli bychom více datových novinářů jako jsem já, tedy těch, kteří příběhy doplňují daty a méně jako Cibulka, který o data často*

bojuje. Datová žurnalistika by byla více o vyprávění příběhů a méně o hrdinech, kteří dokážou data získat.“ Boček zmiňuje, že k české novinářské tradici náleží soubor se státem a „hrdinství žurnalistiky“ se soustředí na to, kdo nad státní správnou zvítězí: *„Dobry novinář je v našem kontextu mnohdy někdo, kdo přepere státní správu. (...) Další datoví novináři, kteří se objevují, už budou více jako Cibulka a budou o dostupnost dat bojovat, protože je to potřeba,*“ uzavírá.

Jaké jsou hlavní limity Vaší práce?

Nedostupnost datových souborů či omezování jejich výběru ale respondenti neuvedli mezi hlavními limity své práce, vyjma situace okolo pandemických dat. Kočí dříve za hlavní omezení vnímal svou nedostatečnou technologickou znalost, nyní je limitem interpretace, kterou shrnuje do hesla: *„Informace jsou levné, porozumění je drahé.*“ Vyvozování závěrů nebo doporučení pro čtenáře je bez hlubší znalosti podstaty věci dle Kočího téměř nemožné. V tomto ohledu podle něj samotná data nejsou dostačující: *„Avantgarda datové žurnalistiky se vrací k tradiční žurnalistice, kdy novináři měli rozdělené beaty, kterým se věnovali denně, a tudíž tomu rozuměli do hloubky. Oklikou se vracím k tomu, že mám pocit, že je nutné něčemu rozumět více.*“

Cibulka limity nachází na straně novinářské obce, která není jednotná ve formulaci svých požadavků na otevřená data: *„Žádná zastřešující organizace nefunguje, čehož státní správa využívá. Nejednotnost funguje v tom, že pokud jedna redakce na něčem pracuje a ostatní kolegové se za ni nepostaví, veřejná správa může říkat, že je to problém dané redakce, nikoliv problém veřejnosti či veřejného zájmu.*“ Dále dodává, že pokud se snaží získat opravdu kriticky důležitý data set, obvykle k němu najde cestu: *„Pokud cestu nenajdeme, o data se soudíme. Podaří se nám ale minimálně získat nějaký menší vzorek dat, díky kterému validujeme danou hypotézu.*“

Pouze Mahdalová uvedla, že ji nedostatek otevřených dat limituje: *„Dat, která bychom mohli zpracovávat, je výrazně více než kolik jich je dostupných. Samozřejmě také potřebujeme, aby redakce, ve které působíme, měla pochopení pro to, v čem je hodnota datové žurnalistiky.*“

Nedostatečné ocenění práce jako hlavní limit vnímá Boček: *„Čtenáři mou práci občas ocení, může mi například napsat i poslanec, že až díky mému textu získal kontext k nějakému tématu. Ale právě proto, že málokdo pozná rozdíl mezi kvalitní a nekvalitní datovou žurnalistikou, může tím trpět má vnitřní motivace se posunout někam dál.*“

Vysvětľujete čtenářům metody práce použité při zpracování dat? Pokud ano, uveďte konkrétní příklad

Vysvětlení postupu zpracování dat je pro respondenty velmi důležité, zejména pro možnost zpětné kontroly. „*Mám zásadu intelektuální poctivosti. To znamená, že bych neměl zneužívat důvěru v datech a prezentovat problém tak, aby vyšel podle toho, jak chci já a vědomě zamlčoval čtenáři skutečnosti, které omezují důvěryhodnost daných dat,*“ shrnuje Cibulka.

Vojř sv  čtenáře nepřeceňuje a ke každé vizualizaci přistupuje s tím, že je nutné vysvětlit konkrétní hodnoty či výpočty hned několikrát: „*Snažím se i na jednu věc poukázat více způsoby, a to jak v samotné grafice, tak i v textu, protože malá písmena v grafikách mohou čtenáři uniknout.*“

Podobný přístup má také Kočí, který u každého grafu nezapomíná na zdroj dat s aktivním odkazem. U analýz, ve kterých došlo k větším rozhodnutím na straně autorů, Kočí metody popisuje doslovně. V podrobnějším vysvětlování metodologie ale vidí překážky na straně publika: „*Pokud do textu přidáme metodologickou poznámku je to moment, kdy většina lidí přestane číst, protože je to obtížně stravitelný text. Říct to co nejpřesněji a zároveň krátce mnohdy nedává smysl. Je to skoro největší umění. Zjistit, pochopit, převyprávět. Důraz kladu na přechod od pochopit k převyprávět,*“ říká Kočí.

Boček zpochybňuje, co vlastně znamená metodika dat. „*Otázka je, jestli můžeme říkat tomu, odkud máme data a co znamenají, metodika. Je to spíš popisek.*“ Metodiku si podle svých slov představuje vysvětlením, proč dané téma zpracovává a proč důvěřuje zdroji, který data poskytl.

Mahdalová své výstupy v minulosti považovala za srozumitelné a bližší vysvětlení k nim neposkytovala. Postupem času si ale dle svých slov všimla v diskusích na sociálních sítích, že zůstávaly nepochopeny. „*Nechci se jakkoliv vysmívat čtenářům, že něco nepochopili. Nemůžu srovnávat sebe, která se zpracováním dat živí, s někým, kdo přijde z práce a rozklikne si článek na internetu. (...) Zároveň se ale snažím si čtenáře dovychovat nebo dovzdělat. Do velké míry to funguje, na pandemii to jde vidět dobře. Leckdy mě čtenáři upozorní na chybu v datech,*“ říká.

2.4. Závěry metodologické části

U všech tuzemských datových novinářů došlo k fúzi informatiky a žurnalistiky, ve svých redakčních rutinách aktivně využívají programovací jazyky Python (Cibulka), statistický program R (Boček a Kočí), ale také open-source nástroje Tableau či Infogram (Mahdalová a Vojíř). Technologie adaptují svépomocí a shodují se na tom, že v začátcích jejich práce technologické znalosti vnímali za limitující, nyní již ne.

Témata, která zpracovávají, odráží jejich žurnalistické profilace, ale také poptávku publika: „*Jsou to měkká témata, ale tvrdý, datový přístup. Jsou komplexní, nesnažím se je jen popisovat, ale pochopit, vyřešit a předat řešení dál,*“ popisuje Boček, který zpracováním dat přechází do žurnalistiky investigativní či žurnalistiky řešení („*solution journalism*“).

Interaktivitu a lokálnost respondenti zmiňují jako klíčové parametry pro svou práci, především Vojíř vyzdvihuje potřebu data zpracovávat na úroveň krajů či obcí s rozšířenou působností.

Podání žádosti o informace prostřednictvím informačního zákona pro datové novináře představuje významnou časovou investici, z tohoto důvodu k ní příliš nepřistupují. Mahdalová zmiňuje, že pro komerční média je získávání dat prostřednictvím informačního zákona téměř nereálné a hodí se proto pro média veřejné služby.

Nástroje Ministerstva vnitra Národní katalog otevřených dat a Open data-wish list respondenti znají, byť jej aktivně nevyužívají. Komunikaci s poskytovateli dat shrnuje Mahdalová do věty: „*Pokud je to něco, co má instituci konfrontovat, ochota je velmi nízká.*“

Novináři pracují s data sety odlišné kvality, důležitá jsou pro ně meta data, tedy popisy jednotlivých parametrů. Respondenti za primární zdroje dat pro svou práci označili veřejná data státní správy, a to s mírnými odlišnostmi danými jejich profilací. Kolektivně ale narazili na nutnost zpracovávat data Ministerstva zdravotnictví během pandemie COVID-19 a v rozhovorech se shodují, že se v oblasti datové žurnalistiky jednalo o přelomovou událost. Poznatky získané na základě hloubkových rozhovorů se autorka práce pokusí validovat či vyvrátit v případové studii.

3. Praktický příklad práce s daty – případová studie

Kvalita i dostupnost datových souborů provázející pandemii COVID-19 je po 21 měsících od jejího vypuknutí dle respondentů uspokojivá. Ministerstvo zdravotnictví společně s Ústavem zdravotnických informací a statistiky (ÚZIS) data ve formátu CSV poskytuje k volnému stažení. Datové sady publikuje na portálech Onemocnění aktuálně.cz, v Národním katalogu otevřených dat a Národním zdravotnickém informačním systému. V národním katalogu je na toto téma k dispozici 55 datových sad, 17 z nich je aktualizováno denně.¹⁰⁹ Na první pohled se může zdát, že ministerstvo naplňuje požadavky na otevřená data dle zákonné definice. Retrospektivní pohled na vývoj pandemie od prvních potvrzených případů nákazy však popisuje situace, kdy společnost neměla k datům dostatečný přístup a docházelo k utajování informací.

Do případové studie byly zahrnuty pouze texty, které se věnují nedostatkům či nemožnosti přístupu k datovým zdrojům a důsledkům těchto překážek. Autorka práce takových článků identifikovala 20, většina z nich s přesahem do publicistiky, mnohdy naplňující znaky komentáře. Analýza proběhla na totožném vzorku napříč celou prací a je zpracována na časové ose ve formátu: datum zveřejnění textu, titulek, autor textu, médium.

3.1. Časová osa

První tři případy nákazy novým typem koronaviru SARS-CoV-2 byly v České republice zaznamenány 01.03.2020.¹¹⁰ Již o měsíc později vychází na iROZHLAS.cz první

¹⁰⁹ Datové sady se dělí do pěti kategorií: epidemiologická data (17 sad), testování (4 sady), očkování (24 sad), kapacity intenzivní péče (4 sady), různé (1 sada).

¹¹⁰ V České republice jsou první tři potvrzené případy nákazy koronavirem: Tisková zpráva. *Ministerstvo zdravotnictví České republiky* [online]. 2020, 01.03.2020 [cit. 2021-11-22]. Dostupné z: <https://koronavirus.mzcr.cz/v-ceske-republice-jsou-prvni-tri-potvrzene-pripady-nakazy-koronavirem/>

publicistický text vyjadřující nespokojenost s poskytovanými informacemi pověřených úřadů.

- *17.4.2020: Jak přesný je vládní model šíření koronaviru? Nikdo neví, data i parametry modelu úředníci tají¹¹¹ (Boček, iROZHLAS.cz)*

Česká vláda s týdenní frekvencí zveřejňovala prognózy vývoje pandemie na základě matematického modelu. Parametry modelu i vstupní data ale tajila, zveřejňovala také kusé informace o nakažených novým typem koronaviru. Boček se snažil od Ústavu zdravotnických informací a statistiky získat odpovědi na otázky: „*Jaké typy testů se u nás nyní používají a které z nich se do statistik dostanou? Jaké je zpoždění od provedení testu k jeho zaznamenání na dashboardu? Bude ve zveřejňovaných datech rozlišení na testy a testované osoby? Jsou dostupná aktuální data o úmrtnosti na ostatní příčiny, aby bylo vidět, kolik si opatření proti nákaze vyžádá obětí z jiných příčin?*“¹¹²

Necelé dva měsíce od začátku pandemie debata o vládních opatřeních podle Bočka nestála na faktech a „*vedla se spíš po linii důvěry nebo nedůvěry v kompetence vlády.*“ Ústav zdravotnických informací a statistiky odmítl údaje zveřejnit z důvodu ochrany soukromí, který informační zákon připouští. Data o pozitivně testovaných proto prezentoval pouze s rozlišením na kraje. Jak poznamenal Boček, „*stejná data s rozlišením na okresy ale zveřejňovaly krajské hygienické stanice, každá ovšem v jiném formátu.*“ Ústav se podle něj tedy pouze snažil analytikům ztížit práci. Ředitel ústavu Ladislav Dušek jako důvod pro

¹¹¹ BOČEK, Jan. Jak přesný je vládní model šíření koronaviru? Nikdo neví, data i parametry modelu úředníci tají. *IROZHLAS.cz* [online]. 2020, 17.04.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/komentare/koronavirus-otevrena-data-uzis_2004171327_jab

¹¹² Tamtéž

neposkytnutí dat opakovaně zmiňoval obavy ze šíření poplašných zpráv, a to, pokud by došlo ze stran novinářů k jejich mylné interpretaci.

- *17.9.2020: Praha jako první zveřejnila podrobná data o epidemii, nejvíc v přepočtu na obyvatele hlásí Praha 3 a Praha 2¹¹³ (Cibulka, iROZHLAS.cz)*

První vlna pandemie z pohledu kvantity dat byla pouze přípravou pro nadcházející období. Zatímco Ústav zdravotnických informací a statistiky stále neposkytoval veškeré požadované údaje, které sbíral, Praha po půl roce zveřejnila první detailnější statistiky o vývoji pandemie. Cibulka Ministerstvo zdravotnictví o detailní data požádal za pomoci informačního zákona, resort ale jeho žádost ve stanovené lhůtě nevyřídil. „*Žádosti se úplně odmítl zabývat Ústav zdravotnických informací a statistiky s tím, že rozhodnout musí ministerstvo.*“¹¹⁴ Nebylo tedy zcela jasné, kdo za zveřejňování dat a jejich kvalitu nese odpovědnost.

- *23.9.2020: Ministerstvo tají informace o počtu nakažených v obcích. ‚Musíme uklidňovat situaci,‘ tvrdí¹¹⁵ (Cibulka, iROZHLAS.cz)*

S týdenním odstupem po publikování textu o nevyřízení žádosti o informace Ministerstvo zdravotnictví Cibulkův požadavek zveřejňovat souhrnná data o pandemii na úroveň obcí s rozšířenou působností oficiálně zamítlo, a to s odvoláním na osobní údaje, které by mohly identifikovat nakažené, a z obav o vyvolání paniky. Starostové obcí však informace dostávali pravidelně v e-mailové korespondenci a Cibulka proto argumentoval názory právníků specializujících se na oblast zdravotnictví, kteří počet „*nakažených na region obce*

¹¹³ CIBULKA, Jan. Praha jako první zveřejnila podrobná data o epidemii, nejvíc v přepočtu na obyvatele hlásí Praha 3 a Praha 2. *IROZHLAS.cz* [online]. 2020, 17.09.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-covid-19-koronavir-cisla-statistiky-uzis-dusek_2009171321_cib

¹¹⁴ Tamtéž

¹¹⁵ CIBULKA, Jan. Ministerstvo tají informace o počtu nakažených v obcích. ‚Musíme uklidňovat situaci,‘ tvrdí. *IROZHLAS.cz* [online]. 2020, 23.09.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-uzis-covid-19-cisla-informace_2009231251_cib

označili za jednoznačně anonymní“ a chování ministerstva za zmatečné. Cibulka se proti rozhodnutí odvolal.

- 2.10.2020: *Nejenom Praha. Podrobnější čísla o epidemii ukazují na ohniska na Náchodsku a Uherskohradištsku*¹¹⁶ (Cibulka, iROZHLAS.cz)

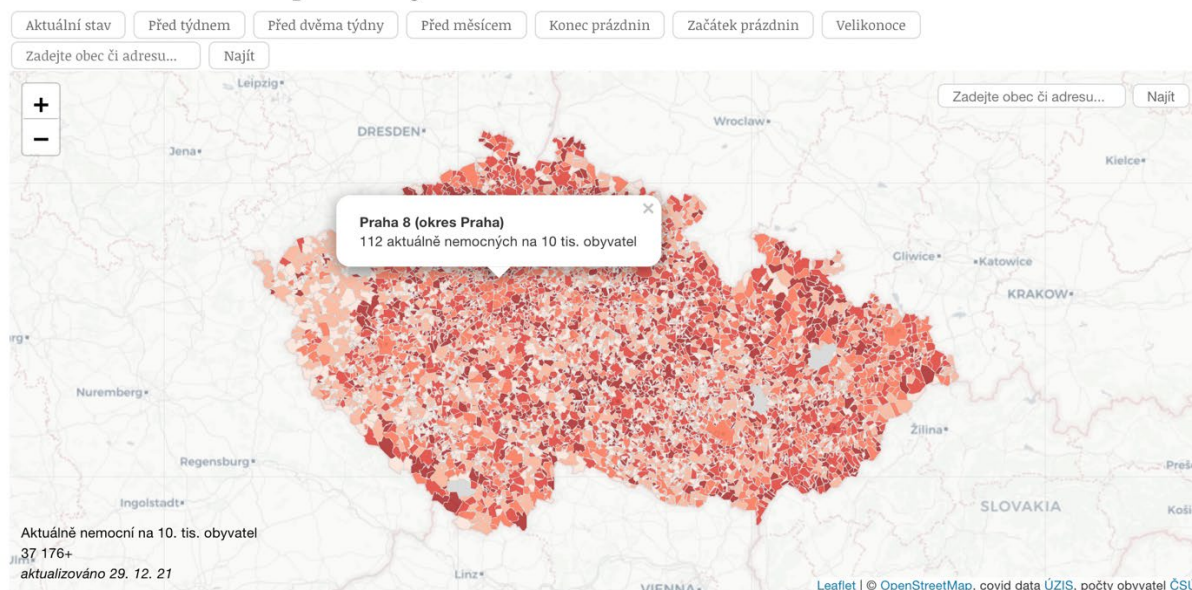
Během následujících dní Ústav zdravotnických informací a statistiky informace na úrovni obcí s rozšířenou působností zveřejnil jako otevřená data na portále Onemocnění aktuálně.cz., na rozšíření datových sad ale neupozornil. Cibulkovo odvolání resort oficiálně nevyřídil, jeho nový ministr Roman Prymula na tiskové konferenci na dotazy novinářů oznámil, že v budoucnu bude resort v oblasti dat více transparentní. Zveřejněná data

¹¹⁶ CIBULKA, Jan. Nejenom Praha. Podrobnější čísla o epidemii ukazují na ohniska na Náchodsku a Uherskohradištsku. *IROZHLAS.cz* [online]. 2020, 02.10.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-covid-nemoc-luzka-obsazenost-statistiky-pravo-na-informace_2010031933_cib

umožnila Cibulkovi vizualizovat ohniska nákazy, nově novináři měli k dispozici také data o obsazenosti nemocničních lůžek.

Příloha č.5: Cibulka na základě získaných dat vytvořil interaktivní mapy, které se aktualizují (obrázek)

Kolik lidí má testem potvrzený covid-19 ve vaší obci



Kliknutím na obec v mapě zobrazíte podrobnosti.

Autor: Cibulka (2020), zdroj: https://www.irozhlas.cz/zpravy-domov/koronavirus-rychlost-sireni-obce-okresy-aplikace-mapa_2010260600_pek

- 14.10.2020: *Koronavirová krize: Vláda se rozhoduje podle neúplných dat*¹¹⁷ (Mahdalová, Seznam Zprávy.cz)

S rostoucími počty nakažených během podzimních měsíců Česká republika znovu vstoupila do nouzového stavu, s čímž rostla poptávka široké i odborné veřejnosti a žurnalistů k zveřejňování úplných a časově relevantních statistik, podle kterých se vláda rozhodovala o zavedení jednotlivých opatření. Mahdalová poukázala na nesystematičnost a nejasnost toho: „*kteří konkrétní opatření bude mít jaký vliv na zpomalení šíření viru: Podklady Ústavu*

¹¹⁷ MAHDALOVÁ, Kateřina. Koronavirová krize: Vláda se rozhoduje podle neúplných dat. *Seznam Zprávy.cz* [online]. 2020, 14.10.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/koronavirova-krize-vlada-se-rozhoduje-podle-neuplnych-dat-124235>

zdravotnických informací a statistiky jsou neúplné. ÚZIS svou analýzu, kterou vládě prezentoval v pondělí, postavil na datech ze soboty, v některých pasážích i na datech z předchozího týdne a na základě epidemiologického modelu kalibrovaného podle zářijových dat.“ Mahdalová zpochybnila kroky vlády a poukázala na nemožnost kontroly, která limitovala práci novinářům.

- *1.11.2020: Jak se čaruje s daty: Hygieny znají více nakažených než státní statistika¹¹⁸ (Mahdalová, Seznam Zprávy.cz)*

Mahdalová denně porovnávala data na webových stránkách čtrnácti Krajských hygienických stanic s oficiálními statistikami Ústavu zdravotnických informací a statistiky. Přišla přitom

¹¹⁸ MAHDALOVÁ, Kateřina. Jak se čaruje s daty: Hygieny znají více nakažených než státní statistika. *Seznam Zprávy.cz* [online]. 2020, 01.11.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/jak-se-caruje-s-daty-hygieny-znaji-vice-nakazenych-nez-statni-statistika-126980>

na zásadní rozdíly v řádech tisíců nakažených. Kontrolními výpočty Mahdalová získala čtyři různé verze toho, kolik je aktuálně vedeno pozitivně testovaných.

Porovnání počtu aktuálně nakažených podle zdrojů¹¹⁹:

- *Dashboard Ministerstva zdravotnictví: 183 020*
- *Otevřená data Ministerstva zdravotnictví (covid-19: Přehled epidemiologické situace dle hlášení krajských hygienických stanic podle ORP): 177 713*
- *Otevřená data Ministerstva zdravotnictví (covid-19: Datové sady pro regiony ČR: Epidemiologické charakteristiky obcí): 127 017*
- *Krajské hygienické stanice: 205 389*

Krajské hygienické stanice, Ministerstvo zdravotnictví ani ÚZIS dostatečné vysvětlení chyb v datech či správné statistiky Mahdalové navzdory opakovaným žádostem neposkytlo.

- *4.11.2020: Máte 107 nakažených, pardon, jednoho. Špatná data „zamořila“ menší obce¹²⁰ (Mahdalová, Seznam Zprávy.cz)*

Šíření nákazy Mahdalová pro Seznam Zprávy.cz vizualizovala pomocí interaktivní mapy, a to s denní aktualizací. Ústav zdravotnických informací a statistiky však v otevřených datech vykazoval nepřesné údaje a nově nakažené „nesprávně přiřadil do malých obcí namísto velkých měst. Na základě toho se na redakci obrátilo několik zástupců obcí s tím, že se údaje významně rozcházejí.“¹²¹ Mahdalová se čtenářům pokusila chyby v datech vysvětlit a uvedla, že „data mají kolísavou spolehlivost a je obtížné je interpretovat.“ Mimo jiné nebyla

¹¹⁹ Tamtéž

¹²⁰ MAHDALOVÁ, Kateřina. Máte 107 nakažených, pardon, jednoho. Špatná data „zamořila“ menší obce. *Seznam Zprávy.cz* [online]. 2020, 04.11.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/mate-107-nakazenych-pardon-jednoho-spatna-data-zamorila-mensi-obce-127568>

¹²¹ Tamtéž

například dostupná data o počtech uzdravených a zemřelých, tedy klíčové parametry pro výpočet poklesu nakažených.

- *17.11.2020: Ředitel ÚZIS by chtěl přestat vydávat denní počty obětí covidu. Není to dobrý nápad¹²² (Boček, iROZHLAS.cz)*

Ústav zdravotnických informací a statistiky s přibývajícím počtem zemřelých hodlal statistiky úmrtí přestat zveřejňovat, a to s obavami o dezinterpretaci ze strany novinářů. Ředitel ústavu Ladislav Dušek podle zjištění médií označil žurnalisty na neveřejném setkání za „tlupu hyen.“ Boček v rozsáhlém komentáři připustil, že interpretace dat není snadná: „Zda jde o úmrtí s covidem, nebo na covid je relevantní otázka a ostrá hranice mezi nimi není; proto není počet obětí covidu tak tvrdé číslo, jak ho většina médií chápe. Řešením ale není přestat data poskytovat. S dezinterpretací se má bojovat tím, že se poskytuje víc a detailnějších informací, ne méně nebo žádné.“

Boček na případu nakažených v menších obcích vyzdvihl, že dezinformace vznikaly na straně poskytovatele dat a jejich zveřejnění pomohlo chyby odhalit. Připomněl také, že na samém začátku pandemie kolega Cibulka žádal ÚZIS o základní anonymizovaná data o testování: kolik osob bylo testováno, zda šlo o test PCR či antigenní, kdy k testování došlo a s jakým výsledkem. Tedy údaje, které by pomohly k přesnější interpretaci dat a možnému uklidnění veřejnosti. Po opakovaném odmítnutí žádosti Cibulka přistoupil k podání žaloby. Boček dále vysvětluje zmiňované důvody zamítnutí: „Nenamítají ochranu soukromí, ale svéhlavě odkazují na paragraf, podle kterého veřejnost zkrátka nemá na zdravotnické statistiky nárok. Důvod přitom není jasný ani ze zákona, ani z důvodové zprávy. Paragraf

¹²² BOČEK, Jan. Ředitel ÚZIS by chtěl přestat vydávat denní počty obětí covidu. Není to dobrý nápad. *IROZHLAS.cz* [online]. 2020, 17.11.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/komentare/komentar-uzis-data-covid_2011170629_jab

do zákona skrze zdravotnický výbor dostal sociálnědemokratický poslanec a lékař Jiří Běhounek. Ten si ovšem na jeho účel nevzpomíná.“

- *24.11.2020: PES a chaos v datech: testy ukazují jiná čísla¹²³ (Mahdalová, Seznam Zprávy.cz)*

Další chyby v datech odhalila Mahdalová při kontrolním výpočtu rizikového skóre PES, které sloužilo k přijímání a orientaci v zavádění vládních opatření. Ústav zdravotnických informací a statistiky bez vysvětlení změnil metodiku výpočtu a z „26 procent pozitivně testovaných se během pár desítek minut stalo 23 procent, což významně změnilo rizikové skóre.“¹²⁴ Na základě výpočtu se například uzavíraly školy či restaurační zařízení, přesnost výpočtu byla klíčová a veřejností bez poskytnutí vstupních dat téměř neověřitelná.

- *5.1.2021: Zdravotníci statistici odmítají zveřejnit podrobné informace o očkováních, uvedou jen týdenní součty¹²⁵ (Cibulka, iROZHLAS.cz)*

S příchodem dostupných vakcín proti koronaviru a jejich nákupu vládou v novinářské obci vyzvuly otázky s očkováním spojené. Ústav zdravotnických informací a statistiky ale odmítl podrobnější statistiky zveřejňovat a poskytoval pouze souhrnné údaje s týdenním odstupem: celkový počet očkováních a počet zakoupených vakcín. Klíčové ale bylo podle novinářů sledovat například proočkovanost jednotlivých věkových skupin. Cibulka za pomoci informačního zákona požadoval: „*anonymizovaná data o očkováních, kromě dne očkování a věku očkování také informace o okresu bydliště a titul pro očkování, tedy zda šlo o*

¹²³ MAHDALOVÁ, Katerina. PES a chaos v datech: testy ukazují jiná čísla. *Seznam Zprávy.cz* [online]. 2020, 24.11.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/jak-ministersky-pes-zastekal-a-onemocnel-130982>

¹²⁴ Tamtéž

¹²⁵ CIBULKA, Jan. Zdravotníci statistici odmítají zveřejnit podrobné informace o očkováních, uvedou jen týdenní součty. *IROZHLAS.cz* [online]. 2021, 05.01.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-ockovani-covid-uzis-informace_2101051741_cib

zdravotníka, seniora či člověka s rizikovým zdravotním stavem.“ Ústav zdravotnických informací a statistiky jeho žádost zamítl.

- *13.1.2021: Zveřejňování čísel o očkování vážne. Ministerstvo s jejich poskytnutím otálí, obchází při tom zákon¹²⁶ (Cibulka, iROZHLAS.cz)*

Cibulka o totožná data požádal Ministerstvo zdravotnictví, resort jeho žádost nevyřídil ve stanoveném termínu, a to z důvodu „přetížení úředníků“. Jak ale Cibulka vysvětluje v komentáři, informační zákon takový důvod pro nevyřízení žádosti nepřipouští.

- *20.1.2021: Data uniklá z ÚZIS: Praha v počtu vakcín válcuje ostatní kraje¹²⁷ (Mahdalová, Seznam Zprávy.cz)*

Mahdalová neoficiální cestou získala první data o očkováných. Poukázala na paradox, že podrobné informace od ÚZIS nedostávají kromě novinářů ani zákonodárci, kteří nemohou provádět dostatečně informovaná rozhodnutí. Netransparentnost informací o dostupnosti vakcín vedla podle Seznam Zpráv.cz k tomu, že se v některých krajích nebylo možné

¹²⁶ CIBULKA, Jan. Zveřejňování čísel o očkování vážne. Ministerstvo s jejich poskytnutím otálí, obchází při tom zákon. *IROZHLAS.cz* [online]. 2021, 13.01.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-ockovani-zakon-106-pravo-na-informace-isin-covid-koronavirus-vakciny_2101131430_cib

¹²⁷ MAHDALOVÁ, Kateřina. Data uniklá z ÚZIS: Praha v počtu vakcín válcuje ostatní kraje. *Seznam Zprávy.cz* [online]. 2021, 20.01.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/data-unikla-z-uzis-praha-v-poctu-vakcin-valcuje-ostatni-kraje-138615>

„oficiálně registrovat k očkování nebo registrace proběhla, ale konkrétní nemocnice vakcínu neměla k dispozici.“

- 22.1.2021: *Ministr Blatný neříká pravdu. Vakcíny se stále rozdělují nerovnoměrně*¹²⁸ (Mahdalová, Seznam Zprávy.cz)

Na základě uniklých dat Mahdalová vyvrátila tvrzení ministra zdravotnictví Blatného, který o distribuci vakcín tvrdil, že je rovnoměrná podle počtu seniorů v daném kraji. Z dat bylo zřejmé, že byly výrazně lépe zásobeny Praha a Jihomoravský kraj. Mahdalová text o počtech naočkovaných seniorů v jednotlivých krajích doplnila o vysvětlení, že Ústav zdravotnických dat a informací odmítá data publikovat oficiální cestou, a proto datoví novináři spoléhají na data neoficiální. Za jejich správnost se zaručila. „Po zveřejnění článku se na redakci obrátila Lenka Svobodová, vedoucí oddělení komunikace a vztahů s veřejností ÚZIS, s tím, že úřad požadovaná data zveřejňovat neodmítá. Na webu spravovaném ÚZIS se však ani na hlavní stránce ani v otevřených datových sadách počty vakcinovaných podle krajů a věkové struktury nenacházejí. Ústav přidal záložku „Očkování“, pod ní se však nacházejí pouze souhrnná data za celou Českou republiku.“¹²⁹

- 12.3.2021: *Neveřejná data z českých nemocnic. Kdo tam bojuje s covidem*¹³⁰ (Mahdalová, Seznam Zprávy.cz)

Statistiky o očkovaných ale nebyly jedinými, které ÚZIS novinářům stále neposkytoval. V průběhu do té doby nejsilnější vlny pandemie neposkytoval údaje o hospitalizovaných pacientech v těžkém stavu, klíčový parametr pro determinaci vytíženosti zdravotnického systému. Mahdalová získala data neoficiální cestou, a to na úroveň jednotlivých nemocnic. Pro citlivost údajů se rozhodla informovat v obecné rovině, a to o tom, že se „v nemocnicích

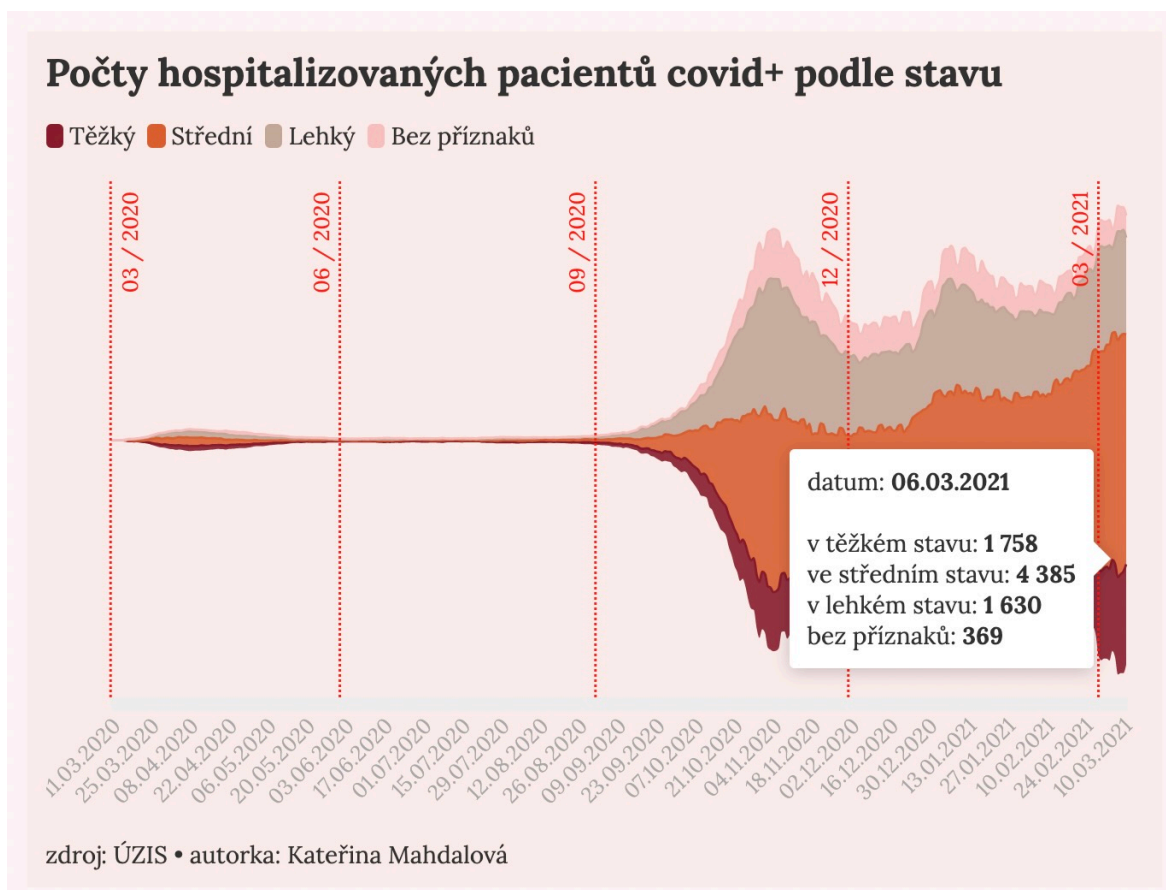
¹²⁸ MAHDALOVÁ, Kateřina. Ministr Blatný neříká pravdu. Vakcíny se stále rozdělují nerovnoměrně. *Seznam Zprávy.cz* [online]. 2021, 22.01.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/ministr-blatny-nerika-pravdu-vakciny-se-stale-rozdeluji-nerovnomerne-138890>

¹²⁹ Tamtéž

¹³⁰ MAHDALOVÁ, Kateřina. Neveřejná data z českých nemocnic. Kdo tam bojuje s covidem. *Seznam Zprávy.cz* [online]. 2021, 12.03.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/data-v-cesku-je-dvakrat-vice-pacientu-v-tezkem-stavu-zez-pocatkem-roku-146945>

nacházelo dvakrát více pacientů v těžkém stavu než počátkem roku. Nejvíce zahlcené kraje byly Liberecký, Královéhradecký a Karlovarský. Výrazně hůře na tom byli muži než ženy.“

Příloha č.6: Mahdalová vizualizovala počty hospitalizovaných pacientů (obrázek)



Autorka: Mahdalová (2020), zdroj: <https://www.seznamzpravy.cz/clanek/data-v-cesku-je-dvakrat-vice-pacientu-v-tezkem-stavu-nez-pocatkem-roku-146945>

- 26.3.2021: *Vláda čte špatně vlastní data: uzavření okresů nemá na šíření viru vliv*¹³¹ (Mahdalová, Seznam Zprávy.cz)

Mahdalová na základě dalších neoficiálně získaných dat ÚZIS společně s daty Atlasu mobility Technické univerzity v Ostravě ověřila, že vládní nařízení o uzavření okresů, tedy

¹³¹ MAHDALOVÁ, Kateřina. Vláda čte špatně vlastní data: uzavření okresů nemá na šíření viru vliv. *Seznam Zprávy.cz* [online]. 2021, 26.03.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/vlada-cte-spatne-vlastni-data-uzavreni-okresu-nema-na-sireni-viru-vliv-148532>

zákaz se mezi okresy volně pohybovat, na šíření viru nemělo zásadní vliv. Vláda podle Mahdalové neměla dostatečné informace o tom, zda dané opatření funguje a rozhodla se toto nepopulární rozhodnutí přijmout pouze na základě apelu ministra zdravotnictví Blatného, který data účelově dezinterpretoval. „(Blatný) Operuje tím, že poklesl pohyb lidí mezi okresy, už ale neříká, že právě tento pohyb byl už v minulosti mnohem menší než pohyb v rámci okresů. Další problém je, že si Blatný účelově vybírá data, která podporují uzávěru, namísto aby podporoval opatření, která reagují na skutečný stav věcí.“¹³²

- 12.4.2021: Soud rozhodl: čísla o pandemii laikům do rukou nepatří¹³³ (Cibulka, iROZHLAS.cz)

Po devíti měsících od podání žádosti podle informačního zákona Cibulka dostal rozhodnutí: 5. senát Městského soudu v Praze rozhodl, že se na statistiky o pandemii zákon o svobodném přístupu k informacím nevztahuje a v platnosti je zákon o veřejném zdraví. „Ministerstvo zdravotnictví získalo oporu v rozsudku, podle kterého můžou být jeho statistiky neúplné a jejich zveřejnění a interpretace by u občanů mohla vyvolávat zbytečné obavy.“¹³⁴

- 8.6.2021: Tajení informací o zdravotnictví je protizákonné, uvedl městský soud. Navrhuje zrušit část zákona¹³⁵ (Cibulka, iROZHLAS.cz)

S rozhodnutím se neztotožnil soudce Městského soudu v Praze Štěpán Výborný, který rozhodnutí označil jako v rozporu s Ústavou České republiky. „Nyní rozhodující senát 14 A se však se závěry 5. senátu neztotožnil, a proto shledal za nezbytné [...] předložit návrh Ústavnímu soudu,“ citoval Výborného iROZHLAS.cz. K odmítání zveřejnění informací ze strany statistiků dodal, že není možné žádosti o informace plošně odmítat a problém by neměl představovat ani fakt, že interpretace zdravotnických dat je složitá: „Možná rizika při

¹³² Tamtéž

¹³³ CIBULKA, Jan. Soud rozhodl: čísla o pandemii laikům do rukou nepatří. *IROZHLAS.cz* [online]. 2021, 12.04.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/komentare/data-covid-statistika-soud-rozsudek-praco-na-informace_2104120629_cib

¹³⁴ Tamtéž

¹³⁵ CIBULKA, Jan. Tajení informací o zdravotnictví je protizákonné, uvedl městský soud. Navrhuje zrušit část zákona. *IROZHLAS.cz* [online]. 2021, 08.06.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/informace-behounek-uzis-mzcr_2106081139_cib

neetickém výkladu daných informací pak zajisté může minimalizovat statistický ústav tím, že případně poskytne jejich náležitou interpretaci,“ dodal Výborný.

- 18.6.2021: Unikátní data: Jak virus postupuje školami. Jsou tři rizikové kraje¹³⁶ (Mahdalová, Seznam Zprávy.cz)

Ústav zdravotnických informací a statistiky disponoval mimo jiné údaji o nakažených ve školách či domovech pro seniory a také o počtech infikovaných v jednotlivých pracovních odvětvích. Data sice poskytoval, ale ne souhrnně a pravidelně, a proto v nich datoví novináři nemohli sledovat trendy. Veřejnost nemohla být dostatečně ujistěna o tom, zda byla uzávěra škol efektivním opatřením. Mahdalová data získala neoficiální cestou a retrospektivně zmapovala, jak se „koronavirus šířil v období celého školního roku ve všech typech škol.“

- 2.9.2021: Jedinečná mapa obcí: Neveřejná data odhalují, jak postupuje očkování¹³⁷ (Mahdalová, Seznam Zprávy.cz)

Po devíti měsících Mahdalová získala podrobná data o očkování na úrovni obcí s rozšířenou působností, a to s denní aktualizací. Popsala, že ačkoliv data nebyla dostupná jako otevřená, ÚZIS je „rozesílal na stovky blíže neurčených mailových adres.“ Mahdalová data poprvé získala již v dubnu, v srpnu ÚZIS data zveřejnil, ale ve formátu přesahujícím velikost 1,5

¹³⁶ MAHDALOVÁ, Kateřina. Unikátní data: Jak virus postupuje školami. Jsou tři rizikové kraje. *Seznam Zprávy.cz* [online]. 2021, 18.06.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/unikatni-data-virus-ze-skol-mizi-zbyvaji-tri-rizikove-kraje-167590>

¹³⁷ MAHDALOVÁ, Kateřina. Jedinečná mapa obcí: Neveřejná data odhalují, jak postupuje očkování. *Seznam Zprávy.cz* [online]. 2021, 02.09.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/jedinecna-mapa-obci-neverejna-data-odhaluji-jak-postupuje-ockovani-152402>

GB, což podle Mahdalové zamezilo zpracování dat běžným kancelářským softwarem a nesplnilo tedy požadavky stanovené Ministerstvem vnitra na otevřená data.

- 4.11.2021: *Stát nemá přehled o tom, jak se šíří covid ve školách*¹³⁸ (Mahdalová, Seznam Zprávy.cz)

Doposud poslední nejasnosti panovaly okolo šíření nákazy ve školách. Mahdalová upozornila, že tyto údaje doposud nejsou publikovány souhrnně, a je tedy na Krajských hygienických stanicích jednotlivých krajů, zda novinářům data poskytnou. „*Ministerstvo zdravotnictví data novinářům neposkytlo s tím, že tiskový odbor zjišťoval, zda údaje vůbec mají.*“¹³⁹ Panovala přitom nejasnost, zda má data evidovat Ministerstvo zdravotnictví, Ministerstvo školství, či krajské hygieny.

3.2. Závěry případové studie

Poznatky získané v případové studii diplomantka validuje za využití třech základních znaků otevřených dat dle definice v teoretické části práce a pokusí se tak ověřit, zda byly poskytované datové sady dostupné. Následovat bude hodnocení situace ze strany datových novinářů Cibulky a Mahdalové, které autorka práce zařadila mezi otázky hloubkových rozhovorů.¹⁴⁰

Data se v průběhu pandemie začala sbírat a zveřejňovat poměrně rychle, pro Ústav zdravotnických informací a statistiky to byla nová situace, na kterou dle Cibulky i Mahdalové nebyl dostatečně připraven. „*Na začátku datové sady obsahovaly chyby. Nezbyvalo než data validovat vůči základním pravidlům, například že v obci nemůže být více nakažených lidí, než kolik v ní žije. Bylo nutné komunikovat s poskytovatelem a pořizovatelem dat,*“ popisuje situaci Cibulka.

¹³⁸ MAHDALOVÁ, Kateřina. Stát nemá přehled o tom, jak se šíří covid ve školách. *Seznam Zprávy.cz* [online]. 2021, 04.11.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/ve-kterych-skolach-se-siri-covid-prehled-nema-ministerstvo-ani-mnohe-hygieny-179532>

¹³⁹ Tamtéž

¹⁴⁰ Všechny uvedené citace autorů jsou dostupné v přílohách diplomové práce

Problém nastává v bodě, kdy si přetížený Ústav drží právo na jedinou správnou interpretaci a další zpracování dat. „*Nestíhají to, neposílili datové týmy. Úřad je přetížen, zároveň ale za rok a půl nezvládnout organizaci není omluvitelné,*“ hodnotí Mahdalová.

Za tři hlavní znaky otevřených dat autorka práce na základě odborné literatury považuje univerzální přístup ve strojově čitelném formátu, neomezenost využití a kvalitní meta data.

3.2.1. Univerzální dostupnost dat ve strojově čitelném formátu

U datových sad opakovaně nebyl naplněn požadavek na poskytování dat ve strojově čitelném formátu, a to bez technických překážek. Mahdalová zmiňuje například sběr dat Krajskými hygienickými stanicemi: „*Snažila jsem se dopočítat prevalenci testů, musela jsem každé ráno navštívit 14 webů Krajských hygienických stanic a ručně z nich data posbírat, a to za každý okres zvlášť.*“ Komunikace s krajskými stanicemi byla dle jejích slov nepružná, ochota poskytnout zdrojová data téměř žádná a úroveň kvality rozdílná: „*Je to pořád Krajská hygienická stanice s nadřízeným úřadem, ale ta spolupráce je s každým krajem absolutně odlišná.*“ Zmiňuje také nemožnost přesné kontroly výpočtů: „*Chtěla jsem si prevalenci dopočítat sama, nechtěli mi dát zdrojová data, jen už vypočítanou prevalenci. Jak mohu bezpečně ověřit, že jsou údaje přesné?*“

3.2.2. Přístup k otevřeným datům není nikomu limitován

Přístup k otevřeným datům by neměl být diskriminován či limitován. Na případu data setu o počtech nakažených podle profesí, ale Mahdalová demonstruje, že k omezení její osoby ze strany ÚZIS opakovaně docházelo. Zatímco České televizi ústav data poskytl, Mahdalová o ně žádala téměř 13 měsíců bez odpovědi. „*Od samého začátku pandemie s ústavem bojuji. Začala jsem podléhat pocitu, zda vedení data schválně nerozmlžuje,*“ říká. Neposkytování kvalitních dat během pandemie vnímá jako „katastrofu“: „*Umírají lidé a jsem přesvědčena o tom, že kdybychom na základě dat loni věděli, ve kterých profesích se nákaza šíří, dalo se s tím něco dělat a mohli na to adekvátně reagovat politici,*“ říká Mahdalová. Za důležité rovněž považuje data zveřejňovat kontinuálně a novinářům umožnit v datech sledovat trendy. Mahdalová značnou část dat získala neoficiální cestou od tajných zdrojů, jednalo se přitom o oficiální data ÚZIS. „*Začali na mě útočit, že jedním neeticky. Nikdy ale*

nerozporovali, že to jsou jejich data. Sami data zveřejňovat nechtěli, nebo je poté začali zveřejňovat s odstupem,“ popisuje.

3.2.3. Data sety jsou opatřeny dokumentací a meta daty dostatečně vysvětlující jejich hodnoty

Nejasnosti ve struktuře poskytovaných dat opakovaně zmiňují Cibulka, Mahdalová i Boček. Mahdalová zdůrazňuje například poskytování dat o pandemii COVID-19 jako absolutních čísel bez poznámky, že se jedná o přepočtené na 100 tisíc obyvatel či problémy interpretačního charakteru. Například u dat o hospitalizovaných pacientech ústav uvádí, zda jsou či nejsou očkovaní. *„Apelovala jsem na ústav, že absolutní čísla nestačí, je potřeba dělat relativní přepočtené na věkové skupiny. Ideální by bylo mít dodatečné informace jako například kdy měl daný člověk ukončené očkování, z toho můžeme usuzovat, zda poklesly protilátky a podobně,*“ říká Mahdalová. Tyto klíčové parametry by ve veřejné debatě napomohly vyvracet dezinformace.

Závěr

Hlavním cílem práce bylo poskytnout ucelený pohled na problematiku otevřených dat a jejich využití v tuzemské datové žurnalistice.

V teoretické části práce autorka popsala spojitost mezi tradiční žurnalistikou a žurnalistikou zaměřenou na práci s daty, její přesah do investigativní novinářiny, a především návaznost na metody společenských věd. Podstatnými pro rozšíření datové žurnalistiky do redakcí zahraničních i tuzemských jsou znaky mediální technologické konvergence a její dopady: interaktivita a lokálnost. Jak zmiňuje odborná literatura datová žurnalistika má potenciál vzdělávat publikum, validovat důvěryhodnost zjištění a vyvracet dezinformace. Limitovat ji mohou technologické dovednosti novinářů, ale také nedostupnost datových souborů. Získávání dat, primárně těch s vysokým veřejným zájmem, se proto stalo součástí práce datových novinářů. Státní správa má jako tradičně silný sběratel a uchovavatel údajů potenciálně moc formovat podobu datové žurnalistiky. Snahy a legislativní závazky vlády v oblasti zveřejňování dat popsala kapitola o informačním zákonu, a to včetně prognóz vývoje po přijetí novely Evropské unie o zveřejňování data setů s vysokou hodnotou. Kvantitativní analýza závěrem teoretické části poskytla vhled do rozšířenosti datové žurnalistiky v českém mediálním prostoru a potvrdila, že ačkoliv se datoví novináři v českých redakcích vyskytují minoritě, v posledních letech o své práci dávají publiku znát.

V metodologické části oslovení datoví novináři dodali svými cennými zkušenostmi kontext mnohým teoretickým poznatkům. Zatímco odborná literatura hovoří o „otevřeném vládnutí“ a nebývale vysokému počtu zveřejňovaných datových sad ze stran státní správy, novináři toto tvrzení dále upřesňují: Ano, dat zveřejňovaných státní správou přibývá, ale jedná se pouze o „datovou mlhu“ která publikováním nerelevantních údajů odvádí pozornost od datových sad, které by mohly úřadům být nepříjemné.

Konkrétní příklady dopadů limitace přístupu k datovým zdrojům a získávání informací v praxi poskytla závěrečná část práce, případová studie. Záměrem diplomantky nebylo hanit kroky zmiňovaných poskytovatelů dat ze strany státní správy v době globální pandemie ale ukázat na obecnější problém dostupnosti a získávání dat v žurnalistice:

neochotu, nesystematičnost a nepochopení role datového novináře jako prostředníka komunikace mezi státem a občany.

Summary

In conclusion, the main aim of the master's thesis was to provide a comprehensive view of the issue of open data and their use in Czech data journalism.

In the theoretical part, the author described the connection between traditional journalism and journalism focused on working with data, its transcendence into investigative journalism, and usage of the methods known from social sciences. Fundamental for the expansion of data journalism to foreign and domestic newsrooms are the features of media convergence: interactivity and locality of news. Data journalism has the potential to educate audiences, validate findings, and contradict misinformation.

Data journalist may be limited by the need of technological skills, but also by the unavailability of data files. Data acquisition, primarily of those with high public interest, has therefore become crucial part of their work. The state administration, as a traditionally strong data collector and custodian, has the power to shape the form of data journalism. Czech government's efforts and legislative obligations in open data publication were described in the chapter on the Freedom of Information Act, including prophesy of developments following the adoption of the European Union amendment on the publication of High-value Data sets. The theoretical part was finalized by the quantitative analysis, which provided an insight into the prevalence of data journalism in the Czech media landscape and confirmed that although data journalists in Czech newsrooms are in the minority, in recent years they let the audience know about their work.

Addressed data journalists in the methodological part provided their valuable insights with context for many theoretical findings. While researchers speak of "open government" and an unusually high totals of published data sets by government, journalists further clarify this statement: Yes, data published by government is increasing, but it is only a "data fog" that leads to the publication of irrelevant data and spans the attention from datasets, which could be bothersome to authorities.

The final part of the thesis, a case study, provided explicit examples of the effects of limiting access to data sources and obtaining information in practice. The author's intention

was not to slander the actions of the mentioned data providers during the global pandemic, but to apprise a more general obstacle of data availability and retrieval in journalism: reluctance, disorganization and misunderstanding of the role of the data journalist as a mediator of communication between the state and its citizens.

Použitá literatura

Tištěné zdroje

BOUNEGRU, Liliana, Lucy CHAMBERS a Jonathan GRAY. *The Data Journalism Handbook*. European Journalism Centre, 2012. ISBN 9781449330064.

DEUZE, Mark. "Convergence Culture in the Creative Industries." *International Journal of Cultural Studies* 10, no. 2, 2007, str. 243–63. <https://doi.org/10.1177/1367877907076793>.

FELLE, Tom. Digital Watchdogs? Data Reporting and the News Media's Traditional 'Fourth Estate' Function. *Journalism* 17, no. 1, 2016, str. 85–96. <https://doi.org/10.1177/1464884915593246>.

GOGGIN, G., MARTIN, F., a DWYER, T., 2015. Locative News. *Journalism Studies*, 16, str. 4-59.

GRANT, August. *Understanding media convergence: the state of the field*. Oxford University Press, 2009. ISBN 9780195327779.

HENDL, Jan. *Kvalitativní výzkum: Základy metody a aplikace* [online]. Praha: Portál, 2005 [cit. 2021-11-12]. ISBN 80-7367-040-2. Dostupné z: https://dl1.cuni.cz/pluginfile.php/882806/mod_resource/content/1/06%20Hendl%2C%20J.%20-%20Kvalitativn%C3%AD%20v%C3%BDzkum%3B%2024-63.pdf

HERMIDA, Alfred a Mary Lynn YOUNG. *Data Journalism and the Regeneration of News* [online]. Routledge, 2019 [cit. 2021-10-02]. ISBN 9781138058934. Dostupné z: <https://www.routledge.com/Data-Journalism-and-the-Regeneration-of-News/Hermida-Young/p/book/9781138058934>

JAROSLAV, Čuřík, Burgr RUDOLF, Jiřina SALAQUARDOVÁ, Martin ŠKOP a Lenka WASCHKOVÁ-CÍSAŘOVÁ. *Nové trendy v médiích I*. [online]. 2. Brno: Masarykova

univerzita, 2012 [cit. 2021-12-02]. ISBN 978-80-210-5826-2. Dostupné z:
https://is.muni.cz/publication/987880/Nove_trendy_v_mediich_I.pdf

KUŽÍLEK, Oldřich. *Dalších 106 odpovědí na Vaše dotazy (s rozsudky): příručka pro občany o svobodném přístupu k informacím a otevřenosti veřejné správy*. Praha: Otevřená společnost, 2004. ISBN 80-903331-1-7.

LEWIS, Seth C a USHER, Nikki. Open Source and Journalism: Toward New Frameworks for Imagining News Innovation. *Media, Culture & Society* 35, no. 5, 2013, str. 02–109. <https://doi.org/10.1177/0163443713485494>.

MCNAIR, Brian. *Sociologie žurnalistiky*. Přeložil Hana ANTONÍNOVÁ. Praha: Portál, 2004. ISBN 80-7178-840-6.

MCQUAIL, Denis. *Úvod do teorie masové komunikace*. 4., rozš. a přeprac. vyd. Praha: Portál, 2009. ISBN 978-80-7367-574-5.

POLLOCK, Rufus. *The Open Revolution: Rewriting the Rules of the Information Age*. Open Revolution, 2018, 126 s. ISBN 9781983033223.

SEDLÁKOVÁ, Renáta. *Výzkum médií: Nejužívanější metody a techniky* [online]. Praha: Grada Publishing, 2014 [cit. 2021-11-02]. ISBN 978-80-247-3568-9. Dostupné z:
<https://books.google.cz/books?id=710nBgAAQBAJ&printsec=frontcover&key=AIzaSyDlPfl89JdFhWBVsMVsavVo6aNh057xlTc#v=onepage&q&f=false>

ŠVAŘÍČEK, Roman, Klára ŠEĐOVÁ. *Kvalitativní výzkum v pedagogických vědách*. Vyd. 2. Praha: Portál, 2014. ISBN 978-80-262-0644-6.

ULMANOVÁ, Jiřina. *Souhrnná výzkumná zpráva o problematice otevřených dat* [online]. In: Praha, 2021, s. 82 [cit. 2021-11-02]. Dostupné z:
<https://vubp.cz/soubory/vyzkum/projekty/V03-S4/Vsouhrn-Souhrnna-vyzkumna-zprava-o-problematice-otevrenych-dat.pdf>

YARNALL, Louise, JOHNSON, J.T., RINNE, Luke. How Post-Secondary Journalism Educators Teach Advanced CAR Data Analysis Skills in the Digital Age. *Journalism & Mass Communication Educator* 63, no. 2, 2008, str. 146–64. <https://doi.org/10.1177/107769580806300204>.

Online zdroje

4 důvody, proč číst Deník. *Vltava Labe Media: Deník* [online]. [cit. 2021-11-07]. Dostupné z: <https://www.vlmedia.cz/denik>

Aleš Vojjř: vedoucí Deník Data. *Deník.cz: My jsme Deník* [online]. [cit. 2021-11-07]. Dostupné z: <https://www.denik.cz/my-jsme-denik/praha/ales-vojir-3115.html>

BOČEK, Jan, Jakub MRÁČEK a Jindřich MYNARZ. *Otevřená data: Příležitost pro Českou republiku* [online]. Nadace Open Society Fund Praha: Fond Otakara Motejla [cit. 2021-11-02]. ISBN 978-80-87725-03-0. Dostupné z: https://osf.cz/wp-content/uploads/2012/08/Opendata2_web_komplet.pdf

BRADSHAW, Noel-Ann. *Florence Nightingale (1820–1910): An Unexpected Master of Data* [online]. 2020 [cit. 2021-12-02]. Dostupné z: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7660360/>. School of Computing and Digital Media, London Metropolitan University, London, UK.

Czech Republic. *Open Government Partnership* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.opengovpartnership.org/members/czech-republic/#commitments>

HUYER, Ester, BLANK, Marit, European Data Portal Publications Office. Analytical Report 15: High-value datasets: understanding the perspective of data providers [online]. Luxembourg: Publications Office of the European Union, 2020, 2020, 21 [cit. 2021-11-29]. ISSN 2600-0601. Dostupné z: doi:10.2830/363773

Jak žádat o informace. *Ministerstvo vnitra České republiky* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.mvcr.cz/clanek/jak-zadat-o-informace-453908.aspx>

KITCHIN, Rob. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences* [online]. California: Sage Publication, 2014 [cit. 2021-11-02]. ISBN 978-1446287484. Dostupné z: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jors.12293>

KUNDRA, Vivek. Digital Fuel of the 21st Century: Innovation through Open Data and the Network Effect. *Harvard Kennedy School* [online]. 2012, 25.01.2012 [cit. 2021-10-08]. Dostupné z: <https://shorensteincenter.org/digital-fuel-of-the-21st-century-innovation-through-open-data-and-the-network-effect/>

LOWREY, Wilson, Ryan BROUSSARD a Lindsey A. SHERRILL. *Data journalism and black-boxed data sets* [online]. 2019 [cit. 2021-11-02]. Dostupné z: <https://journals.sagepub.com/doi/pdf/10.1177/0739532918814451>

MÍŠEK, Jakub. Co jsou otevřená data? *Otevřená data* [online]. 2021 [cit. 2021-11-02]. Dostupné z: <https://opendata.gov.cz/informace:start>

MÍŠEK, Jakub. Otevřená data: Základní přehled právní úpravy. *Otevřená data* [online]. 2021, 30. 09. 2021 [cit. 2021-11-02]. Dostupné z: <https://data.gov.cz/%C4%8D%C3%A1nky/otev%C5%99en%C3%A1-data-z%C3%A1kladn%C3%AD-p%C5%99ehled-pr%C3%A1vn%C3%AD-%C3%BApravy>

Národní katalog otevřených dat na Portálu veřejné správy. *Ministerstvo vnitra České republiky* [online]. [cit. 2021-11-09]. Dostupné z: <https://www.mvcr.cz/clanek/zpravodajstvi-narodni-katalog-otevrenych-dat-na-portalu-verejne-spravy.aspx>

NGUYEN, An a Jairo LUGO-OCANDO. The state of data and statistics in journalism and journalism education: Issues and debates. *Journalism* [online]. 2016, **17**, 3-17 [cit. 2021-10-02]. Dostupné z: [doi:10.1177/1464884915593234](https://doi.org/10.1177/1464884915593234)

Open Data Maturity: Report 2020 [online]. Luxembourg: Publications Office of the European Union, 2020 [cit. 2021-11-20]. ISSN 2600-0512. Dostupné z: https://data.europa.eu/sites/default/files/edp_landscaping_insight_report_n6_2020.pdf

Open Government Data. *OECD: Better policies for better lives* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.oecd.org/gov/digital-government/open-government-data.htm>

PORLEZZA, Colin a SPENDORE, Sergio, 2019, From Open Journalism to Closed Data: *Data Journalism in Italy, Digital Journalism*, 7:9, 1230-1252, DOI: 10.1080/21670811.2019.1657778

PORTER, M.E., What Is Strategy? *Harvard Business Review*, 1996, 74, str. 61-78.

ŘEHÁK, Jan. Data, statistika a sociologie: analýza dat explorační. *Sociologický časopis* [online]. 1981, 11.12.2017, **XVII**. [cit. 2021-11-02]. Dostupné z: https://encyklopedie.soc.cas.cz/w/Anal%C3%BDza_dat_explora%C4%8Dn%C3%AD

STUHLÍKOVÁ, Lucie. Clarifying the complex: Czech Radio's data journalism team. *European Journalism Observatory* [online]. 19.12.2019 [cit. 2021-12-06]. Dostupné z: <https://en.ejo.ch/specialist-journalism/clarifying-the-complex-czech-radios-data-journalism-team>

Tracking the state of open government data. *Global Open Data Index* [online]. [cit. 2021-11-02]. Dostupné z: <https://index.okfn.org/>

Výroční zpráva o stavu otevřených dat v České republice za rok 2019. In: *Otevřená data: Odbor Hlavního architekta eGovernmentu* [online]. Praha: Ministerstvo vnitra, 2020, 16.06.2020, s. 48 [cit. 2021-11-02]. Dostupné z: <https://data.gov.cz/attachments/2020-07-14/V%C3%BDro%C4%8Dn%C3%AD%20zpr%C3%A1va%20o%20stavu%20otev%C5%99en%C3%BDch%20dat%20v%20%C4%8Cesk%C3%A9%20republice%20za%20rok%202019.pdf>

VÝZKUM ČTENOSTI TISKU, MEDIA PROJEKT. *Unie vydavatelů* [online]. [cit. 2021-11-07]. Dostupné z:

http://www.unievychavatelu.cz/cs/unie_vychavatelu/medialni_data/vyzkum_ctenosti

What is open? *Open Knowledge Foundation* [online]. [cit. 2021-11-05]. Dostupné z:

<https://okfn.org/opendata/>

Why open data? *Open Knowledge Foundation* [online]. [cit. 2022-11-07]. Dostupné z:

<https://okfn.org/opendata/why-open-data/>

Zpravodajské texty

BOČEK, Jan a Hana MAZANCOVÁ. VELKÝ PŘEHLED: Ministerské odměny za rok 2016. Kdo data tají a kdo bez problémů zveřejňuje? *IROZHLAS.cz* [online]. 2017, 07.02.2017 [cit. 2021-11-02]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/velky-prehled-ministerske-odmeny-za-rok-2016-kdo-data-taji-a-kdo-bez-problemu-zverejnuje-_201702070600_hmazancova

ČTK. Tým datové žurnalistiky z Economie přešel k rozhlasu. *Mediář* [online]. 2014, 10.10.2014 [cit. 2021-11-07]. Dostupné z: <https://www.mediar.cz/tym-datove-zurnalistiky-z-economie-presel-k-rozhlasu/>

Datová žurnalistika v Česku v roce jedna. *Nadace OSF* [online]. 28.06.2012 [cit. 2021-11-12]. Dostupné z: <https://osf.cz/2012/06/28/datova-zurnalistika-v-cesku-v-roce-jedna/>

IROZHLAS.cz se zařadil do první zpravodajské desítky: v březnu ho navštívily rekordní tři miliony čtenářů. *IROZHLAS.cz* [online]. 2020, 03.04.2020 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/irozhlas-server-zpravodajstvi-rekord-navstevnost-brezen_2004031320_tzr

IROZHLAS.CZ. Novinářská cena 2019: Dataři zvítězili s projektem Rozdělení svobodou, za Český rozhlas uspěl i Lukáš Houdek. *IROZHLAS.cz* [online]. 2020, 17.06.2020 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/novinarska-cena-vyhra-datova-zurnalistika-irozhlas_2006172015_ada

KUBANT, Vít. Náměstci dostali za podzimní boj s covidem i čtvrt milionu. Lékaři a zdravotníci z první linie zatím nic. *IROZHLAS.cz* [online]. 2021, 08.02.2021 [cit. 2021-11-02]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/odmeny-ministerstvo-zdravotnictvi-penize-lekari-hygienici-uzis_2102080500_vtk

KUBIŠTOVÁ, Dominika. Ústavní soud dal datovému novináři za pravdu. Policie nemůže chtít miliony za informace o kriminalitě. *IROZHLAS.cz* [online]. 2021, 08.07.2021 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/policie-data-o-kriminalite-ustavni-soud-novinar-nalez_2107080924_dok

VINKLOVÁ, Julie. Počasí, věc veřejná. Meteorologové po letech sporů začali poskytovat historické záznamy. *IROZHLAS.cz* [online]. 2020, 26.03.2020 [cit. 2021-11-07]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/chmu-pocasi-data-historicke-zaznamy-o-pocasi-pravo-na-informace_2003260832_elev

Zpravodajské weby 2020: Na čele Novinky, v TOP 10 i nový web CNN Prima. *MediaGuru* [online]. 2021, 21.01.2021 [cit. 2021-11-07]. Dostupné z: <https://www.mediaguru.cz/clanky/2021/01/zpravodajske-weby-2020-na-cele-novinky-v-top-10-i-novy-web-cnn-prima/>

Závěrečné práce

HALÁMOVÁ, Anna. *Vývoj a současný stav datové žurnalistiky v České republice* [online]. Olomouc, 2020 [cit. 2021-11-19]. Dostupné z: https://theses.cz/id/2fe1bh/Bakalarska_prace_Halamova_Datova-zurnalistika.pdf. Bakalářská práce. Univerzita Palackého. Vedoucí práce Mgr. et Mgr. Karel Páral.

KOKTOVÁ, Alena. *Svobodný přístup k informacím z pohledu novinářů* [online]. Praha, 2015 [cit. 2021-11-92]. Dostupné z: https://dspace.cuni.cz/bitstream/handle/20.500.11956/69815/BPTX_2013_2_11230_0_388677_0_150987.pdf?sequence=1&isAllowed=y. Bakalářská práce. Univerzita Karlova. Vedoucí práce Mgr. Peter Hanák.

Zákony, novely, usnesení a nálezy

IV.ÚS 1146/16 ze dne 20. 6. 2017: N 101/85 SbNU 679. *Ústavní soud* [online]. 2017 [cit. 2021-11-02]. Dostupné z: https://nalus.usoud.cz/Search/GetText.aspx?sz=4-1146-16_2

Další nálezy Ústavního soudu k poskytování informací o platech zaměstnanců. *Ústavní soud* [online]. 2018, 2018 [cit. 2021-11-02]. Dostupné z: <https://www.usoud.cz/aktualne/dalsi-nalez-ustavniho-soudu-k-poskytovani-informaci-o-platech-zamestnancu>

Usnesení č. 2/1993 Sb.: Usnesení předsednictva České národní rady o vyhlášení LISTINY ZÁKLADNÍCH PRÁV A SVOBOD jako součástí ústavního pořádku České republiky. *Zákony pro lidi* [online]. [cit. 2021-11-22]. Dostupné z: <https://www.zakonyprolidi.cz/cs/1993-2?text=%C3%9Austavn%C3%AD+z%C3%A1kon+%C4%8D.+2%2F1993+Sb>

Zákon č. 2/1969 Sb.: Zákon České národní rady o zřízení ministerstev a jiných ústředních orgánů státní správy České socialistické republiky. *Zákony pro lidi* [online]. [cit. 2021-11-02]. Dostupné z: <https://www.zakonyprolidi.cz/cs/1969-2>

Zákon č. 106/1999 Sb.: Zákon o svobodném přístupu k informacím. *Zákony pro lidi* [online]. [cit. 2021-11-09]. Dostupné z: <https://www.zakonyprolidi.cz/cs/1999-106?text=svobodn%C3%A9m+p%C5%99%C3%ADstupu>

Podklady případové studie

BOČEK, Jan. Jak přesný je vládní model šíření koronaviru? Nikdo neví, data i parametry modelu úředníci tají. *IROZHLAS.cz* [online]. 2020, 17.04.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/komentare/koronavirus-otevrena-data-uzis_2004171327_jab

BOČEK, Jan. Ředitel ÚZIS by chtěl přestat vydávat denní počty obětí covidu. Není to dobrý nápad. *IROZHLAS.cz* [online]. 2020, 17.11.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/komentare/komentar-uzis-data-covid_2011170629_jab

CIBULKA, Jan. Ministerstvo tají informace o počtu nakažených v obcích. ‚Musíme uklidňovat situaci,‘ tvrdí. *IROZHLAS.cz* [online]. 2020, 23.09.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-uzis-covid-19-cisla-informace_2009231251_cib

CIBULKA, Jan. Nejenom Praha. Podrobnější čísla o epidemii ukazují na ohniska na Náchodsku a Uherskohradištsku. *IROZHLAS.cz* [online]. 2020, 02.10.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-covid-nemoc-luzka-obsazenost-statistiky-pravo-na-informace_2010031933_cib

CIBULKA, Jan. Praha jako první zveřejnila podrobná data o epidemii, nejvíc v přepočtu na obyvatele hlásí Praha 3 a Praha 2. *IROZHLAS.cz* [online]. 2020, 17.09.2020 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-covid-19-koronavir-cisla-statistiky-uzis-dusek_2009171321_cib

CIBULKA, Jan. Soud rozhodl: čísla o pandemii laikům do rukou nepatří. *IROZHLAS.cz* [online]. 2021, 12.04.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/komentare/data-covid-statistika-soud-rozudek-praco-na-informace_2104120629_cib

CIBULKA, Jan. Tajení informací o zdravotnictví je protizákonné, uvedl městský soud. Navrhuje zrušit část zákona. *IROZHLAS.cz* [online]. 2021, 08.06.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/informace-behounek-uzis-mzcr_2106081139_cib

CIBULKA, Jan. Zdravotníci statistici odmítají zveřejnit podrobné informace o očkovaných, uvedou jen týdenní součty. *IROZHLAS.cz* [online]. 2021, 05.01.2021 [cit. 2021-11-22]. Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-ockovani-covid-uzis-informace_2101051741_cib

CIBULKA, Jan. Zveřejňování čísel o očkování vážne. Ministerstvo s jejich poskytnutím otálí, obchází při tom zákon. *IROZHLAS.cz* [online]. 2021, 13.01.2021 [cit. 2021-11-22].

Dostupné z: https://www.irozhlas.cz/zpravy-domov/data-ockovani-zakon-106-pravo-na-informace-isin-covid-koronavirus-vakciny_2101131430_cib

MAHDALOVÁ, Kateřina. Data uniklá z ÚZIS: Praha v počtu vakcín válcuje ostatní kraje. *Seznam Zprávy.cz* [online]. 2021, 20.01.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/data-unikla-z-uzis-praha-v-poctu-vakcin-valcuje-ostatni-kraje-138615>

MAHDALOVÁ, Kateřina. Jak se čaruje s daty: Hygieny znají více nakažených než státní statistika. *Seznam Zprávy.cz* [online]. 2020, 01.11.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/jak-se-caruje-s-daty-hygieny-znaji-vice-nakazenych-nez-statni-statistika-126980>

MAHDALOVÁ, Kateřina. Jedinečná mapa obcí: Neveřejná data odhalují, jak postupuje očkování. *Seznam Zprávy.cz* [online]. 2021, 02.09.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/jedinecna-mapa-obci-neverejna-data-odhaluji-jak-postupuje-ockovani-152402>

MAHDALOVÁ, Kateřina. Koronavirová krize: Vláda se rozhoduje podle neúplných dat. *Seznam Zprávy.cz* [online]. 2020, 14.10.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/koronavirova-krize-vlada-se-rozhoduje-podle-neuplnych-dat-124235>

MAHDALOVÁ, Kateřina. Máte 107 nakažených, pardon, jednoho. Špatná data „zamořila“ menší obce. *Seznam Zprávy.cz* [online]. 2020, 04.11.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/mate-107-nakazenych-pardon-jednoho-spatna-data-zamorila-mensi-obce-127568>

MAHDALOVÁ, Kateřina. Ministr Blatný neříká pravdu. Vakcíny se stále rozdělují nerovnoměrně. *Seznam Zprávy.cz* [online]. 2021, 22.01.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/ministr-blatny-nerika-pravdu-vakciny-se-stale-rozdeluji-nerovnomerne-138890>

MAHDALOVÁ, Kateřina. Neveřejná data z českých nemocnic. Kdo tam bojuje s covidem. *Seznam Zprávy.cz* [online]. 2021, 12.03.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/data-v-cesku-je-dvakrat-vice-pacientu-v-tezkem-stavu-nez-pocatkem-roku-146945>

MAHDALOVÁ, Kateřina. PES a chaos v datech: testy ukazují jiná čísla. *Seznam Zprávy.cz* [online]. 2020, 24.11.2020 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/jak-ministersky-pes-zastekal-a-onemocnel-130982>

MAHDALOVÁ, Kateřina. Stát nemá přehled o tom, jak se šíří covid ve školách. *Seznam Zprávy.cz* [online]. 2021, 04.11.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/ve-kterych-skolach-se-siri-covid-prehled-nema-ministerstvo-ani-mnohe-hygieny-179532>

MAHDALOVÁ, Kateřina. Unikátní data: Jak virus postupuje školami. Jsou tři rizikové kraje. *Seznam Zprávy.cz* [online]. 2021, 18.06.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/unikatni-data-virus-ze-skol-mizi-zbyvaji-tri-rizikove-kraje-167590>

MAHDALOVÁ, Kateřina. Vláda čte špatně vlastní data: uzavření okresů nemá na šíření viru vliv. *Seznam Zprávy.cz* [online]. 2021, 26.03.2021 [cit. 2021-11-22]. Dostupné z: <https://www.seznamzpravy.cz/clanek/vlada-cte-spatne-vlastni-data-uzavreni-okresu-nema-na-sireni-viru-vliv-148532>

V České republice jsou první tři potvrzené případy nákazy koronavirem: Tisková zpráva. *Ministerstvo zdravotnictví České republiky* [online]. 2020, 01.03.2020 [cit. 2021-11-22]. Dostupné z: <https://koronavirus.mzcr.cz/v-ceske-republice-jsou-prvni-tri-potvrzene-pripady-nakazy-koronavirem/>

Institut komunikačních studií a žurnalistiky FSV UK
Teze MAGISTERSKÉ diplomové práce

TUTO ČÁST VYPLŇUJE STUDENT/KA:

Příjmení a jméno diplomantky/diplomanta:

Krawiecová Nela

Razítko podatelny:

Imatrikulační ročník diplomantky/diplomanta:
2019

E-mail diplomantky/diplomanta:

nela.krawiecova@gmail.com

Studijní obor/forma studia:

Žurnalistika/prezenční

Univerzita Karlova	
Fakulta sociálních věd	
Došlo dne:	15 -03- 2021 -1-
Čj: BF	Příloh:
<small>PP. číslo:</small>	

Název práce v češtině:

Otevřená data a jejich využití v tuzemské datové žurnalistice

Název práce v angličtině:

Open Data and their use in Czech Data Journalism

Předpokládaný termín dokončení (semestr, akademický rok – vzor: ZS 2012/2013)

(diplomovou práci je možné odevzdat nejdříve po dvou semestrech od schválení tezí)

ZS 2021/2022

Charakteristika tématu a jeho dosavadní zpracování (max. 1800 znaků):

Základním pramenem datové žurnalistiky, která se v posledních letech stala vyhledávaným směrem nejen v tuzemských redakcích, je dostupnost použitelných dat. Neexistuje však žádná ucelená práce, která v rámci žurnalistického diskurzu zkoumá možnosti volně dostupných dat a získávání informací, které přístupné doposud nejsou.

Nezbytnou součástí práce je tedy mapování historického vývoje otevírání tuzemských dat státní správou a s ní spojená legislativa, popis současné situace, a následná analýza rutin a dostupnosti datových zdrojů.

Práce by však kromě teoretické části měla přinést především praktickou ukázkou získávání, čištění a analýzy dat pro účely zpracování ve zpravodajství či publicistice, a to na příkladu dat covid-19 poskytovaných Ústavem zdravotnických informací a statistiky během pandemie v letech 2020-2021.

Předpokládaný cíl práce, případně formulace problému, výzkumné otázky nebo hypotézy (max. 1800 znaků):

Cílem práce je poskytnout ucelený přehled o využitelnosti a dostupnosti datových sad, které pro svou práci mohou využívat tuzemští novináři. Cílem práce je rovněž uvést na konkrétních případech možnosti využití Zákona o svobodném přístupu k informacím a oslovit novináře, kteří s otevřenými daty pracují v rámci svých redakčních rutin. Na základě hloubkových rozhovorů práce přinese vhled do těchto rutin a pokusí se najít limity, na které novináři narážejí. V praktické části pak tyto poznatky budou využity k analýze a zpracování pandemických dat covid-19.

Předpokládaná struktura práce (rozdělení do jednotlivých kapitol a podkapitol se stručnou charakteristikou jejich obsahu):

1. Úvod

- 1.1 vymezení základních pojmů, ukotvení datové žurnalistiky v tuzemském mediálním prostoru
- 1.2 stanovení cíle práce a bližší úvod do problematiky dostupnosti dat

2. Teoretická část

- 2.1 definice otevřených dat
- 2.2 zákonné povinnosti, historie legislativy v oblasti otevřených dat
- 2.3 příklady využití Zákona o svobodném přístupu k informacím v žurnalistice
- 2.4 srovnání se zahraničím

3. Metodologie

- 3.1 charakteristika použitých metod zkoumání
- 3.2 formulace výzkumných otázek pro hloubkové rozhovory, očekávané rutiny práce s daty

4. Empirická část

- 4.1 analýza dostupnosti datových sad
- 4.2 hloubkové rozhovory s novináři pracujícími s daty
- 4.3 praktický příklad práce s otevřenými daty
- 4.4 vyhodnocení a interpretace

5. Závěr

- 5.1 shrnutí výsledků

Vymezení podkladového materiálu (např. titul periodika a analyzované období):

Teoretická část práce bude vycházet z odborné literatury a relevantních článků, a to především zahraničních. Základem pro empirickou část práce bude analýza vybraných datových portálů a sad, doplněna o hloubkové rozhovory (cca s deseti datovými novinářkami/novináři - např. Českého rozhlasu).

Metody (techniky) zpracování materiálu:

Hlavními metodami práce budou kvantitativní a kvalitativní obsahová analýza, srovnání a dotazování.

Základní literatura (nejméně 5 nejdůležitějších titulů k tématu a metodě jeho zpracování; u všech titulů je nutné uvést stručnou anotaci na 2-5 řádků):

YOUNG, Andrew a Stefaan VERHULST. *The Global Impact of Open Data*. O'Reilly Media, 2016. ISBN 9781491964668.

Komplexní úvod do fenoménu otevřených dat, jejich vlivu na společenskou transformaci, včetně nových forem politické a vládní odpovědnosti. Součástí je hloubková analýza dostupnosti dat v několika vybraných státech, která může sloužit jako zdroj informací pro teoretickou část práce.

LEWIS, Seth. *Journalism in an Era of Big Data : Cases, concepts, and critiques*. Routledge, 2018. ISBN 9780367028411.

Knihla se rozsáhle zabývá trendy v digitální žurnalistice a zejména uvádí příklady práce s velkými objemy dat ve zpravodajství. Rovněž analyzuje, jak sílící potřeba zpracování dat mění postupy a role novinářů v novodobých mediálních organizacích.

HERZOG, David. *Data literacy: a user's guide*. Los Angeles: SAGE, 2016. ISBN 1483333469.

Publikace seznamuje čtenáře s datovou gramotností a zkoumá různé typy datových zdrojů a možnosti jejich využití pro následnou analýzu a vizualizaci dat v žurnalistice.

FUREK, Adam a Lukáš ROTHANZL. *Zákon o svobodném přístupu k informacím a související předpisy: komentář. 2. aktualizované a rozšířené vyd.* Praha: Linde Praha, 2012, 1031 s. ISBN 978-807-2018-680.

Komentář reaguje na přijatou českou legislativu v oblasti přístupu k informacím a otevřeným datům, včetně příkladů právních sporů, které vyústily k zveřejnění datových sad později využitých v žurnalistice.

HENDL, Jan, 2016. *Kvalitativní výzkum: základní teorie, metody a aplikace*. Čtvrté, přepracované a rozšířené vydání. Praha: Portál. ISBN 978-80-262-0982-9.

Základní příručka pro studium kvalitativních metod sběru dat, jejich kódování, vyhodnocování a interpretaci včetně příkladů zpracování získaných dat.

Diplomové a disertační práce k tématu (seznam bakalářských, magisterských a doktorských prací, které byly k tématu obhájeny na UK, případně dalších oborově blízkých fakultách či vysokých školách za posledních pět let)

Dostupností a využitelností otevřených dat v žurnalistice se samostatně nezabývala žádná závěrečná práce. Fenomén otevřených dat však pronikl do závěrečných prací jiných oborů a fakult, které zmiňuji níže.

MAREK, Jiří. *Otevřená data v akademické praxi*. Brno, 2019. Diplomová práce. Masarykova univerzita, Právnická fakulta. Vedoucí práce JUDr. Matěj Myška, Ph.D.

VAMBERSKÁ, Dominika. *Otevřená data ve veřejné správě*. Brno, 2017. Diplomová práce. Masarykova univerzita, Právnická fakulta. Vedoucí práce JUDr. Matěj Myška, Ph.D.

ŠTIRBA, Dominik. *Otevřená právní data*. Brno, 2019. Bakalářská práce. Masarykova univerzita, Fakulta informatiky. Vedoucí práce JUDr. Jakub Harašta, Ph.D.

HALAMOVÁ, Anna. *Vývoj a současný stav datové žurnalistiky v České republice*. Olomouc, 2020. Bakalářská práce. Univerzita Palackého v Olomouci, Filozofická fakulta. Vedoucí práce Mgr. et Mgr. Karel Páral.

KOMÁRKOVÁ, Kateřina. *Otevřená data a jejich praktické využití*. Pardubice, 2017. Bakalářská práce. Univerzita Pardubice, Fakulta ekonomicko-správní. Vedoucí práce a JUDr. Matěj Myška, Ph.D.

Datum / Podpis studenta/ky

10.03.2021



TUTO ČÁST VYPLŇUJE PEDAGOG/PEDAGOŽKA:

Doporučení k tématu, struktuře a technice zpracování materiálu:

Případné doporučení dalších titulů literatury předepsané ke zpracování tématu:

Potvrzuji, že výše uvedené teze jsem s jejich autorem/kou konzultoval(a) a že téma odpovídá mému oborovému zaměření a oblasti odborné práce, kterou na FSV UK vykonávám.

Souhlasím s tím, že budu vedoucí(m) této práce.

PhDr. Václav MORAVEC, Ph.D.

15. března 2021

Příjmení a jméno pedagožky/pedagoga

Datum / Podpis pedagožky/pedagoga



TEZE JE NUTNO ODEVZDAT VYTIŠTĚNÉ, PODEPSANÉ A VE DVOU VYHOTOVENÍCH DO TERMÍNU UVEDENÉHO V HARMONOGRAMU PŘÍSLUŠNÉHO AKADEMICKÉHO ROKU, A TO PROSTŘEDNICTVÍM PODATELNÝ FSV UK. PŘIJATÉ TEZE JE NUTNÉ SI VYZVEDNOUT V SEKRETARIÁTU PŘÍSLUŠNÉ KATEDRY A NECHAT VEVÁZAT DO OBOU VÝTIŠKŮ DIPLOMOVÉ PRÁCE.

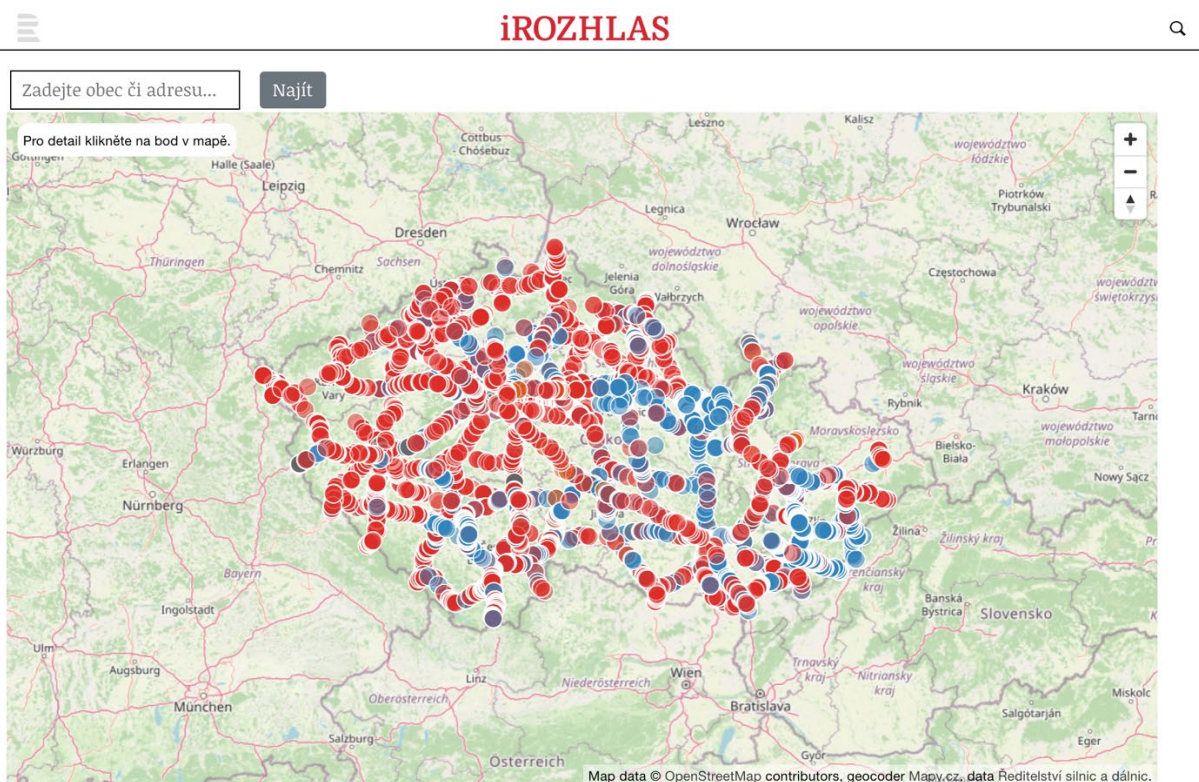
TEZE NA IKSŽ SCHVALUJE GARANT PŘÍSLUŠNÉHO STUDIJNÍHO OBORU.

Seznam příloh

- Příloha č. 1: Interaktivní mapa: vizualizace ilegálních poutačů u silnic (obrázek)
- Příloha č. 2: Ukázka datové žurnalistiky na iROZHLAS.cz (obrázek)
- Příloha č. 3: Ukázka datové žurnalistiky na Deník.cz (obrázek)
- Příloha č. 4: Ukázka datové žurnalistiky na Seznam Zprávy.cz (obrázek)
- Příloha č. 5: Cibulka na základě získaných dat vytvořil interaktivní mapy, které se aktualizují (obrázek)
- Příloha č. 6: Mahdalová vizualizovala počty hospitalizovaných pacientů (obrázek)
- Příloha č.7: Znění Zákona č.106/1999 Sb. (text)
- Příloha č. 8: Přepis hloubkového rozhovoru s Janem Bočkem (text)
- Příloha č. 9: Přepis hloubkového rozhovoru s Janem Cibulkou (text)
- Příloha č. 10: Přepis hloubkového rozhovoru s Petrem Kočím (text)
- Příloha č. 11: Přepis hloubkového rozhovoru s Alešem Vojířem (text)
- Příloha č. 12: Přepis hloubkového rozhovoru s Kateřinou Mahdalovou (text)

Přílohy

Příloha č. 1: Interaktivní mapa: vizualizace ilegálních poutačů u silnic (obrázek)

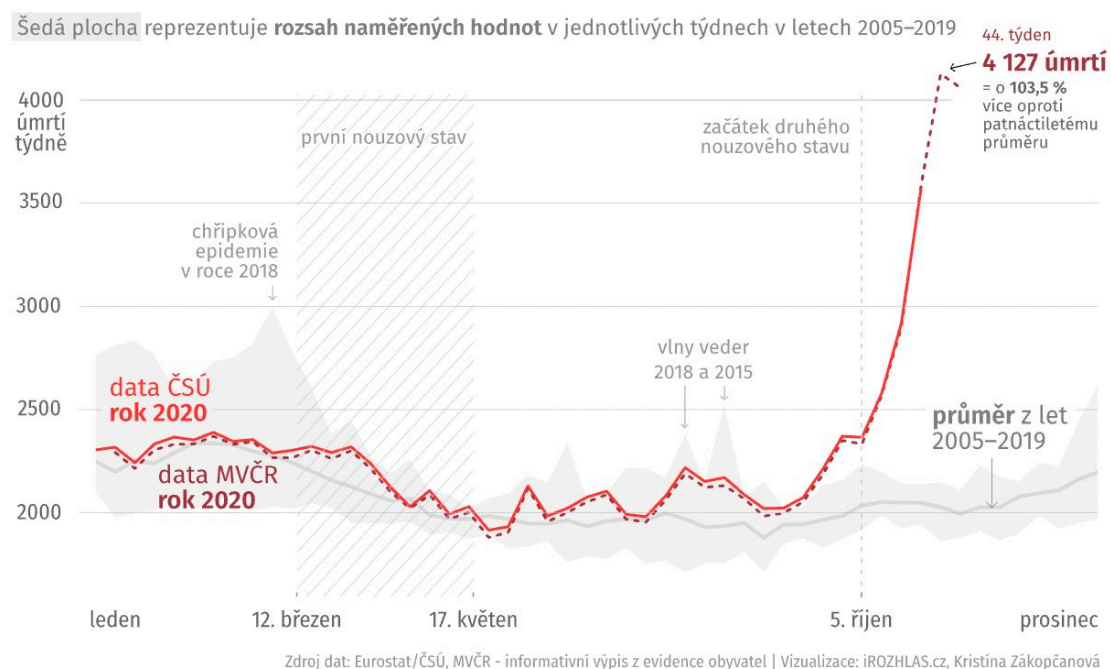


Autoři: Cibulka a Krawiecová (2020), zdroj: https://www.irozhlas.cz/zpravy-domov/billboardy-reklama-poutace-dalnice-d1-ministerstvo-dopravy-reditelstvi-silnic-a_2003020600_nkr

Příloha č. 2: Ukázka datové žurnalistiky na iROZHLAS.cz (obrázek)

Koronavirus počtem obětí překonal velké chřipkové epidemie i vlny veder

Počty zemřelých po týdnech v České republice mezi lety 2005–2020

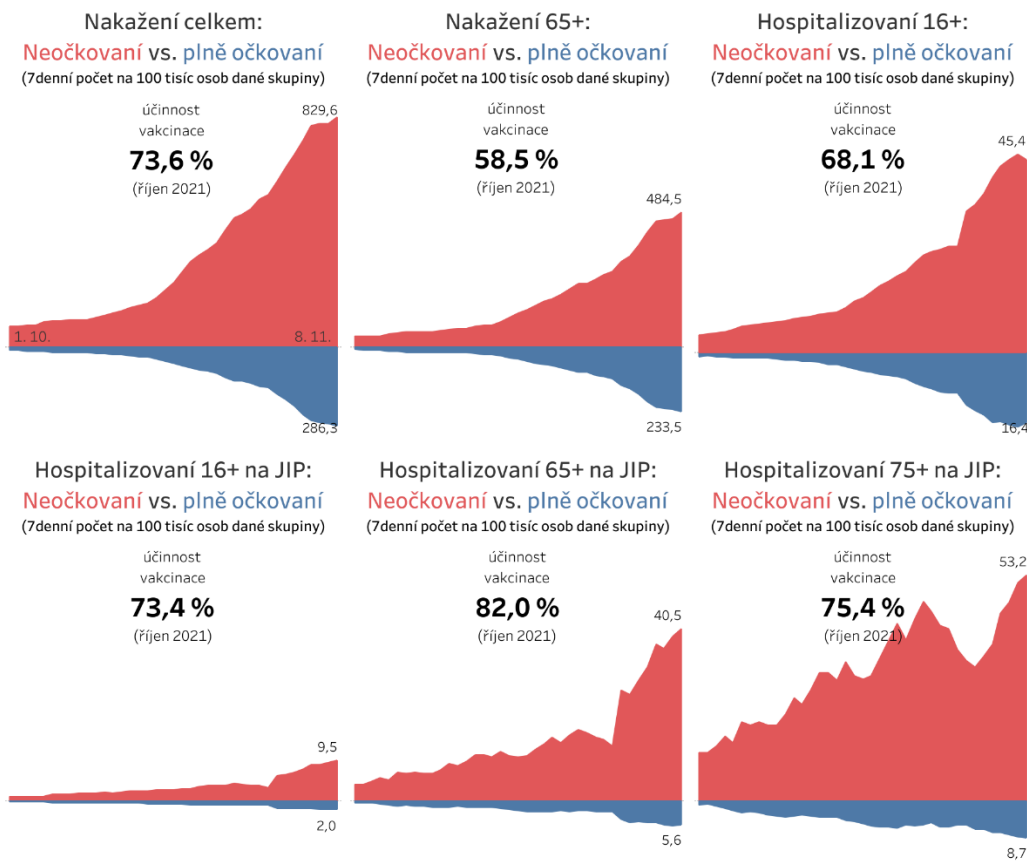


Autor: Kočí (2020), zdroj: https://www.irozhlas.cz/zpravy-domov/celkova-umrti-listopad_2012070635_pek

Příloha č. 3: Ukázka datové žurnalistiky na Deník.cz (obrázek)

Jak funguje očkování proti covidu v Česku

(data za jednotlivé dny října a listopadu 2021)



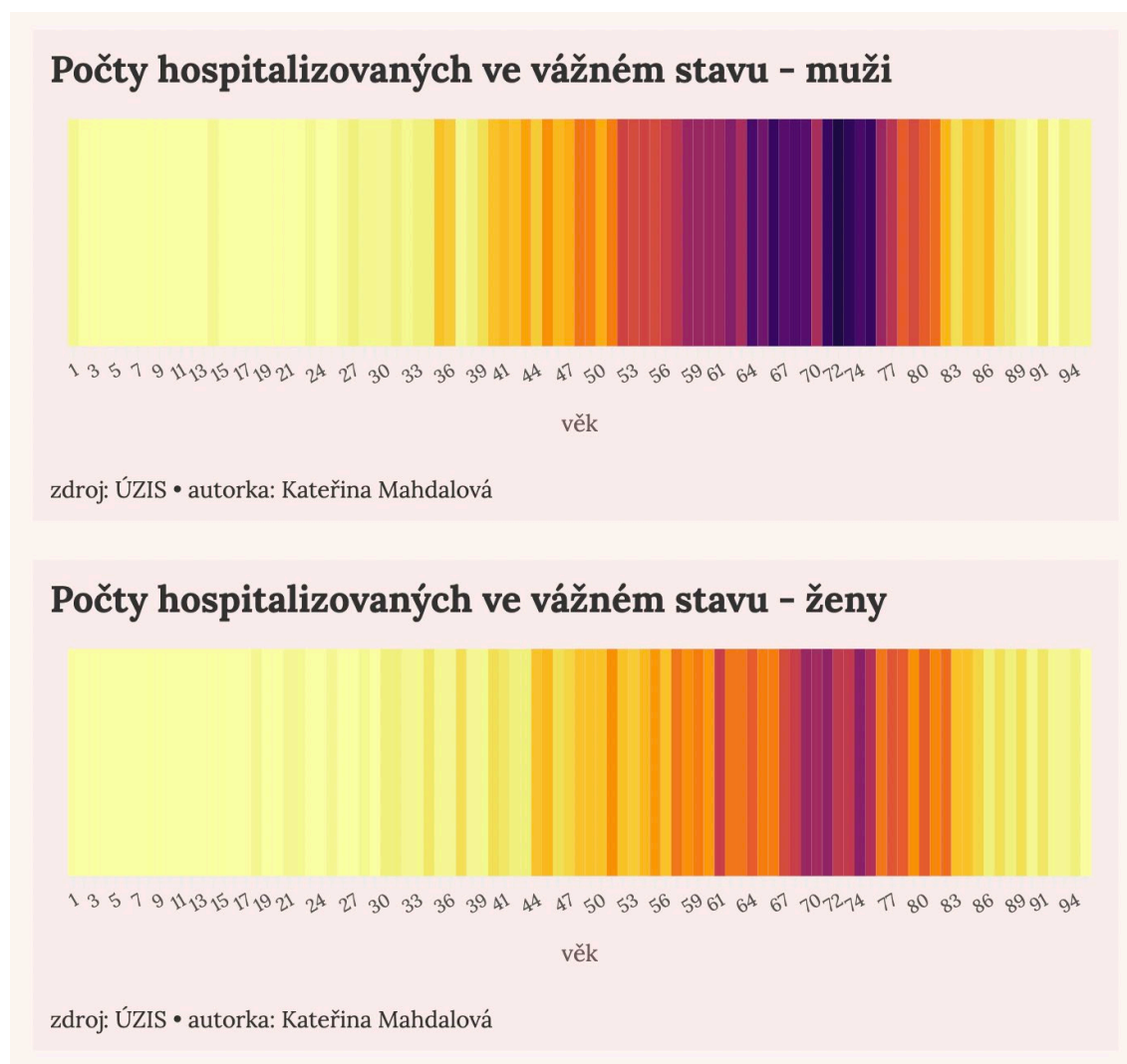
Zdroj: ÚZIS

Infografika: Aleš Vojíř, Deník Data



Autor: Vojíř (2020), zdroj: <https://data.denik.cz/cesi-v-cislech/ockovani-proti-koronaviru-funguje.html>

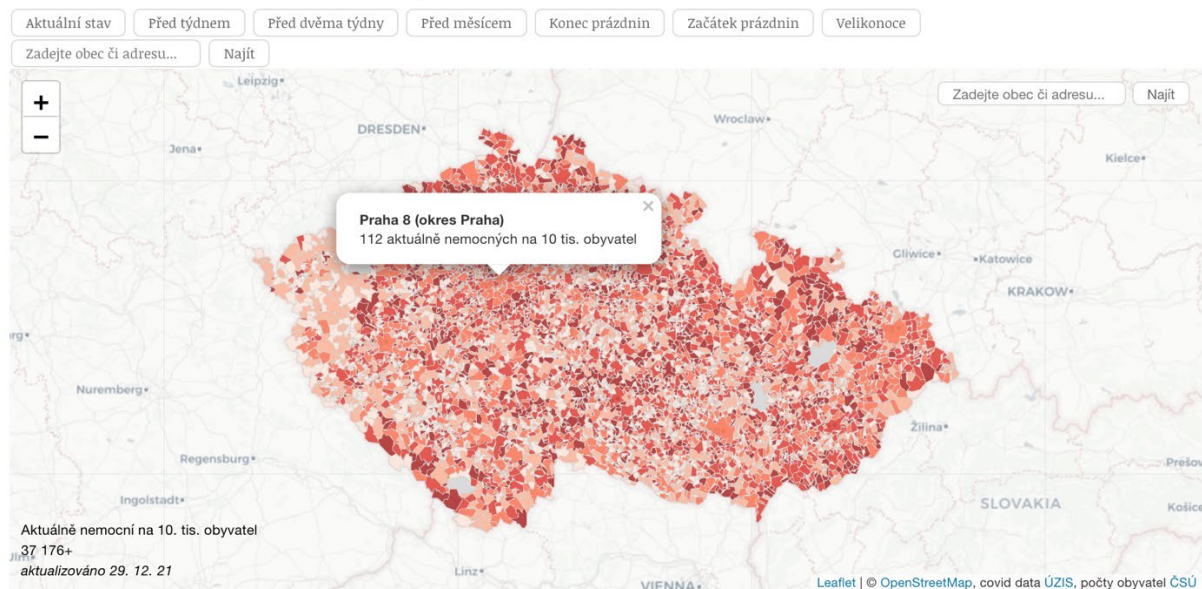
Příloha č. 4: Ukázka datové žurnalistiky na Seznam Zprávy.cz (obrázek)



Autor: Mahdalová (2020), zdroj:<https://www.seznamzpravy.cz/clanek/data-v-cesku-je-dvakrat-vice-pacientu-v-tezkem-stavu-nez-pocatkem-roku-146945>

Příloha č. 5: Cibulka na základě získaných dat vytvořil interaktivní mapy, které se aktualizují (obrázek)

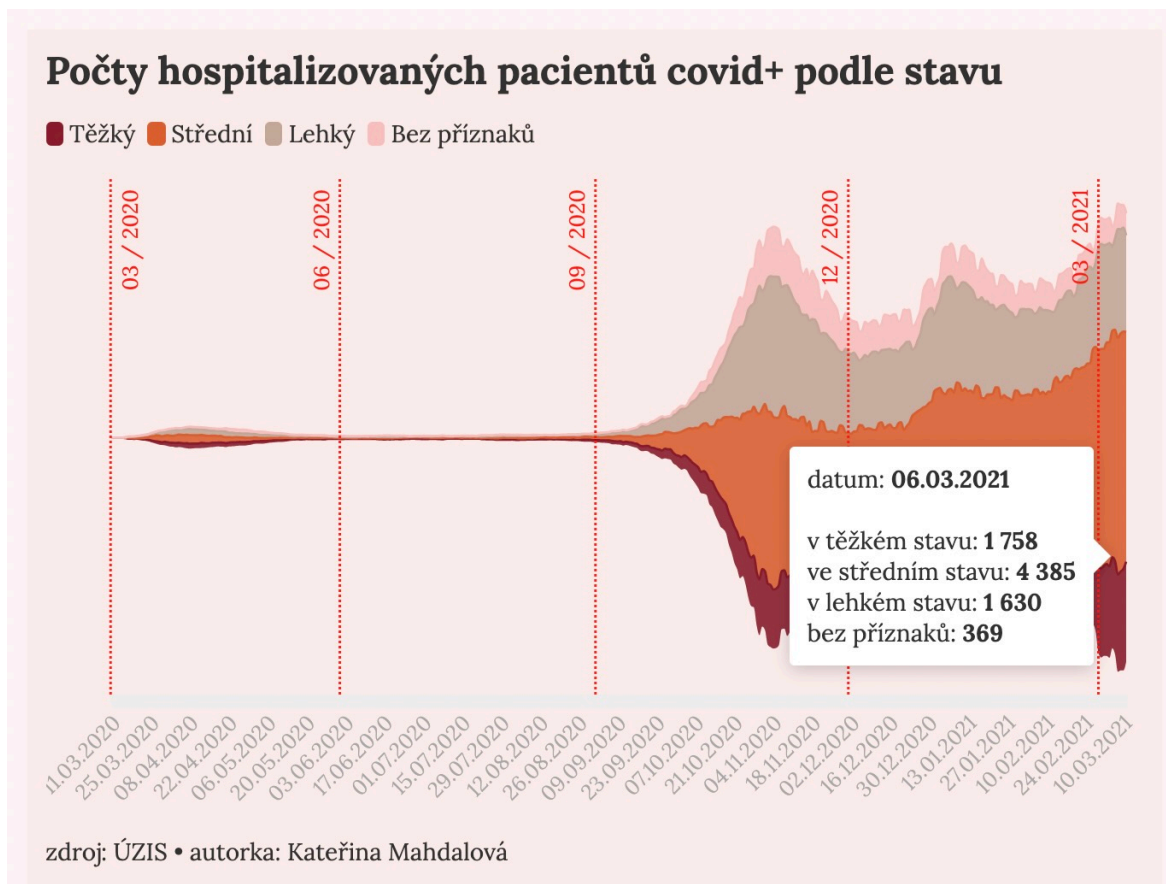
Kolik lidí má testem potvrzený covid-19 ve vaší obci



Kliknutím na obec v mapě zobrazíte podrobnosti.

Autor: Cibulka (2020), zdroj: https://www.irozhlas.cz/zpravy-domov/koronavirus-rychlost-sireni-obce-okresy-aplikace-mapa_2010260600_pek

Příloha č. 6: Mahdalová vizualizovala počty hospitalizovaných pacientů (obrázek)



Autorka: Mahdalová (2020), zdroj: <https://www.seznamzpravy.cz/clanek/data-v-cesku-je-dvakrat-vice-pacientu-v-tezkem-stavu-nez-pocatkem-roku-146945>

Příloha č.7: Celé znění Zákona č. 106/1999 Sb. o svobodném přístupu k informacím (text)¹⁴¹

106

ZÁKON

ze dne 11. května 1999

o svobodném přístupu k informacím

Parlament se usnesl na tomto zákoně České republiky:

ČÁST PRVNÍ

¹⁴¹ Zákon č. 106/1999 Sb.: Zákon o svobodném přístupu k informacím. *Zákony pro lidi* [online]. [cit. 2021-11-09]. Dostupné z: <https://www.zakonyprolidi.cz/cs/1999-106?text=svobodn%C3%A9m+p%C5%99%C3%ADstupu>

§ 1

Tento zákon zpracovává příslušné předpisy Evropské unie¹⁾ a upravuje pravidla pro poskytování informací a dále upravuje podmínky práva svobodného přístupu k těmto informacím.

§ 2

Povinnost poskytovat informace

- (1) Povinnými subjekty, které mají podle tohoto zákona povinnost poskytovat informace vztahující se k jejich působnosti, jsou státní orgány, územní samosprávné celky a jejich orgány a veřejné instituce.
- (2) Povinnými subjekty jsou dále ty subjekty, kterým zákon svěřil rozhodování o právech, právem chráněných zájmech nebo povinnostech fyzických nebo právnických osob v oblasti veřejné správy, a to pouze v rozsahu této jejich rozhodovací činnosti.
- (3) Zákon se nevztahuje na poskytování informací o údajích vedených v centrální evidenci účtů a v navazujících evidencích, informací, které jsou předmětem průmyslového vlastnictví^{1a)}, a dalších informací, pokud zvláštní zákon^{1b)} upravuje jejich poskytování, zejména vyřízení žádosti včetně náležitostí a způsobu podání žádosti, lhůt, opravných prostředků a způsobu poskytnutí informací.
- (4) Povinnost poskytovat informace se netýká dotazů na názory, budoucí rozhodnutí a vytváření nových informací.

§ 3

Základní pojmy

- (1) Žadatelem pro účel tohoto zákona je každá fyzická i právnická osoba, která žádá o informaci.
- (2) Možností dálkového přístupu pro účel tohoto zákona je přístup k informaci neomezeného okruhu žadatelů pomocí sítě nebo služby elektronických komunikací²⁾.
- (3) Informací se pro účely tohoto zákona rozumí jakýkoliv obsah nebo jeho část v jakémkoliv podobě, zaznamenaný na jakémkoliv nosiči, zejména obsah písemného záznamu na listině, záznamu uloženého v elektronické podobě nebo záznamu zvukového, obrazového nebo audiovizuálního.
- (4) Informací podle tohoto zákona není počítačový program.
- (5) Zveřejněnou informací pro účel tohoto zákona je taková informace, která může být vždy znovu vyhledána a získána, zejména vydaná tiskem nebo na jiném nosiči dat umožňujícím zápis a uchování informace, vystavená na úřední desce, s možností

dálkového přístupu nebo umístěná v knihovně poskytující veřejné knihovnické a informační služby podle knihovního zákona^{2a}).

(6) Doprovodnou informací pro účel tohoto zákona je taková informace, která úzce souvisí s požadovanou informací (například informace o její existenci, původu, počtu, důvodu odepření, době, po kterou důvod odepření trvá a kdy bude znovu přezkoumán, a dalších důležitých rysech).

(7) Strojově čitelným formátem se pro účely tohoto zákona rozumí formát datového souboru s takovou strukturou, která umožňuje programovému vybavení snadno nalézt, rozpoznat a získat z tohoto datového souboru konkrétní informace, včetně jednotlivých údajů a jejich vnitřní struktury.

(8) Otevřeným formátem se pro účely tohoto zákona rozumí formát datového souboru, který není závislý na konkrétním technickém a programovém vybavení a je zpřístupněn veřejnosti bez jakéhokoli omezení, které by znemožňovalo využití informací obsažených v datovém souboru.

(9) Otevřenou formální normou se pro účely tohoto zákona rozumí pravidlo, které bylo vydáno písemně a obsahuje specifikace požadavků na zajištění schopnosti různých programových vybavení vzájemně si poskytovat služby a efektivně spolupracovat.

(10) Metadata jsou pro účely tohoto zákona data popisující souvislosti, obsah a strukturu zaznamenaných informací a jejich správu v průběhu času.

(11) Otevřenými daty se pro účely tohoto zákona rozumí informace zveřejňované způsobem umožňujícím dálkový přístup v otevřeném a strojově čitelném formátu, jejichž způsob ani účel následného využití není omezen a které jsou evidovány v národním katalogu otevřených dat.

§ 4

Poskytování informací

(1) Povinné subjekty poskytují informace na základě žádosti nebo zveřejněním.

(2) V případě, že je žadatelem povinný subjekt, je mu poskytována informace za stejných podmínek jako jiným žadatelům.

§ 4a

Poskytování informací na žádost

(1) Je-li informace poskytována na základě žádosti, poskytuje se ve formátech a jazycích podle obsahu žádosti o poskytnutí informace, včetně k ní se vztahujících metadat, pokud tento zákon nestanoví jinak. Povinný subjekt není povinen měnit formát nebo jazyk

informace ani vytvářet k informaci metadata, pokud by taková změna nebo vytvoření metadat byly pro povinný subjekt nepřiměřenou zátěží; v tomto případě vyhoví povinný subjekt žádosti tím, že poskytne informaci ve formátu nebo jazyce, ve kterých byla vytvořena. Pokud je požadovaná informace součástí většího celku a její vynětí by bylo pro povinný subjekt nepřiměřenou zátěží, poskytne povinný subjekt takový celek v souladu s tímto zákonem. Pokud je to možné s přihlédnutím k povaze podané žádosti a způsobu záznamu požadované informace, poskytne povinný subjekt informaci v elektronické podobě.

(2) Je-li informace poskytována na základě žádosti, poskytuje se způsobem podle obsahu žádosti, zejména

- a) sdělením informace v elektronické nebo listinné podobě,
 - b) poskytnutím kopie dokumentu obsahujícího požadovanou informaci,
 - c) poskytnutím datového souboru obsahujícího požadovanou informaci,
 - d) nahlédnutím do dokumentu obsahujícího požadovanou informaci,
 - e) sdílením dat prostřednictvím rozhraní informačního systému, nebo
 - f) umožněním dálkového přístupu k informaci, která se v průběhu času mění, obnovuje, doplňuje nebo opakovaně vytváří, nebo jejím pravidelným předáváním jiným způsobem.
- (3) Pokud způsob poskytnutí informace podle odstavce 2 není možný nebo by pro povinný subjekt představoval nepřiměřenou zátěž, vyhoví povinný subjekt žádosti tím, že poskytne informaci jiným způsobem umožňujícím její účinné využití žadatelem.

§ 4b

Poskytování informací zveřejněním

Informace poskytovaná zveřejněním se poskytuje ve všech formátech a jazycích, ve kterých byla vytvořena; při zveřejnění takové informace v elektronické podobě musí být jeden z těchto formátů otevřený a, je-li to možné, též strojově čitelný. Je-li to možné a vhodné, zveřejní povinný subjekt spolu s informací též metadata, která se k ní vztahují.

Formát i metadata by měly co nejlépe splňovat otevřené formální normy.

§ 4c

Národní katalog otevřených dat

(1) Národní katalog otevřených dat je informační systém veřejné správy přístupný způsobem umožňujícím dálkový přístup sloužící k evidování informací zveřejňovaných jako otevřená data.

(2) Správcem národního katalogu otevřených dat je Ministerstvo vnitra.

§ 5

Zveřejňování informací

(1) Každý povinný subjekt musí pro informování veřejnosti ve svém sídle a svých úřadovnách zveřejnit na místě, které je všeobecně přístupné, jakož i umožnit pořízení jejich kopie, tyto informace:

- a) důvod a způsob založení povinného subjektu, včetně podmínek a principů, za kterých provozuje svoji činnost,
- b) popis své organizační struktury, místo a způsob, jak získat příslušné informace, kde lze podat žádost či stížnost, předložit návrh, podnět či jiné dožádání anebo obdržet rozhodnutí o právech a povinnostech osob,
- c) místo, lhůtu a způsob, kde lze podat opravný prostředek proti rozhodnutím povinného subjektu o právech a povinnostech osob, a to včetně výslovného uvedení požadavků, které jsou v této souvislosti kladeny na žadatele, jakož i popis postupů a pravidel, která je třeba dodržovat při těchto činnostech, a označení příslušného formuláře a způsob a místo, kde lze takový formulář získat,
- d) přehled nejdůležitějších předpisů, podle nichž povinný subjekt zejména jedná a rozhoduje, které stanovují právo žádat informace a povinnost poskytovat informace a které upravují další práva občanů ve vztahu k povinnému subjektu, a to včetně informace, kde a kdy jsou tyto předpisy poskytnuty k nahlédnutí,
- e) sazebník úhrad za poskytování informací,
- f) výroční zprávu za předcházející kalendářní rok o své činnosti v oblasti poskytování informací (§ 18),
- g) výhradní licence poskytnuté podle § 14a odst. 4,
- h) usnesení nadřízeného orgánu o výši úhrad vydaná podle § 16a odst. 7,
- i) elektronickou adresu podatelny.

(2) Povinné subjekty jsou ve svém sídle povinny v úředních hodinách zpřístupnit

- a) seznamy hlavních dokumentů, zejména koncepční, strategické a programové povahy, které mohou být poskytnuty podle tohoto zákona včetně případných návrhů licenčních smluv podle § 14a,
- b) postup, který musí povinný subjekt podle § 2 odst. 1 dodržovat při vyřizování podání nebo podnětů, zveřejněný Ministerstvem vnitra na portálu veřejné správy v podobě popisu úkonů orgánu veřejné moci a úkonů subjektu, který při vykonávání těchto úkonů nepůsobí

jako orgán veřejné moci, vedeného v základním registru agend, orgánů veřejné moci, soukromoprávních uživatelů údajů a některých práv a povinností,

a to tak, aby do nich mohl každý nahlédnout a pořídit si opis, výpis nebo kopii.

(3) Do 15 dnů od poskytnutí informací na žádost povinný subjekt tyto informace zveřejní způsobem umožňujícím dálkový přístup. O informacích poskytnutých způsobem podle § 4a odst. 2 písm. e) a f), informacích poskytnutých v jiné než elektronické podobě, nebo mimořádně rozsáhlých elektronicky poskytnutých informacích postačí zveřejnit doprovodnou informaci vyjadřující jejich obsah.

(4) Povinné subjekty jsou povinny zveřejňovat informace uvedené v odstavci 1 a 2 též způsobem umožňujícím dálkový přístup. Tato povinnost se nevztahuje na povinné subjekty, které jsou pouze fyzickými osobami. V případě informací uvedených v odstavci 1 písm. b) a c) a odstavci 2 písm. a) a c) postačuje ke splnění této povinnosti uvedení odkazu na místo, kde jsou tyto informace již zveřejněny způsobem umožňujícím dálkový přístup.

(5) Povinný subjekt může informace podle odstavce 1 zveřejnit i dalšími způsoby a může zveřejnit i další informace, nestanoví-li zákon jinak.

§ 5a

Informace zveřejňované způsobem umožňujícím dálkový přístup

(1) Povinné subjekty, které na základě zákona vedou a spravují registry, evidence, seznamy nebo rejstříky obsahující informace, které jsou na základě zákona každému přístupné (dále jen „registr“), jsou povinny tyto informace zveřejňovat v přehledné formě způsobem umožňujícím i dálkový přístup.

(2) Povinné subjekty zveřejňují informace obsažené v registrech, s výjimkou dokumentů ze sbírky listin, jsou-li součástí registru, rovněž jako otevřená data. Podle věty první se nezveřejní jméno a příjmení, datum narození s výjimkou roku narození, rodné číslo a adresa místa trvalého pobytu nebo bydliště s výjimkou názvu obce; to neplatí, jsou-li tyto údaje uvedeny v registrech v souvislosti s podnikáním nebo jinou obdobnou výdělečnou činností nebo v souvislosti s členstvím fyzické osoby ve statutárním nebo jiném orgánu právnické osoby nebo s výkonem funkce statutárního orgánu nebo v souvislosti s postavením skutečného majitele podle zákona upravujícího evidenci skutečných majitelů.

§ 6

Odkaz na zveřejněnou informaci

- (1) Pokud žádost o poskytnutí informace směřuje k poskytnutí zveřejněné informace, může povinný subjekt co nejdříve, nejpozději však do sedmi dnů, místo poskytnutí informace sdělit žadateli údaje umožňující vyhledání a získání zveřejněné informace, zejména odkaz na internetovou stránku, kde se informace nachází.
- (2) Pokud žadatel trvá na přímém poskytnutí zveřejněné informace, povinný subjekt mu ji poskytne; to neplatí, pokud byla žádost o poskytnutí informace podána elektronicky a pokud je požadovaná informace zveřejněna způsobem umožňujícím dálkový přístup a žadateli byl sdělen odkaz na internetovou stránku, kde se informace nachází.

§ 7

Ochrana utajovaných informací

Je-li požadovaná informace v souladu s právními předpisy⁴⁾ označena za utajovanou informaci, k níž žadatel nemá oprávněný přístup, povinný subjekt ji neposkytne. Povinný subjekt neposkytne rovněž osobní údaje o osobě, která je držitelem osvědčení fyzické osoby pro přístup k utajovaným informacím pro stupeň utajení Přísně tajné a Tajné, pokud by to mohlo ohrozit ochranu utajovaných informací.

§ 8zrušeno

§ 8a

(1) Informace týkající se osobnosti, projevů osobní povahy, soukromí fyzické osoby a osobní údaje povinný subjekt poskytne jen v souladu s právními předpisy, upravujícími jejich ochranu^{4a)}.

(2) Povinný subjekt poskytne osobní údaje o veřejně činné osobě, funkcionáři nebo zaměstnanci veřejné správy, které vypovídají o jeho veřejné nebo úřední činnosti nebo o jeho funkčním nebo pracovním zařazení.

§ 8b

Příjemci veřejných prostředků

(1) Povinný subjekt poskytne základní osobní údaje^{4b)} o osobě, které poskytl veřejné prostředky.

(2) Ustanovení odstavce 1 se nevztahuje na poskytování veřejných prostředků podle zákonů v oblasti sociální, poskytování zdravotních služeb, hmotného zabezpečení v nezaměstnanosti, státní podpory stavebního spoření a státní pomoci při obnově území^{4c)}.

(3) Základní osobní údaje podle odstavce 1 se poskytnou pouze v tomto rozsahu: jméno, příjmení, rok narození, obec, kde má příjemce trvalý pobyt, výše, účel a podmínky poskytnutých veřejných prostředků.

§ 9

Ochrana obchodního tajemství

(1) Pokud je požadovaná informace obchodním tajemstvím⁶⁾, povinný subjekt ji neposkytne.

(2) Při poskytování informace, která se týká používání veřejných prostředků, se nepovažuje poskytnutí informace o rozsahu a příjemci těchto prostředků za porušení obchodního tajemství.

§ 10

Ochrana důvěrnosti majetkových poměrů

Informace o majetkových poměrech osoby, která není povinným subjektem, získané na základě zákonů o daních, poplatcích, penzijním nebo zdravotním pojištění anebo sociálním zabezpečení⁸⁾ povinný subjekt podle tohoto zákona neposkytne.

§ 11

Další omezení práva na informace

(1) Povinný subjekt může omezit poskytnutí informace, pokud:

- a) se vztahuje výlučně k vnitřním pokynům a personálním předpisům povinného subjektu,
- b) jde o novou informaci, která vznikla při přípravě rozhodnutí povinného subjektu, pokud zákon nestanoví jinak; to platí jen do doby, kdy se příprava ukončí rozhodnutím,
- c) jde o informaci poskytnutou Organizací Severoatlantické smlouvy nebo Evropskou unií, která je v zájmu bezpečnosti státu, veřejné bezpečnosti nebo ochrany práv třetích osob chráněna uvedenými původci označením "NATO UNCLASSIFIED" nebo "LIMITE" a v České republice je toto označení respektováno z důvodů plnění povinností vyplývajících pro Českou republiku z jejího členství v Organizaci Severoatlantické smlouvy nebo Evropské unii, pokud původce nedal k poskytnutí souhlas^{8a)},
- d) její poskytnutí významně nebo přímo ohrožuje účinnost bezpečnostního opatření stanoveného na základě zvláštního předpisu pro účel ochrany bezpečnosti osob, majetku a veřejného pořádku, nebo
- e) její poskytnutí významně nebo přímo ohrožuje výkon zahraniční služby při ochraně zájmů České republiky a jejích občanů v zahraničí²⁰⁾.

(2) Povinný subjekt informaci neposkytne, pokud:

- a) jde o informaci vzniklou bez použití veřejných prostředků, která byla předána osobou, již takovouto povinnost zákon neukládá, pokud nesdělila, že s poskytnutím informace souhlasí,

- b) ji zveřejňuje na základě zvláštního zákona⁹⁾ a v předem stanovených pravidelných obdobích až do nejbližšího následujícího období,
- c) by tím byla porušena ochrana práv třetích osob k předmětu práva autorského nebo práv souvisejících s právem autorským (dále jen „právo autorské“)^{2b)}, nebo
- d) jde o informaci, která se týká stability finančního systému¹⁸⁾.

(3) Informace, které získal povinný subjekt od třetí osoby při plnění úkolů v rámci kontrolní, dozorové, dohledové nebo obdobné činnosti prováděné na základě zvláštního právního předpisu¹¹⁾, podle kterého se na ně vztahuje povinnost mlčenlivosti anebo jiný postup chránící je před zveřejněním nebo zneužitím, se neposkytují. Povinný subjekt poskytne pouze ty informace, které při plnění těchto úkolů vznikly jeho činností.

(4) Povinné subjekty dále neposkytnou informace o

- a) probíhajícím trestním řízením, nebo týkající se trestního řízení, pokud by její poskytnutí ohrozilo či zmařilo účel trestního řízení, zejména zajištění práva na spravedlivý proces,
 - b) rozhodovací činnosti soudů s výjimkou rozsudků,
 - c) plnění úkolů zpravodajských služeb,¹²⁾ nebo o činnosti zpravodajských služeb, pokud by poskytnutí této informace ohrozilo plnění jejich úkolů či ochranu utajovaných informací, nebo o činnosti zpravodajských služeb, pokud by poskytnutí této informace ohrozilo plnění jejich úkolů či ochranu utajovaných informací,
 - d) přípravě, průběhu a projednávání výsledků kontrol v orgánech Nejvyššího kontrolního úřadu,
 - e) činnosti Finančního analytického úřadu podle zákona o některých opatřeních proti legalizaci výnosů z trestné činnosti a financování terorismu nebo podle zákona o provádění mezinárodních sankcí,
 - f) činnosti České národní banky v souvislosti s vedením centrální evidence účtů.
- Ustanovení zvláštních zákonů¹³⁾ o poskytování informací v uvedených oblastech tím nejsou dotčena.

(5) Povinný subjekt neposkytne informaci, která je předmětem ochrany práva autorského^{2b)}, je-li v držení

- a) provozovatelů rozhlasového nebo televizního vysílání, kteří toto vysílání provozují na základě zvláštních právních předpisů^{13a)},
- b) škol a školských zařízení, které jsou součástí vzdělávací soustavy podle školského zákona^{13b)} a podle zákona o vysokých školách^{13c)},

- c) Akademie věd České republiky a dalších veřejných institucí, které jsou příjemci nebo spolupříjemci podpory výzkumu a vývoje z veřejných prostředků podle zákona o podpoře výzkumu a vývoje^{13d}), nebo
- d) kulturních institucí hospodařících s veřejnými prostředky, jako jsou divadla, orchestry a další umělecké soubory, s výjimkou knihoven poskytujících veřejné knihovnické a informační služby podle knihovního zákona^{2a}) a muzeí a galerií poskytujících standardizované veřejné služby¹⁹).

Poskytování těchto informací v souladu se zvláštními předpisy tím není dotčeno.

- (6) Povinný subjekt neposkytne informaci o činnosti orgánů činných v trestním řízení nebo bezpečnostních sborů, která se týká předcházení, vyhledávání, odhalování nebo stíhání trestné činnosti nebo ochrany bezpečnosti osob, majetku a veřejného pořádku, pokud by její poskytnutí ohrozilo práva třetích osob anebo schopnost orgánů veřejné moci předcházet trestné činnosti, vyhledávat nebo odhalovat trestnou činnost, stíhat trestné činy nebo zajišťovat veřejný pořádek a bezpečnost České republiky.

§ 12

Podmínky omezení

Všechna omezení práva na informace provede povinný subjekt tak, že poskytne požadované informace včetně doprovodných informací po vyloučení těch informací, u nichž to stanoví zákon. Právo odepřít informaci trvá pouze po dobu, po kterou trvá důvod odepření. V odůvodněných případech povinný subjekt ověří, zda důvod odepření trvá.

§ 13

Žádost o poskytnutí informace

- (1) Žádost o poskytnutí informace se podává ústně nebo písemně, a to i prostřednictvím sítě nebo služby elektronických komunikací.
- (2) Není-li žadateli na ústně podanou žádost informace poskytnuta anebo nepovažuje-li žadatel informaci poskytnutou na ústně podanou žádost za dostačující, je třeba podat žádost písemně.
- (3) Ustanovení § 14 až 16a a § 18 platí pouze pro žádosti podané písemně.

§ 14

Postup při podávání a vyřizování písemných žádostí o poskytnutí informace

- (1) Žádost je podána dnem, kdy ji obdržel povinný subjekt.
- (2) Ze žádosti musí být zřejmé, kterému povinnému subjektu je určena, a že se žadatel domáhá poskytnutí informace ve smyslu tohoto zákona. Fyzická osoba uvede v žádosti

jméno, příjmení, datum narození, adresu místa trvalého pobytu nebo, není-li přihlášená k trvalému pobytu, adresu bydliště a adresu pro doručování, liší-li se od adresy místa trvalého pobytu nebo bydliště. Právnická osoba uvede název, identifikační číslo osoby, adresu sídla a adresu pro doručování, liší-li se od adresy sídla. Adresou pro doručování se rozumí též elektronická adresa.

(3) Je-li žádost učiněna elektronicky, musí být podána prostřednictvím elektronické adresy podatelny povinného subjektu, pokud ji povinný subjekt zřídil. Pokud elektronické adresy podatelny nejsou zveřejněny, postačí podání na jakoukoliv elektronickou adresu povinného subjektu.

(4) Neobsahuje-li žádost náležitosti podle odstavce 2 věty první a adresu pro doručování, případně není-li elektronická žádost podána podle odstavce 3, není žádostí ve smyslu tohoto zákona.

(5) Povinný subjekt posoudí žádost a:

- a) brání-li nedostatek údajů o žadateli podle odstavce 2 postupu vyřízení žádosti o informaci podle tohoto zákona, zejména podle § 14a nebo 15, vyzve žadatele ve lhůtě do 7 dnů ode dne podání žádosti, aby žádost doplnil; nevyhoví-li žadatel této výzvě do 30 dnů ode dne jejího doručení, žádost odloží,
- b) v případě, že je žádost nesrozumitelná, není zřejmé, jaká informace je požadována, nebo je formulována příliš obecně, vyzve žadatele ve lhůtě do sedmi dnů od podání žádosti, aby žádost upřesnil, neupřesní-li žadatel žádost do 30 dnů ode dne doručení výzvy, rozhodne o odmítnutí žádosti,
- c) v případě, že požadované informace se nevztahují k jeho působnosti, žádost odloží a tuto odůvodněnou skutečnost sdělí do 7 dnů ode dne doručení žádosti žadateli,
- d) nerozhodne-li podle § 15, poskytne informaci v souladu se žádostí ve lhůtě nejpozději do 15 dnů ode dne přijetí žádosti nebo ode dne jejího doplnění; je-li zapotřebí licence podle § 14a, předloží v této lhůtě žadateli konečnou licenční nabídku.

(6) O postupu při poskytování informace se pořídí záznam.

(7) Lhůtu pro poskytnutí informace podle odstavce 5 písm. d) může povinný subjekt prodloužit ze závažných důvodů, nejvýše však o deset dní. Závažnými důvody jsou:

- a) vyhledání a sběr požadovaných informací v jiných úřadovnách, které jsou oddělené od úřadovny vyřizující žádost,
- b) vyhledání a sběr objemného množství oddělených a odlišných informací požadovaných v jedné žádosti,

c) konzultace s jiným povinným subjektem, který má závažný zájem na rozhodnutí o žádosti, nebo mezi dvěma nebo více složkami povinného subjektu, které mají závažný zájem na předmětu žádosti.

Žadatel musí být o prodloužení lhůty i o jeho důvodech vždy prokazatelně informován, a to včas před uplynutím lhůty pro poskytnutí informace.

§ 14a

Některá ustanovení o licenční nebo podlicenční smlouvě při poskytování informací

(1) Má-li být informace, která je předmětem ochrany práva autorského^{2b}), poskytnuta na základě licenční nebo podlicenční smlouvy a majetková práva k předmětu ochrany práva autorského vykonává povinný subjekt, který není vyňat z povinnosti poskytovat informaci podle § 11 odst. 5, postupuje se ve věcech neupravených tímto zákonem podle autorského zákona^{2b}).

(2) Odměna za oprávnění informaci užít nesmí být vyšší než úhrada podle § 17, nestanoví-li jinak zvláštní právní předpis nebo licenční smlouva mezi povinným subjektem a tím, kdo právo nakládat s předmětem práva autorského na povinný subjekt převedl.

(3) Podmínky poskytnutí informace v licenční nebo podlicenční smlouvě (dále jen "licenční smlouva") musí umožňovat další užití informace žadatelem v souladu se žádostí, pokud tento zákon nestanoví jinak. Licence nebo podlicence (dále jen "licence") se poskytuje jako nevýhradní, s výjimkou případů podle odstavce 4.

(4) Povinný subjekt může poskytnout výhradní licenci pouze tehdy, je-li výhradní licence pro další šíření poskytované informace nezbytná a je-li to ve veřejném zájmu. Poskytne-li povinný subjekt výhradní licenci podle věty první, přezkoumá alespoň každé 3 roky trvání důvodů, na základě kterých byla výhradní licence poskytnuta; to neplatí v případě poskytnutí výhradní licence k digitalizaci kulturního zdroje, kdy povinný subjekt přezkoumá trvání důvodů, na základě kterých byla výhradní licence poskytnuta, v jedenáctém roce trvání, případně každých následujících 7 let. V případě poskytnutí výhradní licence k digitalizaci kulturního zdroje musí být povinnému subjektu bezplatně poskytnuta kopie digitální reprodukce kulturního zdroje, kterou bude možné po skončení trvání výhradní licence poskytnout podle tohoto zákona.

(5) Povinný subjekt zveřejní způsobem umožňujícím dálkový přístup vzor licenční smlouvy, která může být žadatelem přizpůsobena konkrétní žádosti a použita jako návrh na uzavření licenční smlouvy.

(6) Na ustanovení licenčních smluv uzavřených při poskytování informací podle tohoto zákona se nevztahuje ochrana obchodního tajemství.

§ 15

Rozhodnutí o odmítnutí žádosti

(1) Pokud povinný subjekt žádosti, byť i jen zčásti, nevyhoví, vydá ve lhůtě pro vyřízení žádosti rozhodnutí o odmítnutí žádosti, popřípadě o odmítnutí části žádosti (dále jen "rozhodnutí o odmítnutí žádosti"), s výjimkou případů, kdy se žádost odloží.

(2) Pokud nebylo žádosti vyhověno z důvodů ochrany obchodního tajemství podle § 9 nebo ochrany práv třetích osob k předmětu práva autorského podle § 11 odst. 2 písm. c), musí být v odůvodnění rozhodnutí uvedeno, kdo vykonává právo k tomuto obchodnímu tajemství nebo kdo vykonává majetková práva k tomuto předmětu ochrany práva autorského, je-li tato osoba povinnému subjektu známa. Pro knihovny poskytující veřejné knihovnické a informační služby podle knihovního zákona^{2a}) a muzea a galerie poskytující standardizované veřejné služby¹⁹) se věta první nepoužije.

§ 16

Odvolání

(1) Proti rozhodnutí povinného subjektu o odmítnutí žádosti lze podat odvolání.

(2) Povinný subjekt předloží odvolání spolu se spisovým materiálem nadřízenému orgánu ve lhůtě 15 dnů ode dne doručení odvolání.

(3) Nadřízený orgán rozhodne o odvolání do 15 dnů ode dne předložení odvolání povinným subjektem. Lhůta pro rozhodnutí o rozkladu je 15 pracovních dnů ode dne doručení rozkladu povinnému subjektu. Lhůtu nelze prodloužit.

(4) Neshledá-li nadřízený orgán důvody pro odmítnutí žádosti, zruší rozhodnutí povinného subjektu o odmítnutí žádosti nebo jeho část a řízení v tomto rozsahu zastaví. Současně rozhodnutím přikáže povinnému subjektu požadovanou informaci žadateli poskytnout ve lhůtě, která nesmí být delší než 15 dnů ode dne oznámení rozhodnutí o odvolání povinnému subjektu. Proti rozhodnutí nadřízeného orgánu podle věty první se nelze odvolat. Poskytnutí informace povinným subjektem lze exekučně vykonat.

(5) Při soudním přezkumu rozhodnutí o odvolání na základě žaloby podle zvláštního právního předpisu soud přezkoumá, zda jsou dány důvody pro odmítnutí žádosti. Nejsou-li žádné důvody pro odmítnutí žádosti, soud zruší rozhodnutí o odvolání a rozhodnutí povinného subjektu o odmítnutí žádosti a povinnému subjektu nařídí požadované informace poskytnout.

§ 16a

Stížnost na postup při vyřizování žádosti o informace

- (1) Stížnost na postup při vyřizování žádosti o informace (dále jen "stížnost") může podat žadatel,
- a) který nesouhlasí s vyřízením žádosti způsobem uvedeným v § 6,
 - b) kterému po uplynutí lhůty podle § 14 odst. 5 písm. d) nebo § 14 odst. 7 nebyla poskytnuta informace nebo předložena konečná licenční nabídka a nebylo vydáno rozhodnutí o odmítnutí žádosti,
 - c) kterému byla informace poskytnuta částečně, aniž bylo o zbytku žádosti vydáno rozhodnutí o odmítnutí, nebo
 - d) který nesouhlasí s výší úhrady sdělené podle § 17 odst. 3 nebo s výší odměny podle § 14a odst. 2, požadovanými v souvislosti s poskytováním informací.
- (2) Stížnost lze podat písemně nebo ústně; je-li stížnost podána ústně a nelze-li ji ihned vyřídit, sepíše o ní povinný subjekt písemný záznam.
- (3) Stížnost se podává u povinného subjektu, a to do 30 dnů ode dne
- a) doručení sdělení podle § 6, § 14 odst. 5 písm. c) nebo § 17 odst. 3,
 - b) uplynutí lhůty pro poskytnutí informace podle § 14 odst. 5 písm. d) nebo § 14 odst. 7.
- (4) O stížnosti rozhoduje nadřízený orgán.
- (5) Povinný subjekt předloží stížnost spolu se spisovým materiálem nadřízenému orgánu do 7 dnů ode dne, kdy mu stížnost došla, pokud v této lhůtě stížnosti sám zcela nevyhoví tím, že poskytne požadovanou informaci nebo konečnou licenční nabídku, nebo vydá rozhodnutí o odmítnutí žádosti.
- (6) Nadřízený orgán při rozhodování o stížnosti podle odstavce 1 písm. a), b) nebo c) přezkoumá postup povinného subjektu a rozhodne tak, že
- a) postup povinného subjektu potvrdí,
 - b) povinnému subjektu přikáže, aby ve stanovené lhůtě, která nesmí být delší než 15 dnů ode dne doručení rozhodnutí nadřízeného orgánu, žádost vyřídil, případně předložil žadateli konečnou licenční nabídku, a neshledá-li důvody pro odmítnutí žádosti v případě, kdy dostupné informace o právním a skutkovém stavu nevyvolávají důvodné pochybnosti, postupuje obdobně podle § 16 odst. 4, nebo
 - c) usnesením věc převezme a informaci poskytne sám nebo vydá rozhodnutí o odmítnutí žádosti; tento postup nelze použít vůči orgánům územních samosprávných celků při výkonu samostatné působnosti.

(7) Nadřízený orgán při rozhodování o stížnosti podle odstavce 1 písm. d) přezkoumá postup povinného subjektu a rozhodne tak, že

a) výši úhrady nebo odměny potvrdí,

b) výši úhrady nebo odměny sníží; v případě, kdy dostupné informace o právním a skutkovém stavu nevyvolávají důvodné pochybnosti, postupuje obdobně podle § 16 odst. 4 věty druhé s tím, že povinnému subjektu přikáže požadovanou informaci žadateli poskytnout ve lhůtě, která nesmí být delší než 15 dnů ode dne zaplacení úhrady či odměny.

(8) Nadřízený orgán o stížnosti rozhodne do 15 dnů ode dne, kdy mu byla předložena.

(9) Rozhodnutí podle odstavců 6 a 7 se oznamuje žadateli a povinnému subjektu. Proti rozhodnutí vydanému podle odstavců 6 a 7 se nelze odvolat. Jde-li však o rozhodnutí podle odstavce 6 písm. c), nelze se odvolat pouze v případě, kdy rozhodl nadřízený orgán určený podle § 178 odst. 2 věty poslední správního řádu nebo podle § 20 odst. 5 tohoto zákona.

(10) Je-li poskytnuta informace podle odstavce 6 písm. c), žadatel může ve smyslu odstavce 1 písm. a) nebo c) postupovat obdobně.

§ 16b

Přezkumné řízení a ochrana proti nečinnosti

(1) Rozhodnutí nadřízeného orgánu lze přezkoumat v přezkumném řízení, pro jehož vedení podle zvláštního právního předpisu je příslušný Úřad pro ochranu osobních údajů.

(2) Dospěje-li správní orgán při přezkumu k závěru, že informace byly odepřeny nezákonně, a dostupné informace o právním a skutkovém stavu nevyvolávají důvodné pochybnosti, postupuje obdobně podle § 16 odst. 4.

(3) K opatřením proti nečinnosti nadřízeného orgánu podle zvláštního právního předpisu je příslušný Úřad pro ochranu osobních údajů.

§ 17

Hrazení nákladů

(1) Povinné subjekty jsou v souvislosti s poskytováním informací oprávněny žádat úhradu ve výši, která nesmí přesáhnout náklady spojené s pořízením kopií, opatřením technických nosičů dat a s odesláním informací žadateli. Povinný subjekt může vyžádat i úhradu za mimořádně rozsáhlé vyhledání informací.

(2) Pokud byla v licenční smlouvě sjednána odměna, nelze požadovat úhradu nákladů.

(3) V případě, že bude povinný subjekt za poskytnutí informace požadovat úhradu, písemně oznámí tuto skutečnost spolu s výší úhrady žadateli před poskytnutím informace. Z oznámení musí být zřejmé, na základě jakých skutečností a jakým způsobem byla výše

úhrady povinným subjektem vyčíslena. Součástí oznámení musí být poučení o možnosti podat proti požadavku úhrady nákladů za poskytnutí informace stížnost podle § 16a odst. 1 písm. d), ze kterého je patrné, v jaké lhůtě lze stížnost podat, od kterého dne se tato lhůta počítá, který nadřízený orgán o ní rozhoduje a u kterého povinného subjektu se podává.

(4) Nesplní-li povinný subjekt vůči žadateli oznamovací povinnost podle odstavce 3, ztrácí nárok na úhradu nákladů.

(5) Poskytnutí informace podle odstavce 3 je podmíněno zaplacením požadované úhrady.

Pokud žadatel do 60 dnů ode dne oznámení výše požadované úhrady úhradu nezaplatí, povinný subjekt žádost odloží. Po dobu vyřizování stížnosti proti výši požadované úhrady lhůta podle věty druhé neběží.

(6) Úhrada je příjmem povinného subjektu.

§ 18

Výroční zpráva

(1) Každý povinný subjekt musí vždy do 1. března zveřejnit výroční zprávu za předcházející kalendářní rok o své činnosti v oblasti poskytování informací podle tohoto zákona obsahující následující údaje:

a) počet podaných žádostí o informace a počet vydaných rozhodnutí o odmítnutí žádosti,

b) počet podaných odvolání proti rozhodnutí,

c) opis podstatných částí každého rozsudku soudu ve věci přezkoumání zákonnosti rozhodnutí povinného subjektu o odmítnutí žádosti o poskytnutí informace a přehled všech výdajů, které povinný subjekt vynaložil v souvislosti se soudními řízeními o právech a povinnostech podle tohoto zákona, a to včetně nákladů na své vlastní zaměstnance a nákladů na právní zastoupení,

d) výčet poskytnutých výhradních licencí, včetně odůvodnění nezbytnosti poskytnutí výhradní licence,

e) počet stížností podaných podle § 16a, důvody jejich podání a stručný popis způsobu jejich vyřízení,

f) další informace vztahující se k uplatňování tohoto zákona.

(2) Pokud má povinný subjekt zvláštním zákonem uloženou povinnost předkládat veřejnou výroční zprávu obsahující informace o jeho činnosti, začleňuje údaje podle odstavce 1 do této výroční zprávy jako její samostatnou část s názvem "Poskytování informací podle zákona č. 106/1999 Sb., o svobodném přístupu k informacím".

§ 19

Umožnění přístupu k informacím nebo poskytnutí informací za podmínek a způsobem stanoveným tímto zákonem není porušení povinnosti zachovávat mlčenlivost uložené zvláštními zákony.15)

ČÁST DRUHÁ PŘECHODNÁ A ZÁVĚREČNÁ USTANOVENÍ

§ 20

- (1) Povinnost uvedená v § 5 odst. 2 nastává dnem 1. ledna 2001. Pro obce, které nejsou městy,16) povinnost podle § 5 odst. 2 nastává dnem 1. ledna 2002.
- (2) Povinnost uvedená v § 5 odst. 3 nastává dnem 1. ledna 2002.
- (3) Lhůta pro poskytnutí informace [§ 14 odst. 3 písm. c)] a prodloužení této lhůty (§ 14 odst. 5) se v prvních 12 měsících od účinnosti zákona prodloužují na dvojnásobek, a dalších 12 měsíců se prodloužují o polovinu.
- (4) Pokud tento zákon nestanoví jinak, použijí se při postupu podle tohoto zákona
- a) pro rozhodnutí o odmítnutí žádosti,
 - b) pro odvolací řízení,
 - c) pro vykonatelnost příkazu poskytnout informace, a
 - d) v řízení o stížnosti pro počítání lhůt, doručování a náklady řízení ustanovení správního řádu¹⁷⁾; dále se při postupu podle tohoto zákona použijí ustanovení správního řádu o základních zásadách činnosti správních orgánů, ustanovení o ochraně před nečinností, v rozsahu § 16b ustanovení o přezkumném řízení a ustanovení § 178; v ostatním se správní řád nepoužije.
- (5) Nelze-li podle § 178 správního řádu nadřízený orgán určit, rozhoduje v odvolacím řízení a v řízení o stížnosti Úřad pro ochranu osobních údajů.
- (6) Informace, které se týkají přenesené působnosti územního samosprávného celku, poskytují orgány územního samosprávného celku v přenesené působnosti.

§ 21

- (1) Vláda vydá nařízení, kterým upraví součinnost orgánů státní správy s obcemi při zajišťování povinností obcí podle § 5 tohoto zákona.
- (2) Vláda stanoví nařízením zásady stanovení úhrad a licenčních odměn za poskytování informací.
- (3) Ministerstvo vnitra stanoví vyhláškou strukturu informací zveřejňovaných o povinném subjektu podle § 5 odst. 1 a 2 způsobem umožňujícím dálkový přístup.

§ 22

Účinnost

Tento zákon nabývá účinnosti dnem 1. ledna 2000.

Přechodná ustanovení zavedena zákonem č. 61/2006 Sb. Čl. II

1. Pro vyřízení žádostí, které povinný subjekt obdržel přede dnem nabytí účinnosti tohoto zákona, se použijí dosavadní právní předpisy.
2. Licenční nebo podlicenční smlouvy, jimiž byla poskytnuta výhradní licence týkající se poskytování informací podle zákona č. 106/1999 Sb., ve znění účinném do dne nabytí účinnosti tohoto zákona, a které byly uzavřeny přede dnem nabytí účinnosti tohoto zákona, pozbývají účinnosti uplynutím sjednané doby, nejdéle však 31. prosince 2008.

Přechodná ustanovení zavedena zákonem č. 222/2015 Sb. Čl. II

1. Žádost o poskytnutí informace podaná přede dnem nabytí účinnosti tohoto zákona se vyřizuje podle zákona č. 106/1999 Sb., ve znění účinném přede dnem nabytí účinnosti tohoto zákona.
2. Povinnost zveřejňovat informace podle zákona č. 106/1999 Sb., ve znění účinném ode dne nabytí účinnosti tohoto zákona, se nevztahuje na informace, k jejichž zveřejnění došlo přede dnem nabytí účinnosti tohoto zákona.
3. Licenční nebo podlicenční smlouva, na základě níž byla poskytnuta výhradní licence týkající se poskytování informací podle zákona č. 106/1999 Sb., ve znění účinném přede dnem nabytí účinnosti tohoto zákona, a která je platná k 17. červenci 2013, pozbývá účinnosti uplynutím sjednané doby, nejpozději však 18. července 2043. Věta první se nepoužije, jde-li o licenční nebo podlicenční smlouvu týkající se digitalizace kulturních zdrojů.

Přechodné ustanovení zavedeno zákonem č. 298/2016 Sb. Čl. XIX

Povinný subjekt zveřejní informace podle § 4b odst. 2 zákona č. 106/1999 Sb., ve znění účinném ode dne nabytí účinnosti tohoto zákona, jako otevřená data a zaeviduje tyto informace v národním katalogu otevřených dat do 12 měsíců ode dne nabytí účinnosti tohoto zákona.

Přechodné ustanovení zavedeno zákonem č. 111/2019 Sb. Čl. XVI

Do 1. ledna 2020 se postupuje podle § 16, 16a, 16b a § 20 tohoto zákona ve znění dosavadních předpisů.

Přechodné ustanovení zavedeno zákonem č. 261/2021 Sb. Čl. LI

Povinné subjekty zveřejní dosud vedené registry, evidence, seznamy nebo rejstříky obsahující informace, které jsou na základě zákona každému přístupné, jako otevřená data podle § 5a odst. 2 zákona č. 106/1999 Sb., ve znění účinném od patnáctého dne ode dne vyhlášení tohoto zákona, do 31. prosince 2023.

Příloha č. 8: Přepis hloubkového rozhovoru s Janem Bočkem (text)

Jaká je délka Vaší novinářské praxe? Studoval jste žurnalistiku, programování či jiný obor na vysoké škole?

Pracoval jsem v IT časopisech Extra PC a Computer. Před tím, než jsme s Petrem Kočím začali dělat datovou žurnalistiku, každý jsme působili v takovém specializovaném médiu. Vystudoval jsem Žurnalistiku a Politologii, předtím Informatiku na Masarykově univerzitě v Brně. Na informatice jsem zjistil, že nechci dělat programátora. Na žurnalistice jsem poté zjistil, že nechci dělat denní zpravodajství. Teď dělám programátora v denním zpravodajství.

Byl u Vás tedy propojen zájem o technologie a žurnalistiku už od začátku?

V rodině mám tři novináře, proto jsem samozřejmě nechtěl jít dělat novináře. Když jsem ale začal studovat informatiku a nebavila mě, tak jsem zkusil přijímací řízení na novinářinu a dostal jsem se.

Kde jste získal určité základy sociologie, které jsou pro Vaši práci specifické?

Základní znalosti jako je analýza dat jsme se s Petrem Kočím začali učit postupně. Na žádné škole jsem nikdy neměl kurz statistiky, nepracoval s excelem, neanalyzoval data. Až v okamžiku, kdy jsme v Hospodářských novinách řekli, že chceme dělat datovou žurnalistiku, jsme se takové věci začali učit. Naučili jsme se programovat v R.

Kde se učíte nové věci?

Samouka. Tehdy jsme se učili přes online kurzy, všechno ostatní funguje tak, že když to potřebuji, tak se to naučím. V poslední době se ale stává málo kdy, že bych potřeboval technologii nebo nástroj, který neznám. Teď se učím nástroje, když chci něco ukázat studentům. Já a Cibulka učíme na vysokých školách. V redakcích ale na to není prostor. Na začátku jsme si mysleli, že budeme v redakcích učit základy práce s daty, de facto to, co teď učíme studenty. V redakci o to ale není zájem.

Jaká nejčastější témata zpracováváte? Odráží se to na výběru datových zdrojů?

Zdravotnictví, školství, lesnictví, energetika. Měkká témata, tvrdý přístup (datový). Genomika, matematická témata, detektivka. Témata, která jsou komplexní. Už nepíšu ani tak long read, kde bych tohle všechno popisoval, ale už se snažím téma pochopit, vyřešit a předat už řešení. Dotace, národní plán obnovy, fond spravedlivé transformace – opakují se stejné vzorce, korupce.

Je pro Vás první je téma, pro které hledáte data nebo naopak máte data, ve kterých hledáte téma?

Funguje to úplně stejně jako jakákoliv jiná komplexní žurnalistika nebo i denní zpravodajství. Mám téma, celý život se učím, co je pointa a důležité v tom tématu. Když se objeví další impuls, výroční zpráva, nebo se stane něco důležitého, tak se od dat odpíchnu a vyprávím. Data jsou v mém případě až druhotné. Primární je story, kdy já si myslím, že je tam nějaký posun, díky kterému má cenu tu story převyprávět znovu a najít komplexnější obraz. Data jsou spíš způsobem vyprávění toho příběhu, slouží k tomu abych příběh vysvětlil ne jako impuls, abych ho vůbec psal.

Jaká je podle Vás komunikace s poskytovateli dat? Je možné navázat vztah s konkrétním úřadem, který Vám data poskytuje?

Občas to tak samozřejmě je. Najdu úředníka, který má data a poskytne mi je. V tématu příčin úmrtí, kdy jsem data požadoval od ÚZIS jsem na takovou ochotnou úřednici narazil. Mám ale pocit, že čím dál více je to tak, že všichni úředníci jsou pod dohledem tiskových mluvčích úřadů a ti mají čím dál větší moc a používají ji k tomu, aby blokovali cokoli, co by někdo mohl chtít. Pro mě je poslední odstrašující případ, kdy chci napsat o reformě psychiatrické péče, napíšu hezký e-mail co potřebuji, proč to potřebuji. Po 14 dnech nemám ani od ÚZIS ani od Ministerstva zdravotnictví žádnou odpověď a neberou mi telefony. Jsem tedy donucen podat žádost o informace 106. Bohužel je to v mnoha případech jediný způsob, jak z úřadů něco dostat. Nyní už to lze říct dost obecně: dříve se s úřady bavit šlo, dostávat z nich data, bylo to méně formální. V poslední době se ale bez 106 člověk nedostane skoro k ničemu.

Sledujete tedy opačný trend? Čím více se o datech mluví a je po nich větší poptávka, tím je jejich získávání složitější?

Já ten pocit mám. Samozřejmě to ale může být i jinými věcmi. U spousty úřadu jsme v „nemilosti“ a to právě proto, že jsme se s nimi v minulosti o data soudili. Tohle by ale nemělo přece vůbec hrát roli, úřady by neměly koukat na vaše články a říkat „aha tenhle člověk píše něco, co nemáme rádi, a proto mu nebudeme vycházet vstříc.“ Mám pocit, že za vlády Andreje Babiše se rozjel tady ten trend. Je to nepříjemné, všechny komunikace s jakýmkoliv úřadem jsou nepříjemné.

Kolik žádostí o informace jste podal a jaké máte zkušenosti?

Poslední, kterou jsem podal, byla 106 k reformě psychiatrické péče. Data jsou nekonfliktní, ale stejně mám takový pocit, že to nedopadne. První 106, kterou jsem podal bylo, když jsme řešili Městskou policii v Brně a odtahy aut. Po roce a půl, kdy jsme ta data dostali jiným způsobem, to nakonec vedlo k tomu, že se naučili data přes 106 dávat. Pak jsem o data žádal ještě 2x. Zajímavá zkušenost byla, když jsem podával 106 na to, kde jsou v Brně přechody, a to BKOM – Brněnské komunikace a Magistrát města Brna. Dostal jsem legrační, zamítavou, odpověď, že mi data nemohou dát, protože to jeden z těch dvou úřadů jako podřízená akciovka zakázal. Nedal jsem jich moc, ale mám pocit, že pokaždé, když nějakou podám, tak vyjma velkých ministerstev, kde s tím mají zkušenost, tak je to pro ně nová pohádka, drama, někoho naštvu. Úředníci, kteří to například neznají, si to mohou vzít osobně. Jako novinář data potřebuji hned. Obzvláště v denním zpravodajství je obtížné vysvětlovat, proč jedno téma může trvat takovou dobu. Když ale podám žádost o informace, odpověď dostanu za náhodnou dobu. Je to nepraktické a bylo by krásné, kdyby všechna data, která potřebujeme, byla jako otevřená, někdo se o ně staral, anebo s námi alespoň komunikoval. My bychom 106 nemuseli podávat, nebo by to k něčemu bylo. To se teď ale neděje.

Máte zkušenosti s Národním katalogem otevřených dat a Open data wish-listem Ministerstvem vnitra?

Vím, že to existuje. V minulosti jsem nějaké podněty dával Michalu Kubáňovi, který působil jako koordinátor otevřených dat na Ministerstvu vnitra. Za mě jsou otevřená data problematická především v tom, že je tam strašně málo věcí, které by byly k něčemu užitečné (v Národním katalogu otevřených dat). Vyžaduje to totiž značné úsilí, mluvit s menšími úřady, organizacemi, vysvětlovat proč by měli data otvírat. Aby se tam dostaly užitečná data, vyžadovalo by to značné úsilí. Jsou ale jednotlivci, v Brně, kteří si vezmou

nějaký data set, vyčistí ho a najdou v něm nějaké příběhy. Je potřeba lidí, kteří jsou aktivní, možná až aktivističtí, kteří se o jednotlivé data sety starají. Vysočina byla jedním z prvních krajů, který zveřejnil otevřená data. Tehdy jsem se na ně podíval a byly to seznamy názvů obcí. Jako novinář vlastně nevím, jak bych z něčeho takového měl poskládat článek. V Národním katalogu otevřených dat je tedy mnoho data setů, které může státní správa vykázat, ale nebývají to data, která jsou komukoliv k čemukoliv.

Jak hodnotíte kvalitu a dostupnost otevřených dat?

Dostupnost i kvalita dobrá, témata nic moc. Možná, že když se do toho člověk zahrabe, tak tam něco najde. Ale pro mě v tom postupu, kdy já mám téma a chci ho odvyprávět by bylo zdržením kontrolovat na NKOD, jestli tam náhodou nejsou data, která se mi hodí. Zeptám se přímo u zdroje, kde mi třeba data nedají, což pravděpodobně znamená, že je nedali ani do toho NKOD.

Ověřujete nebo zpochybňujete kvalitu dat?

Možná je to podmět pro mě začít to dělat jinak. Teď k tomu ale přistupuji tak, že pokud jsou data použitelná, považuji je za pravdivá. Klasický problém je totiž v tom, že jsou data nepoužitelná. Nejsou dostatečně vyčištěná. Jakmile vidím, že nejsou v dostatečné kvalitě, беру je s rezervou. Typicky, pokud dostanu stejný data set na stejné téma od dvou zdrojů a říkají jiné věci. Není ani výjimkou, že od stejné instituce jsou v ročenkách každý rok jiné údaje za uplynulá období. Je tedy na místě se ptát, co se stalo, že jsou ta čísla za stejný rok pokaždé jiná. V okamžiku, kdy data nesedí, to musím začít zkoumat. Ale v okamžiku, kdy nemám takový impuls, mám čistá data, dostanu je autoritativně od ministerstva tak s nimi pracuji, jakože jsou pravdivá. Nemyslím si, že by v Česku někomu stálo za to si data vymýšlet. Problém je opravdu spíše v tom, že je v datech nepořádek, a ne v tom, že by lidé nechtěli říct pravdu. Navíc co můžeme my dělat s tím, pokud Ministerstvo zdravotnictví například poskytne data, která si vymyslí.

Reportujete v datech chyby?

Toho se týká poslední žádost o informace 106. Vysvětlil jsem, co potřebuji a co nesedí. Reportuji, protože se zároveň potřebuji zeptat, jak to tedy je. Jsou ale i případy, kdy my jsme první, kdo data analyzuje, a proto je to legrační, že sami úředníci nemusí vědět, co v datech je.

Jak podle Vás státní správa výběrem poskytovaných dat formuje datovou žurnalistiku?

Představil jsem si, jak by vypadal „ideální stav“. Kdybychom všechna data měli v NKOD, kam by člověk sáhl a měl data k čemukoliv. Což samozřejmě ani teoreticky nejde, protože my vždy potřebujeme specifický úhel pohledu. Asi bychom měli více datových novinářů jako jsem já, kteří příběhy doplňují daty a méně jako Jan Cibulka, který o data často bojuje. Datová žurnalistika by byla více o vyprávění příběhů a méně o hrdinech, kteří dokážou data vydolovat z úřadů. Případá mi, že k naší východoevropské novinářině tradice souboje se státem je výraznější a více opěvovaná, a to hrdinství novinářiny se soustředí na to, že jste jako novinář schopný přeprat stát, takže nějaké mé vyprávění dlouhých příběhů o tom, co věci znamenají, není moc úspěšné. Dobrý novinář je v našem kontextu mnohdy někdo, kdo přepere státní správu. Zároveň by to byla škoda, protože mi připadá, že je napínavé, když data dostupná nejsou. Další datoví novináři, kteří se objevují, tak už budou více jako Cibulka a budou o dostupnost dat bojovat, protože je to potřeba.

Co považujete za hlavní limity Vaší práce?

Největší limit své práce vnímám v tom, že je komplikované ji pochopit a dostatečně ocenit. Čtenáři mou práci občas ocení, může mi například napsat i poslanec, že až díky mému textu získal kontext k nějakému tématu. Ale právě proto, že málo kdo pozná rozdíl mezi kvalitní a nekvalitní datovou žurnalistikou může tím trpět má vnitřní motivace se posunout někam dál.

Vysvětlujete čtenářům metodiku dat?

Otázka je, jestli můžeme říkat tomu, odkud máme data a co znamenají, metodika. Je to spíš popisek. Metodiku si představuji například tak, že pokud mám data z anonymního zdroje, tak podstatnou část textu věnuji tomu, že čtenářům vysvětlím, proč právě tomuto zdroji důvěřuji, jaká je metodika celé práce s daty, co to znamená. Možná je to část, kterou většina lidí neocení, ale dodává důležitý kontext. Novinář by měl být schopen popsat, i když se to od něj nečeká, svou motivaci, proč daný článek píše. To ale nepatří přímo do zpravodajství, jehož jsme součástí.

Příloha č. 9: Přepis hloubkového rozhovoru s Janem Cibulkou (text)

Jaká je délka Vaší novinářské praxe? Studoval jste žurnalistiku, programování či jiný obor na vysoké škole?

Programování jsem se začal učit na střední škole, kde jsme se učili spíše webový vývoj. K novinářské práci jsem se poprvé dostal také již na gymnáziu, a to v radničních novinách. Potom jsem byl na stáži v deníku Blesk, střídal jsem to s prací v IT firmě. Šel jsem studovat Mediální studia, nejspíše právě proto, že to byl obor, který mě živil. Rychle jsem zjistil, že to není úplně pro mě a školu jsem nedokončil. Míchal jsem to od začátku. Pak jsem nastoupil do Hospodářských novin, a to projektem Data Retention, kdy jsem se „soudil“ s T-mobile, který mi nakonec poskytl moje lokalizační údaje.

Kde se učíte nové věci?

Většinu nových technologií se učím sám, na internetu. Je dobré sledovat, jak se vyvíjí nástroje, které se osvědčily. Kolega Marcel Šulek měl zásadu, že se učil jednu novou technologii za rok. Nemám to takto striktně, ale je pravda, že minimálně jednu věc by měl člověk pochopit a třeba ji posléze i opustit, pokud zjistí, že mu nevyhovuje. V rámci datové redakce jsme vždy dělali nárazové přednášky pro studenty, nejen žurnalistiky, ale také na ekonomických školách a podobně. Honza (*Jan Boček*) učí v Brně a já učím již dva roky na Univerzitě Karlově, původně na bakalářské stupni, nyní je to na magisterském stupni. V budoucnu plánuji učit analýzu a zpracování dat, což bude vyloženě technické.

Učíte základy datové žurnalistiky také v redakci?

Základy kyber bezpečnosti pro novináře učím také v redakci i mimo ni. Kyber bezpečnosti ani datová žurnalistika ale nejsou věci, které umožňují práci dělat rychleji, ale precizněji. To ale není vždy požadavek, požadavek většinou bývá více výstupů a rychleji. Na druhou stranu není to úplně ztraceno, protože vidíme, že se datová žurnalistika dostává do více redakcí. Je tedy snaha tyto složitější přístupy používat.

Jaká nejčastější témata zpracováváte? Odráží se to na výběru datových zdrojů?

Pokud jsou to témata datové žurnalistiky, protože v poslední době zjišťuji, že nedělám jen datovou žurnalistiku, ale technologická znalost mě natlačila dělat technologickou žurnalistiku. V Česku se technologická žurnalistika redukovala na recenzi produktů, což je

legitimní, ale nevysvětlí to jak technologie působí na společnost. V datové žurnalistice se v poslední době nejvíce věnuji koronaviru. Pak jsou to témata s přesahem do demografie, kriminality, životního prostředí. Majoritně využívám data veřejné správy.

Kolik žádostí o informace jste podal a jaké máte zkušenosti?

Zákon je fajn, kde se berou jeho slabosti: byl napsán tak, aby se lidem dobře používal. Nikdo nepředpokládal, že veřejná správa mu bude intenzivně klást odpor. Vymahatelnost toho zákona, pokud jde o něco, co je sporné, je zdoluhavá a náročná. Vyžaduje to značnou časovou investici, v rámci redakce tedy i investici finanční. Výsledek se může táhnout i měsíce. O lepší vymahatelnost se v minulém volebním období pokusil Michálek, když prosadil přezkum z pozice Úřadu pro kontrolu osobních údajů, ale zákonodárci efektivně zabránili tomu, aby se to dělo tím, že úřad nemá dostatek prostředků na takovou kontrolu. Vidíme opravdu velký odpor státní správy k tomuto zákonu, který nemizí. Představa byla, že časem bude odpor klesat, ale to se ani náznakem nestalo.

Komunikace s poskytovateli dat je tedy podle Vás horší než v minulosti?

Problém administrativy posledních osmi let byl v tom, že se obsadili i nižší rozhodovací úrovně. Zatím co předtím se vyměnila vláda, ministři, náměstci a občas šéf odborů, tak teď došlo k obměně míst i na nižších stupních. Úředníci dostali jistou míru nezávislosti, ale zároveň se „bojí“ svého politického pána a rozhodují ve prospěch politických linií. Zejména věci, jako je poskytování informací novinářům, se rozhoduje striktně politicky, žádná nezávislost úřadů neexistuje. Viděli jsme to teď během koronaviru na Ministerstvu zdravotnictví, kdy to šlo v politické linii, rozhodovali lidé, kteří by neměli mít rozhodovací pravomoc. To je velký problém, a to nejen informačního zákona, ale obecně zpolitizovanost státní správy a poplatnost vládnoucí garnituře je extrémní problém. Typický případ ÚZIS, šéf ústavu Ladislav Dušek poskytuje informace lidem, které má rád nebo je s nimi kamarád. S lidmi, se kterými není za dobře na datech dusí, netýká se to jen novinářů ale i výzkumníků. Tohle se během koronaviru obrovským veřejným tlakem podařilo lehce oslabit, zůstává otázkou, co se bude dít, až se vymění Sněmovna. Naše veřejná správa je institucionálně slabá. To znamená, že je velmi jednoduché ji z politické pozice „ohnout“ k obrazu svému. To vede k tomu, že veřejná moc se vykonává k zájmu jednotlivců v čelech úřadů. O tom, jestli data přes informační zákon 106 dostanu, rozhoduje spousta věcí. Stačí, když je ze strany úřadu obava, že by se to nějakým způsobem mohlo dotknout politických zájmů úřadu.

Zákon přesně stanovuje, z jakých důvodů se data neposkytují. Úřady mnohdy hledají právní důvody, za které své pravé zájmy maskují.

Máte zkušenosti s Národním katalogem otevřených dat a Open data wish-listem Ministerstvem vnitra?

Je to skvělá věc, protože je to způsob veřejné konzultace. Ministerstvo vnitra a lidé, kteří mají konkrétní data sety na starosti, pořádají na toto téma diskuse. Je potřeba, aby veřejná správa takovou komunikaci prováděla. V ideálním světě jsou data zakotvena v Zákoně o svobodném přístupu k informacím. To znamená, že ve chvíli, kdy se na to sahá, tak by se mělo jít po linii: ministerstvo, které má 106 v gesci, tedy Ministerstvo vnitra, osloví podatele žádosti, jimiž jsou v České republice nejčastěji média. Navazují se ale častěji neformální kontakty mezi novináři a úřadem. Česká republika má open data komunitu, která se schází na různých akcích, na které chodí také zástupci různých resortů, kteří si jdou poslechnout, co by ta komunita chtěla. Dalším způsobem je právě ten wish list. Ten má výhodu v tom, že Ministerstvo vnitra systematicky seznam prochází a pokud něco zveřejnit nejde, vysvětlí žadateli z jakých důvodů to nejde. Přičemž se podařilo prosadit změny i v tom zákoně, teď například všechny rejstříky, které jsou veřejné, jsou i jako open data.

Takže nejlepší cesta je oslovit přímo toho člověka, který daný data set spravuje?

Ano, pokud žádám o konkrétní data set je nejlepší jít za původcem data setu, například i zjistím, že se nedá použít kvůli technickému stavu atp. Často důvody, proč úřad klade odpor ke zveřejnění nejsou ani racionální. Je to obava z toho, že z toho budou problémy, tedy si řeknou „pojd'me radši nedělat nic.“

Jak hodnotíte kvalitu a dostupnost otevřených dat?

Kvalita dat je téma, které se začíná řešit, protože je mizerná. Je to způsobeno hlavně tím, že státní správa nevyužívá data, která má k dispozici. Jejich cílem je říct, že ne všechna data mohou být open data. Nicméně ty, které open být mohou tak je zveřejníme. A veřejná správa je ze zákona povinná využívat ve svých systémech open data. Teď to funguje tak, že si státní správa data předává nějak a způsobem B je tlačí veřejnosti. Veřejné správě je jedno, jak se budou data veřejnosti využívat, protože ona sama není tím příjemcem. Kvalita se řeší na formální úrovni, řeší se nastavení frame worků, jak by teda kvalitní data měla vypadat. Aby byla data podle zákona otevřená, musí splňovat určité parametry. Ne všechno lze ale

legislativně popsat. Ta kvalita, protože to často mohou být velmi komplexní otázky, u jednoho konkrétního data setu je to poplatné tomu, jak daný data set vzniká.

Jak hodnotíte kvalita u sad, které používáte?

Pandemická data se nabízí na prvním místě. Oproti začátku pandemie jsou nyní data již dobře dostupná. Jejich kvalita se po technologické stránce také zvyšuje. Co trochu zaostává jsou meta data, tedy konkrétní popis toho, jak daný data set vznikl a co přesně znamenají jeho hodnoty. Je ale snaha věci vysvětlovat. Hodně používám také geoprostorová data, tam je poskytovatelem primárně Český úřad zeměměřický a katastrální, u věcí, které jsou dostupné, je technická kvalita velmi vysoká, protože je daná oborovým standardem.

Je pro Vás první je téma, pro které hledáte data nebo naopak máte data, ve kterých hledáte téma?

Na začátku mívám primárně hypotézu, kterou se snažím otestovat a k tomu používám data. Ano v minulosti, pokud nějaký úřad otevřel data, která do té doby nebyla veřejná, tak jsme hledali způsob, jak je využít. Teď už se to neděje tak často.

Ověřujete nebo zpochybňujete kvalitu dat?

V tomto případě byla velká škola pandemie. Data se zveřejňovala rychle, a i pro ÚZIS to byla nová situace, pro kterou neměli personál a datové sady obsahovaly chyby. Nezbyvalo než data validovat vůči „základním pravidlům“ – v obci nemůže být více nakažených lidí, než kolik v ni žije. Bylo nutné komunikovat s poskytovatelem a pořizovatelem dat. Data je nezbytné kontrolovat. Samozřejmě, že když na první, druhý, třetí pohled vypadají, že jsou v pořádku tak může člověk přehlédnout jednotlivost, která ve velkém data setu může být, ale chce to alespoň zkontrolovat, zda sedí rozptýlení hodnot, vyjet si histogram.

Reportuje chyby? Pokud ano, jak hodnotíte komunikaci s poskytovateli dat?

Pokud je to chyba technického charakteru, tak mám zkušenosti dobré. Obecně ÚZIS doteď nejdřív zveřejní datový set a potom zkoumá, kdo ho používá. Místo toho, aby se ptali těch lidí, co přesně potřebují. ÚZIS je ale politický přetlačen. Stalo se to opakovaně, kdy jsme viděli, že se zveřejňování dat dělo na politické úrovni. Z pohledu toho, jak funguje státní správa, do zveřejňování dat o koronaviru vstupovali lidé, kteří neměli mít rozhodovací pravomoci.

Jak podle Vás státní správa výběrem poskytovaných dat formuje datovou žurnalistiku?

Je to ten stejný problém jako vidíme v jiných oborech, a to, že Česká republika spoustu informací nesbírá, nesleduje, nebo je sleduje špatně. Zatímco například ekonomové z těchto důvodů nemohou navrhovat politiky nebo zkoumat daný problém, tak my o něm nemůžeme psát. Nemáme k dispozici dostatečně robustní datovou základnu, abychom mohli například říct, jestli se u nás zneužívají nebo nezneužívají sociální dávky a v jaké míře a podobně. Není to primárně problém datových novinářů, veřejná správa nemá tuto datovou základnu a není proto schopna dělat analytickou činnost. My o tom nejsme schopni psát.

Co považujete za hlavní limity Vaší práce?

Nedostupnost dat to zřejmě není, protože pokud je to opravdu kriticky důležitý data set, tak kromě toho, že se o něj soudíme, obvykle najdeme cestu, jak se k němu dostat. Minimálně se podaří získat nějaký menší vzorek dat, díky kterému se nám podaří danou hypotézu validovat. Je to hodně spojené s poslední vládou, poslední garnitura je ale posledních osm let. Primárně jsem zažil tohle a ten problém byl možná i předtím, ale to jsem na něj nenarážel. Obecně problém je, že veřejná správa v Česku nerada komunikuje, nebo nekomunikuje otevřeně. Nerada vysvětluje svoje kroky, protože právě s vědomím, že často kroky vznikají na politické lince, a ne analytické, tak často vysvětlení je „pan ministr to tak chtěl“. Často to zmiňuje také Nejvyšší kontrolní úřad: státní správa kašle na analytickou činnost, která by měla být podkladem jejího rozhodování. My na to narážíme jako novináři tak, že víme, že je někde nějaký problém a snažíme se ho popsat a veřejná správa nám není možná k tomu poskytnout zpětnou vazbu a říct „vy to chápete špatně“ nebo „ten důvod je širší“. Novinářům by pomohlo, kdyby mohli za celou novinářskou obec formulovat své požadavky a předpoklady. Žádná zastřešující organizace ale nefunguje, čehož státní správa využívá. Ta nejednotnost funguje v tom, že když nějaká redakce na něčem pracuje tak se za ni ostatní kolegové nepostaví a veřejná správa může říkat to je problém té a té redakce, to není problém veřejnosti či veřejného zájmu. Česká datová žurnalistika pracuje s malými rozpočty, a proto se jeden autor musí věnovat více tématům a čtenářům opakovaně vysvětlovat své postupy, což dělá přechod ke komplexnějším věcem složitý.

Vysvětlujete čtenářům metodiku dat?

Mám zásadu, říká se tomu intelektuální poctivost – neměl bych zneužívat děr v datech a prezentovat problém tak, aby vyšel podle toho, jak chci já a vědomě zamlčoval čtenáři skutečnosti, které omezují důvěryhodnost těch daných dat. Abych já data použil, musím jim dostatečně věřit k tomu danému účelu, ke kterému je použiji. Poté necítím potřebu jít do detailů a čtenáři vysvětlovat kde se data vzala, jak jsem je analyzoval – pokud to není podstatné pro jejich užití. Pokud to podstatné je a v těch datech něco chybí, tak to samozřejmě tomu čtenáři musím říct, že je tam nedostatek. Upřesním to na příkladu pandemických dat. Můžeme říct, že data mají zpoždění, a tedy neukazují výsledek úplně přesně k danému dni, ale nic to nemění na tom, že ukazují určitý trend, a proto je můžeme použít.

Příloha č. 10: Přepis hloubkového rozhovoru s Petrem Kočím (text)

Jaká je délka Vaší novinářské praxe? Studoval jste žurnalistiku, programování či jiný obor na vysoké škole?

K prvnímu počítači jsem se dostal v 80. letech, a to díky matce, která se k embargovanému zboží dostala v práci. Byla tak hodná, že mi počítač na víkendy nosila domů a já na tom hrál hry. Bavilo mě to, zkoušel jsem i základy programování. V 90. letech už jsem měl doma počítač na stálo, ale hrál jsem spíš hry. Na gymnáziu jsem měl humanitní zaměření, nabízelo se studium žurnalistiky. Bylo to spojení toho, že mě vždycky bavilo psát a dozvídat se nové věci. Vystudoval jsem žurnalistiku, měl jsem práci v novinách. Asi po deseti nebo dvanácti letech v novinách jsem začal mít pocit, že aby ta práce měla smysl, měl bych se zaměřit na téma, kterému rozumím více než ostatní. Když jsem přemýšlel, čemu rozumím více než většina kolegů, tak jsem měl pocit, že jsou to technologie, počítače, internet. Začal jsem o tom psát sloupky do Respektu. Z předchozího působení jsem měl návyk, že se musím potkávat s lidmi z oboru, takže jsem se zúčastňoval různých IT akcí, festivalů, konferencí. Nejdříve jsem psal o lidech, které jsem potkal a měl jsem pocit, že mé téma je, jak technologie mění společnost. Postupem času mě ale čím dál více lákalo si webové technologie vyzkoušet.

Jak se učíte nové technologie?

Na kurzy programování jsem nikdy nechodil. Co pro mě byl výrazný krok, byly začátky datové žurnalistiky, kdy jsme využívali nástroj Google Sheet. Velmi rychle jsme narazili na různá omezení a já se rozhodl naučit programovací jazyk R, který byl velmi populární.

Základy jsem získal v online kurzu, spolu se základy statistiky. Nadchlo mě to a u nástroje už jsem zůstal.

Učíte se nové technologie i nyní, nebo zůstáváte u těch, které již znáte?

Pořád. Vyplývá to z různých situací, kdy potřebuji použít něco, co jsem ještě nedělal. Z velké části samozřejmě při programování různých aplikací používám dostupný kód jiných uživatelů, ale zároveň se snažím pochopit, co daný kód dělá. Člověk se pořád učí, já s tím začal poměrně pozdě, bylo mi 35. Postupem času člověk narazí na technologie, které mu vyhovují a vrací se k nim. Datové sady nejsou vždy úplně stejné. Každá instituce data zveřejňuje v trochu jiném formátu a je potřeba mít flexibilitu a volnost si je přeskupit do formátu, který se ti hodí, protože ho například vyžaduje ten vizualizační nástroj, který chci použít. Principiálně je proces zpracování dat od veřejných institucí stejný, liší se ale právě v různých drobnostech typických pro dané sady.

Jaká témata nejčastěji zpracováváte?

V poslední době jsem silnou tematickou specializaci neměl. Opakují se témata volební, koronavirus, ekonomické věci.

Které datové zdroje pro svou práci nejčastěji používáte?

Za poslední rok dva určitě data Českého statistického úřadu a otevřená data Ministerstva zdravotnictví.

Kolik žádostí o informace jste podal a jaké máte zkušenosti?

Žadostí o informace jsem podal zhruba pět. Výsledky jsou smíšené, od toho, že přišlo, co jsem požadoval, až po neposkytnutí informací. Většinou se spíše snažím vycházet z toho, co je k dispozici. Myslím, že to souvisí také s tím, jak jsem tematicky méně vyhraněný.

Je pro Vás první je téma, pro které hledáte data nebo naopak máte data, ve kterých hledáte téma?

Řekl bych, že v mém případě jsou první data, ve kterých hledám příběh. Nedokážu říct, jestli to je správně nebo ne. Vychází to z dob, kdy dat bylo málo a abychom ukázali, co je to datová žurnalistika, tak jsme museli začít u toho, co bylo a s tím zkusit něco vymyslet. Teď už je nejspíše častější mít téma a shánět k němu data.

Jak hodnotíte kvalitu a dostupnost poskytovaných dat?

Těžce se to zobecňuje. Některé úřady si na zveřejňování dat staví PR a kdykoliv narazí na nějaký data set tak ho zveřejní. Jiné úřady to nevnímají jako prioritu a dělají jiné věci a když po nich někdo chce data, tak to vnímají spíše jako komplikaci. Hlavní zdroje, o kterých jsem mluvil, mají dostupnost i kvalitu dat uspokojivou. S výjimkou Českého statistického úřadu, který má veřejnou databázi nepřívětivou. Je dost náročné dokliktat se k tomu, co potřebuji, popřípadě to tam není. Na druhou stranu mají celé informační oddělení, které má v popisu práce lidem radit a pomáhat. Co se týče kvality, každá data mají chyby a je důležité si toho při jejich zpracování být vědom a vzít to v potaz při vyvozování závěrů.

Ověřujete či zpochybňujete kvalitu poskytovaných dat?

Zkouším si udělat explorační analýzu, z dat si vykreslím rychlé grafy a když vidím nějaké nepravděpodobné výkyvy nebo cokoliv, pro co nemám vysvětlení, tak často dojdou k tomu, že je to například metodou sběru dat nebo se může jednat o technickou chybu.

Reportujete chyby v datech?

Napsal bych e-mail, ale dál již nesleduji vývoj a nečekám, až to opraví, pokud to nemá vliv na výsledek. Rychle se podívám na web, napíšu s dotazem.

Jak podle Vás státní správa výběrem poskytovaných dat formuje datovou žurnalistiku?

Hodně, protože novináři jsou pod tlakem něco vydávat. Nejjednodušším způsobem jak „něco vydat“ je podívat se na data státní správy a vymyslet z nich téma. Je to podstatně jednodušší než dlouhé měsíce a roky páčit ze státní správy něco, co z ní možná ani nevypadne. Do značné míry úřady zveřejňováním mohou ovlivňovat, o čem se bude mluvit. Z toho také vyplývá, že obecně dat zveřejňovaných státní správou přibývá, ale dat, které by těch úřadům mohly být nějak nepříjemné přibývá méně pokud jich dokonce neubývá. Vzniká taková datová mlha, publikují velké množství nepříliš užitečných data setů, aby na oko uspokojili poptávku, ale skutečně podstatné věci nebo věci, ze kterých by mohly vyplynout nějaké problémy si nechávají pro sebe.

Limituje Vaši práci nedostatek otevřených dat?

Musím říct, že dříve bych nejspíše řekl, že ano. V poslední době mám ale spoustu témat a mám co dělat, abych se probral tím, co k dispozici je. Kdybych teď měl trochu volnější ruce nebo větší odhodlání tak bych odpověděl ano.

Jaké jsou tedy hlavní limity Vaší práce?

Shrnu bych to do hesla: informace jsou levné, porozumění je drahé. Dat, které můžeme zpracovávat je moc, ale říct, co to znamená a případně z toho vyvodit nějaké doporučení pro občany je těžké bez hlubší znalosti podstaty věci, ke které nám samotná data nestačí. Dříve jsem jako hlavní limit vnímal nedostatek technických dovedností, ale teď bych řekl, že interpretace je problém, co je ta zpráva. Tím se avantgarda datové žurnalistiky vrací k tradiční žurnalistice, že novináři měli rozdělené beaty, kterým se věnovali denně, a tudíž tomu rozuměli do hloubky a oklikou se tomu vracím, že mám pocit, že je nutné něčemu rozumět více.

Vysvětlujete čtenářům metodiku dat?

Vždycky je potřeba data vysvětlit. U každého grafu má být minimálně zdroj dat s aktivním proklikem. U věcí, u kterých došlo k nějakému většímu počítání na naší straně tak to sebou nese nějaké rozhodnutí, jako které proměnné zařadím a které ne, jaké použiji měřítko, apod. Určitě pokud mám pocit, že kdybych se u některých z těchto věcí rozhodl jinak a výrazně by to změnilo vyznění nebo výsledek, tak se to snažím v tom textu výslovně říct. Problém je, že pokud do textu přidáme metodologickou poznámku je to ten moment, kdy většina lidí přestane číst, protože je to obtížně stravitelný text. Je to náročné, teď nad tím přemýšlím při zpracování dat o očkování proti koronaviru. Jak to říct, co vlastně zobrazujeme, jaký je rozdíl mezi maximální nebo nevyužitou kapacitou. Říct to co nejpřesněji a zároveň krátce mnohdy nedává smysl. Je to skoro největší umění. Zjistit, pochopit, převyprávět. Tak bych kladl důraz na ten přechod od pochopit k převyprávět.

Příloha č. 11: Přepis hloubkového rozhovoru s Alešem Vojířem (text)

Jaká je délka Vaší novinářské praxe? Studoval jste žurnalistiku, programování či jiný obor na vysoké škole?

Mám to dost propojené a před pěti lety se mi to v datové novinářině vše projmulo. Byla to náhoda, protože na konci 90. let jsem studoval Vysokou školu ekonomickou, fakultu Informatiky a statistiky a jako hlavní specializaci jsem měl znalostní inženýrství, ale jako

vedlejší specializaci jsem měl ekonomickou žurnalistiku. Díky ekonomické žurnalistice jsem v praktické části začal spolupracovat s Hospodářskými novinami, které potřebovaly nové elévy. Už při škole jsem v Hospodářských novinách začal pracovat, v Economii jako vydavatelství jsem zůstal přes deset let, později na pozici zástupce šéfredaktora v Ekonomu. Přes Týden jsem se dostal do Deníku, kde jsem původně měl vykonávat jinou práci. Když jsem se po zkušební době bavil s vedením, z jejich strany byl zájem založit datové oddělení. V ten moment mi došlo, že už nechci být manažer nebo editor, ale chtěl jsem dělat více vlastních věcí. Navíc se mi sešly kompetence, které tam byly už od vysoké školy. Sešla se mi jak praxe novináře, manažera v médiích a technické znalosti, které mě díky škole vždy bavily.

Jaká je tedy Vaše role v rámci redakce? Kolik lidí se datové žurnalistice věnuje?

Začalo to tak, že jsem měl založit redakci datové žurnalistiky, ale v podstatě jsem v ní byl sám. Vtipkoval jsem, že se má redakce skládá z Vojíře, Vojíře a Vojíře. Tedy nikdo neodmlouvá, ale co si neudělám, to nemám. Změnilo se to zhruba před dvěma lety, kdy se moje výstupy líbily tehdejšímu šéfredaktorovi Deníku, který ale odcházel na Slovensko do týdeníku Trend, což je něco mezi Ekonomem a Týdnem a chtěl, abychom dělali stejné věci a na to už by má kapacita nestačila. Na online zpravodajství jsme měli šikovného kolegu, kterého jsem si vzal do týmu a jsme dva. Pro finální výstupy převážně do tištěné verze využíváme grafika, který funguje pro celou redakci. Mojí výhodou je, že obsah dodávám jako službu. Tudíž i když je redakce velká a tištěné vydání má oddělenou redakci, tak já své výstupy publikuji na všech těchto platformách, včetně sociálních sítích. Funguje to tak, že výstup zpracuji, připravím text a vizualizace. Jako člověk ekonomický říkám „když už má firma náklady na výrobu jen jedny, tak je hloupost to neprodat různými kanály.“ Takže vždy to vychází v printu, zároveň na webu, zřídka se stává, že to vyjde jen na jednom místě. Z pravidla, protože se věnujeme zpravodajství s přesahem, že prvně výstupy vychází na webu Deníku a poté v tištěné verzi. Je výjimka, a to jsou mapy ke koronaviru, které aktualizují každý den, ale ne vždy vycházejí v tištěné verzi. Máme také třetí způsob, jak výstupy publikovat, a to v Deníku Extra, což je převážně inzertní nosič rozdáváný zhruba do milionu schránek. Některé projekty, které děláme, se dostanou i do dalších časopisů vydavatelství jako například Květy.

Jak se učíte nové technologie?

Historicky jsem začínal na Infogramu, který je jednoduchý a doporučuji ho začátečníkům. Mám to tak, že když mi nástroj přestane stačit, tak se koukám po něčem lepším. Teď pracuji s nástrojem Tableau, který mi dlouhodobě hodně vyhovuje, i když samozřejmě dostat se do něj trvá delší dobu. Máme na něj připravené také naše systémy, takže to funguje poměrně jednoduše. Začínám ale narážet na různé limity, a to především v tom finálním kroku, kdy se snažím, aby vše bylo precizní. Na doporučení Jana Cibulky jsem začal s Java Scriptem, který by tyto limity mohl vyřešit. Koukám, jak to dělají ostatní a snažím se z toho pochopit více. Vidina je, že bych se od Tableau v budoucnu odprostil.

Jaká témata nejčastěji zpracováváte?

Tím, že jsem na to sám, musím přeskakovat z tématu na téma. Například teď začínám se sportovní datovou žurnalistikou, což jsem nikdy nedělal. Mohu říct, že zpracovávám pouze témata, která jsou přínosná pro naše čtenáře. Když si představím naše čtenáře, jsou to lidé v regionech, trochu starší nebo středního věku. Zejména jsou to populační data nebo socio ekonomická témata. Co vždy funguje jsou důchody, teď například inflace a jak zdražují jednotlivé položky ve spotřebním koši. Snažili jsme se zpracovávat zdravotní data i před pandemií koronaviru. Výstupy na tato témata jsou ale vždy v kratším rozsahu, a ne tak podrobné, protože mám přesně dán rozsah, do kterého se musí vejít. Dále se věnujeme sociálním tématům, exekuce, nehody. U populačních dat Českého statistického úřadu poté sňatky, rozvody. Jsou to témata, která se vždy dají najít v regionálním členění alespoň na úroveň krajů, pro nás optimálně na úroveň okresů, protože máme přes 70 mutací. Proto když říkám, že máme datovou redakci jen jednu, a to jedno nebo dvou člennou, tak navíc kolegům do regionů posílám předem zpracované podklady maximálně jednoduše, a oni se mohou rozhodnout, zda se tématu budou věnovat více. Témata tedy vybíráme s ohledem na to, zda to zajímá naše čtenáře a data se dají zpracovávat dále na nižším detailu na kraje a okresy.

Které datové zdroje nejčastěji využíváte?

Nejčastěji pracuji s daty Českého statistického úřadu, Ústavu zdravotnických informací a statistiky, policejními daty, daty České správy sociálního zabezpečení.

Máte zkušenosti s podáním žádosti o informace?

O data, která nebyla dostupná jsem žádal, ale nikdy ne přes zákon 106. Zeptal jsem se, zda data jsou a buď jsem je dostal nebo nedostal. To, že osobně nepoužívám Zákon o svobodném

přístupu k informacím neznamena, že ho nemám rád, spíše naopak. Jsem ale takový „černý pasažér“ všech snah kolegů, například zmíněného Jana Cibulky, protože dělá skvělou práci, na které vyděláváme my všichni. Nejen novináři, ale i různé neziskové organizace. Jsem rád, že zkušenější kolegové zákon využívají, protože se ukazuje, že se data za poslední dobu hodně otvírají, samozřejmě záleží na tom, jak která. Lepší se ale forma, v jaké úřady data poskytují a myslím si, že je to právě zásluhou konkrétních nadšenců. Například zkušenost se statistickým úřadem je pro mě naprosto skvělá. Stačí napsat na tiskový odbor a pokud daná data existují, tak je dostanu.

Máte zkušenosti s Národním katalogem otevřených dat a Open data wish-listem Ministerstvem vnitra?

Databázi znám, ale nikdy jsem nic nežádal. Je to dáno i tím, že jsme malí a tušíme, co naši čtenáři chtějí. V tématech se navíc dost točíme, protože víme, co funguje. A i když si jednou za čas proklikám v katalogu NKOD co je tam nového a říkám si, že je tam spousta zajímavých věcí, tak je reálně nezvládám zpracovat.

Je pro Vás první je téma, pro které hledáte data nebo naopak máte data, ve kterých hledáte téma?

Když mám hypotézu, tak na ni netrvám. Pokud se podívám do dat a ty mi hypotézu vyvrátí, tak jsem vlastně ještě šťastnější, že jsem se něco nového naučil. Nemám rád, když novináři hledají data, která jen potvrdí jejich hypotézy, protože to vnímám za špatně odvedenou novinářskou práci. Pokud hypotézu mám, tak na základě zkušenosti ze znalosti dat z minulosti a je mi jedno, jestli šel trend nahoru nebo dolů, ale vím, že to bude fungovat a bude to zajímavé. Dat je ale spousta, konkrétně na statistickém úřadu a pomáhá mi prohrabávat se daty, které nikdo nezpracovává a hledat v nich příběhy. Na to mám ale čas jen pokud nejsem zatížen jinými tématy. Takže abych odpověděl na Vaši otázku, používám oba přístupy. Pokud mám hypotézu, vychází z minulosti. Ale když přistupuji k úplně novým datům, tak se spíše „prohrabávám“ a hledám nové věci. Občas vidím inspiraci v zahraničí, zejména v britském Economistu nebo Financial Times, a na základě toho hledám podobná data pro Českou republiku.

Jak hodnotíte kvalitu a dostupnost poskytovaných dat?

V tomhle asi budu v menšině oproti ostatním kolegům, jsem velmi spokojen. Vychází to z mých zkušeností s daty Slovenska. Když jsem zmiňoval, že se kolega věnuje datové žurnalistice pro Trend, tak se nám témata prolínají. Ukazuje se, a nejen na pandemických datech, že oproti Slovensku se máme úžasně. A to nejen proto, že máme ohromné množství data setů, se kterými lze pracovat, zatím co na Slovensku to není. Překvapuje mě, že to není případ jen pandemických dat, ale i u dat statistického úřadu, u kterých jsem si dříve myslel, že metodika bude jednotná. Když se podíváme na počet zveřejněných datových sad u nás, mám pocit, že se to dlouhodobě zlepšuje. Pro nás jako regionální vydavatelství je podstatné, že i municipality se snaží data zveřejňovat.

Jste tedy spokojen i s kvalitou poskytovaných dat?

V jednom by se kvalita všech data setů mohla zlepšit, a to v popisu atributů. Pokud není přesně popsáno, jestli se například jedná o přepočtení na sto tisíc lidí, tak si nikdy nemůžete být jistý, jestli to náhodou nejsou jablka nebo hrušky. I když vím na 95 procent, jak to je, nechci mít jen 95 procent jistoty, ale stoprocentní. Když je špatný popis, zvyšuje se šance, že data špatně interpretuji nebo se musím doptávat.

Ověřujete či zpochybňujete kvalitu poskytovaných dat?

U každých dat udělám prvotní vizualizaci, rozložení hodnot. Tak abych viděl, jestli tam nejsou nějaké výkyvy. Například u dat zemřelých na koronavirus jsem takto ověřoval extrémní hodnoty jak u nejmladších, tak u nejstarších, protože jsem chtěl mít jistotu, že se nestala chyba v reportování rodných čísel a podobně. Věřím tomu, že právě poučená veřejnost nebo novináři, kteří zpracovávají data mohou pomáhat s ověřováním dat. Zveřejnění dat může vést k jejich čištění, nezazlívám úřadům, že chyby v datech mají. Když chyby najdu, ozvu se. Častěji se mi to stává u historických dat, které třeba nikdo přede mnou nezpracovával.

Jak podle Vás státní správa výběrem poskytovaných dat formuje datovou žurnalistiku?

Vůbec by mi nevadilo, kdyby dat bylo méně, ale kvalitnějších. Například je skvělé, že ÚZIS zveřejňuje podrobné statistiky o hospitalizovaných podle dávek očkování, ale ať je nezveřejňuje bez kvalitního vysvětlení a poznámky o tom, že se jedná o přepočtení na sto tisíc obyvatel. Když jste datový novinář a přemýšlíte o tom, tak víte, že je důležité data přepočítat

a neinterpretovat je špatně. Když po stejných datech ale sáhnou novináři, kteří nemají se zpracováním dat tolik zkušeností, tak je snadné udělat chybu. Než jsou data zveřejněna a mají nějaký společenský vliv tak by lidé, kteří je zveřejňují, měli přemýšlet o tom, jak je správně popsat. Věřím v to, že méně je více. Sám, když zpracovávám data, velkou část informací vyházím a snažím se komunikovat jen to nejpodstatnější. Velmi často, když materiál dodělám, mám stejný pocit, že jsem těch informací nechal moc. Podstatná část mé práce tedy není to, co všechno do textu dám, ale spíše to, co jsem vynechal. Čím jsou data sety složitější, tím je pro novináře těžší v nich informace najít. Jsou ale i případy, kdy bych si přál dat více, a to třeba u policejních dat o nehodách.

Vysvětlujete čtenářům metodiku dat?

Moje novinářské pravidlo je, že nejlepší příběh je ten, který když přijdu domů mám chuť převyprávět manželce. Mám dva přístupy. Nesmím čtenáře přeceňovat a počítat s tím, že mu bude vše jasné. Nebude. Snažím se tedy i na jednu věc poukázat více způsoby, a to jak v samotné grafice, tak i v textu, protože malá písmena v grafikách mohou čtenáři uniknout. Zároveň ale čtenáře nemůžete podceňovat a snažit se je obalamutit. Tedy, nepřeceňuji čtenáře, a to nejpodstatnější zdůrazňuji klidně několikrát, ale zároveň je nepodceňuji.

Příloha č. 12: Přepis hloubkového rozhovoru s Kateřinou Mahdalovou (text)

Jaká je délka Vaší novinářské praxe? Studoval jste žurnalistiku, programování či jiný obor na vysoké škole?

První články mi vycházely v sedmnácti, osmnácti letech, v lokálním Děčínském deníku. Vždy jsem byla psavec, psala jsem opravdu dlouho, ale žurnalistiku jsem nikdy nestudovala. Studovala jsem bohemistiku, na gymnáziu mi ale šla i matematika, měla jsem široký základ. Profesor mě ale od studia matematiky nebo technický oborů odrazoval. V rámci studia jsem zhruba před deseti lety byla na výměnném pobytu ve Velké Británii a tam jsem zjistila, že existuje něco jako datová žurnalistika. Tehdy Guardian začínal s Datablogy. K datové žurnalistice jsem se tedy dostala oklikou. V českých médiích jsem vystřídala několik redakcí. Psala jsem pro Lidové noviny, viditelné angažmá bylo v Respektu, tam jsem se věnovala investigativně a analytickým textům. Později jsem se rozhodla zkusit založit vlastní web, po nocích jsem se začala učit, jak se dělá takový web, uměla jsem základy HTML, nikdy jsem ale nechodila na žádné kurzy. Těžila jsem z toho, co jsem se naučila. Zkoušela jsem různé vizualizační nástroje, potom jsem zjistila, že existují první kurzy datové

žurnalistiky, na které jsem se přihlásila. Věděla jsem, že nikdy nebudu masmédiem, dávala jsem o sobě vědět přes sociální sítě, kde si textů všímali primárně kolegové. Často texty přebírali, tímto jsem se vlastně infiltrovala zpátky do těch masmédií. Na jednom z hackatonů jsem potkala generálního ředitele České tiskové kanceláře, který mi nabídl spolupráci. V agentuře jsem vedla datové projekty, navrhovala jsme šablony pro vizualizace. Projekt byl ale časově limitovaný, po zhruba čtyřech letech bylo vše nachystáno a redakce uměla s daty velmi rychle a dobře pracovat. S vlastním webem jsem oslovovala redakce, jestli by stály o spolupráci. Skvěle zareagovaly nevládní organizace, se kterými jsem navázala spolupráci. Když mi skončila spolupráce s ČTK, začala jsem zjišťovat, na jaké úrovni se využívají data v různých redakcích. Věděla jsem o Českém rozhlasu, v jiných redakcích ale datová žurnalistika nebyla. Rozhodla jsem se zkusit Novinky.cz, věděla jsem, že mají velký dosah a myslela jsem, že by to mohlo fungovat. V redakci jsem působila přibližně rok, ze strany vedení byl tlak na rychlost a odnášela jsem si poměrně častou frustraci. Pro hlubší analýzy prostor nebyl. V Seznam Zprávách jsem přes rok, kde je to výrazně kultivovanější a prostor je větší. Pořád je ale tlak na produkci. Poměrně často si odnáším pocit, že bych potřebovala více času.

Jaká je tedy Vaše role v rámci redakce? Kolik lidí se datové žurnalistice věnuje?

V redakci působí Jiří Burýšek, Petr Juna. Vyloženě datová redakce nejsme, spolupracujeme s redakcí Seznam Zpráv například tak, že pokud vytvořím něco, co má opakované využití, mohou redaktoři použít pro další články, na kterých se já už autorsky nepodílím. Jsem propojena s grafiky, ale ne ve smyslu, že by mi připravovali vizualizace, to je vše na mě. Jsem více méně na vše sama, ke spolupráci jsem ale přivedla Michala Škopa, se kterým fungujeme v tandemu a je to pro mě klíčové, protože mi pomáhá s více statistickými zpracováními, volební data.

Jak se učíte nové technologie?

Nové nástroje a technologie aktivně vyhledávám. Samozřejmě koukám, jak to dělají v jiných médiích, do zahraniční redakcí. Ne všechny technologie ale jdou použít v redakčním systému Seznam Zpráv. Proto i když některé nové technologie zkouším, nepovede se je použít. Limity v redakci jsou v tomto směru poměrně velké.

Jaká témata nejčastěji zpracováváte?

Je toho spousta, poměrně hojně jsem začala vytěžovat Český statistický úřad už během práce pro ČTK. Další jsou data Ministerstva práce a sociálních věcí. Věnuji se také tématům školství, inflace, obsazenost věznic, jsou to desítky a desítky věcí. Kdybych to měla zobecnit jsou to veřejná data státní správy. Snažila jsem se dostat data například také z exekutorské komory. Nedařilo se mi to oficiální cestou, tak jsem dělala řešerše výroků představitelů exekutorské komory. Veliké úskalí je, že právě exekutorské úřady jsou soukromé čili exekutorská komora je soukromý subjekt a stát ty data nemá a vymahatelnost dat je katastrofální. Data, která jsou ochotni poskytnout jsou data, která se jim hodí. Data jsem si sama sesbírala z tiskových zpráv a zjistila, že prohlášení komory, že exekuce meziročně poklesly nebyla úplně vypovídající. Ano mírně poklesly, ale po několika letech strmého nárůstu. Data nám dodala širší kontext. Na tom se dalo hezky ukazovat, k čemu datová žurnalistika je. Ještě zmíním data, o která hodně stojím a celá léta je nemám – zdravotnická data ÚZIS. Ústav data nerad poskytuje, nebo je poskytuje v nepoužitelných formátech. Nejčastější argument byl, že kdyby mi poskytli primární data, mohla bych je zfalšovat. Snažila jsem se vysvětlit, že pokud mi data poskytnou ve formátu PDF tak je pravděpodobnější, že udělám při zpracování nějakou chybu. Dalším příkladem je Česká správa sociálního zabezpečení, u které se mi nikdy nepovedlo dostat se k někomu, kdo data spravuje. Podezírám je, že minimálně o data, která by byla novinářsky zajímavá se nikdo moc nestará. Na mé dotazy reagují tím, že vůbec neví, co po nich žádám. Témata, kterým se věnuji jsou do určité míry dána tím, čemu jsem se věnovala ještě před datovou žurnalistikou. Lidská práva, od justice přes exekuce, témata tohoto typu. Zároveň zpracovávám od každého něco, je to dáno tím, co se například zrovna projednává ve Sněmovně.

Vnímáte ochotu či neochotu státní správy data poskytovat?

Pokud potřebuji data obracím se rovnou na úřad, který data má. Úplně nejlepší je, pokud má úřad oddělení statistiky, jako například Policie České republiky. Spolupráce s nimi je dobrá, jsou ochotní. Záleží na tom, co chci. Někdy si vystačím s dostupnými daty, například Ministerstva financí. Pokud jsem žádala o něco navíc, byla to data takového typu, že s tím neměli problém. Pokud je to něco, co napomůže i té instituci, kdy já funguji jako spojnice mezi institucí a veřejností, tak je ochota data zveřejňovat vysoká. Pokud je to něco, co má naopak konfrontovat tu instituci, tak ta ochota je nízká. Zmíním ještě zkušenost s Českým hydrometeorologickým ústavem, který média příliš nepotřebuje v tom smyslu, že jim stačí, když média zveřejní už hotový výstup. Část dat jsem získala po osobním setkání, ale

potřebovala jsem data kontinuálně. Ve chvíli, kdy jsem je měla zpracovány, ale už byla zastaralá a oni spolupráci přerušili.

Máte zkušenosti s podáním žádosti o informace?

Tím, že jsem převážně v komerčních médiích, neměla jsem na podávání žádostí 106 prostor. Přestože jsou k dispozici formuláře žádosti, je to pro mě příliš mnoho času. Myslím si, že veřejnoprávní média jsou ta, kdo by toto měl primárně dělat. Nezříkám se toho, ale má kapacita je omezená.

Můžete popsat zkušenosti s poskytováním dat během pandemie COVID-19?

Snažila jsem se dopočítat prevalenci testů, musela jsem každé ráno obejít 14 webů Krajských hygienických stanic a ručně z nich data posbírat, a to za každý okres zvlášť. Krajské hygienické stanice byly nepružné, nebo se dobře komunikovalo s jednou, či dvěma. Chtěla jsem si prevalenci dopočítat sama, nechtěli mi dát zdrojová data, ale už vypočítanou prevalenci. Jak mohu bezpečně ověřit, že jsou údaje přesné? Úroveň je tak strašně rozdílná, je to pořád krajská hygienická stanice s nadřízeným úřadem, ale ta spolupráce je s každým kraje absolutně odlišná. Na jednu stranu je nejspíše dobře, že mají svou autonomii, ale v případě pandemie koronaviru to zoufale nefunguje. Data jsou velmi různého charakteru, velmi různé spolehlivosti. Rozepře uvnitř některých úřadů ovlivňovaly to, jestli mi poskytnou nebo neposkytnou data. Od samého začátku pandemie s ÚZIS bojuji. Začala jsem podléhat pocitu, jestli vedení úřadu data schválně nerozmlžuje. Nebylo možné si cokoliv ověřit, cokoliv dopočítat. Teď je to 13 měsíců od podání žádosti o data nakažených podle profesí. Data nejsou tajná, ÚZIS je poskytl například České televizi, která je měla ve vysílání. Poslala jsem tedy žádost ÚZIS o stejná data, nikdy mi je neposlali. Takových situací je spousta, ústav data poskytuje výběrově. Mají otevřené datové sady, ze kterých těžíme. Ale spousta dat jsme si sehnali až díky různým poslancekým iniciativám a podobně. Byly to oficiální data ÚZIS ale nebyly poskytnuty přímo ústavem. Začali na mě útočit, že jedním neeticky. Nikdy ale nerozporovali, že to jsou jejich data. Sami data zveřejňovat nechtěli, nebo je poté začali zveřejňovat s odstupem, například počty nakažených v obcích. Soubory jsou navíc obtížně zpracovatelné, mají přes 2 GB.

Proč je takové chování ze stran poskytovatelů dat problematické?

Neposkytování dat během pandemie vnímám jako katastrofu, protože umírají lidé a jsem přesvědčena o tom, že kdybychom na základě dat loni věděli, ve kterých profesích se nákaza šíří, tak se s tím dalo něco dělat a mohli na to adekvátně reagovat politici. Na místo toho zástupci ústavu prezentovali data v reportáži České televize, jak chtěli oni. Žádná hlubší analýza neproběhla, chtěla jsem ji zpracovat, ale nebylo mi to umožněno. Podobné je to o datech hospitalizovaných pacientů, u kterých se sleduje, zda jsou či nejsou očkovaní. ÚZIS publikuje absolutní čísla, která samozřejmě mohou lidé špatně interpretovat. Apelovala jsem na ústav, že absolutní čísla nestačí, že je potřeba dělat relativní přepočty na věkové skupiny. Ideální by bylo mít dodatečné informace jako například kdy měl daný člověk ukončené očkování, z toho můžeme usuzovat, zda poklesly protilátky a podobně. Jsou to klíčové věci, které kdybychom měli, tak by v debatě napomáhaly vyvracet dezinformace. Nepomáhá nám, když ústav jednorázově některá data zveřejní, potřebujeme je mít kontinuální a trendy sledovat. ÚZIS si z nějakého důvodu uzurpuje právo na jedinou správnou interpretaci a další zpracování dat. Ale nestíhají to, neposílili datové týmy. Úřad je přetížen, zároveň ale za rok a půl nezvládnout organizaci není omluvitelné. Výsledkem je, že data mi nedají jako novináři, protože mě chápou často až nepřátelsky. Vysvětluji jim, že nejsem nepřítel ale spojnice mezi nimi a veřejností a data do nějaké míry analyzovat umíme a pokud bychom si nevěděli rady, tak se na ně obrátíme. Na to ale neslyší a výstupy, které poskytují poslancům jsou tragické. Nenárokuji po zákonodárcích, aby uměli číst či zpracovávat data. Spíše si nárokuji, aby jim ÚZIS dával data v takové kvalitě a podobě, aby mohli kvalifikovaně rozhodovat, a to se doteď neděje. To vnímám jako tu tragédii, že ústav na datech sedí, a když už nějaká data pustí, tak z nich politici nic relevantního nevyčtou. Logicky je ta pandemie o to hůř zvládnutá. Je to velmi nebezpečné a sama za sebe viním ÚZIS z toho, že je průběh pandemie na českém území výrazně horší, než by mohl.

Jaké jsou hlavní limity Vaší práce?

Hlavní limity vidím spíše než v možnostech vizualizace, kterých je již tolik, v tom, že dat, která bychom mohli zpracovávat je výrazně více než kolik jich je dostupných. Samozřejmě také potřebujeme, aby redakce, ve které působíme, měla pochopení pro to, v čem je hodnota datové žurnalistiky.

Ověřujete či zpochybňujete kvalitu poskytovaných dat?

Ověřuji, porovnávám buď s výročními zprávami nebo výstupy toho typu a koukám, jestli třeba někde není rozdíl. Několikrát jsem takto reportovala chyby. Stává se to. Před lety jsem dokonce reportovala chybu v datech Bílému domu. Mám pro to pochopení, většinou poskytovatelé dat jsou rádi, když je na chybu upozorním. Kvalita je kolísavá. Nemůžu říct, zda je to půl na půl. Ty extrémní polohy jsou opravdu velké. Setkávám se s mlčením, nepochopením toho, co chci, špatnými výstupy anebo dostanu data v profesionální kvalitě, vědí přesně, co chci, nabízí něco navíc.

Vysvětľujete čtenářům metodiku dat?

Ze začátku jsem si říkala, že jsou některé výstupy jasně srozumitelné a vysvětlení jsem k nim nedávala. Pak jsem si ale v diskusi pod články všimla, že zůstaly nepochopeny. Došlo mi, že to nejspíše dělám špatně a začala jsem vysvětlovat, jak to vzniklo. U velkých vizualizací typu mapy nákazy koronavirem jsem začala vysvětlení dávat, to stejné u výpočtů volebních modelů. Je to rozhodně věc, kterou musím v komunikaci zlepšit, ať už s redakcí tak s veřejností. Tam, kde si myslím, že je to opravdu velmi jednoduché na čtení tam legendu nedávám, popřípadě to zmíním ve dvou větách přímo v textu. Zdroje dat uvádím vždy. Nechci se, jakkoliv vysmívat čtenářům, protože mám o nich nějakou představu a nemůžu srovnávat sebe, která data zpracovává a živí se tím s někým, kdo přijde z práce a rozklikne si někde článek na internetu. Bývám překvapená, co všechno se musí ještě dovysvětlit, ale to jde spíše za mnou. Zároveň se snažím si čtenáře dovychovat nebo dovzdělat. Do velké míry to funguje, na koronaviru to jde vidět dobře. Leckdy mě čtenáři upozorní na chybu v datech. Jsem ráda, když žádnou chybu neudělám, ale nebudu zastírat, že se to nestává. Člověk se ocitá mezi mlýnskými koly z obou stran. Cítím, že novinářina je o tom oslovovat veřejnost na druhé straně máme akademickou obec, která si potrpí na přesnost a nechápu, že musím zjednodušovat.

Můžete popsat zkušenosti s poskytováním dat během pandemie COVID-19?

Snažila jsem se dopočítat prevalenci testů, musela jsem každé ráno obejít 14 webů Krajských hygienických stanic a ručně z nich data posbírat, a to za každý okres zvlášť. Krajské hygienické stanice byly nepružné, nebo se dobře komunikovalo s jednou, či dvěma. Chtěla jsem si prevalenci dopočítat sama, nechtěli mi dát zdrojová data, ale už vypočítanou prevalenci. Jak mohu bezpečně ověřit, že jsou údaje přesné? Úroveň je tak strašně rozdílná, je to pořád krajská hygienická stanice s nadřízeným úřadem, ale ta spolupráce je s každým

kraje absolutně odlišná. Na jednu stranu je nejspíše dobře, že mají svou autonomii, ale v případě pandemie koronaviru to zoufale nefunguje. Data jsou velmi různého charakteru, velmi různé spolehlivosti. Rozepře uvnitř některých úřadů ovlivňovaly to, jestli mi poskytnou nebo neposkytnou data. Od samého začátku pandemie s ÚZIS bojuji. Začala jsem podléhat pocitu, jestli vedení úřadu data schválně nerozmlžuje. Nebylo možné si cokoliv ověřit, cokoliv dopočítat. Teď je to 13 měsíců od podání žádosti o data nakažených podle profesí. Data nejsou tajná, ÚZIS je poskytl například České televizi, která je měla ve vysílání. Poslala jsem tedy žádost ÚZIS o stejná data, nikdy mi je neposlali. Takových situací je spousta, ústav data poskytuje výběrově. Mají otevřené datové sady, ze kterých těžíme. Ale spousta dat jsme si sehnali až díky různým poslaneckým iniciativám a podobně. Byly to oficiální data ÚZIS ale nebyly poskytnuty přímo ústavem. Začali na mě útočit, že jedním neeticky. Nikdy ale nerozporovali, že to jsou jejich data. Sami data zveřejňovat nechtěli, nebo je poté začali zveřejňovat s odstupem, například počty nakažených v obcích. Soubory jsou navíc obtížně zpracovatelné, mají přes 2 GB.

Proč je takové chování ze stran poskytovatelů dat problematické?

Neposkytování dat během pandemie vnímám jako katastrofu, protože umírají lidé a jsem přesvědčena o tom, že kdybychom na základě dat loni věděli, ve kterých profesích se nákaza šíří, tak se s tím dalo něco dělat a mohli na to adekvátně reagovat politici. Na místo toho zástupci ústavu prezentovali data v reportáži České televize, jak chtěli oni. Žádná hlubší analýza neproběhla, chtěla jsem ji zpracovat, ale nebylo mi to umožněno. Podobné je to o datech hospitalizovaných pacientů, u kterých se sleduje, zda jsou či nejsou očkovaní. ÚZIS publikuje absolutní čísla, která samozřejmě mohou lidé špatně interpretovat. Apelovala jsem na ústav, že absolutní čísla nestačí, že je potřeba dělat relativní přepočty na věkové skupiny. Ideální by bylo mít dodatečné informace jako například kdy měl daný člověk ukončené očkování, z toho můžeme usuzovat, zda poklesly protilátky a podobně. Jsou to klíčové věci, které kdybychom měli, tak by v debatě napomáhaly vyvracet dezinformace. Nepomáhá nám, když ústav jednorázově některá data zveřejní, potřebujeme je mít kontinuální a trendy sledovat. ÚZIS si z nějakého důvodu uzurpuje právo na jedinou správnou interpretaci a další zpracování dat. Ale nestíhají to, neposílili datové týmy. Úřad je přetížen, zároveň ale za rok a půl nezvládnout organizaci není omluvitelné. Výsledkem je, že data mi nedají jako novináři, protože mě chápou často až nepřátelsky. Vysvětluji jim, že nejsem nepřítel ale spojnice mezi nimi a veřejností a data do nějaké míry analyzovat umíme a pokud bychom si

nevěděli rady, tak se na ně obrátíme. Na to ale neslyší a výstupy, které poskytují poslancům jsou tragické. Nenárokuje po zákonodárcích, aby uměli číst či zpracovávat data. Spíše si nárokuje, aby jim ÚZIS dával data v takové kvalitě a podobě, aby mohli kvalifikovaně rozhodovat, a to se doteď neděje. To vnímám jako tu tragédii, že ústav na datech sedí, a když už nějaká data pustí, tak z nich politici nic relevantního nevyčtou. Logicky je ta pandemie o to hůř zvládnutá. Je to velmi nebezpečné a sama za sebe viním ÚZIS z toho, že je průběh pandemie na českém území výrazně horší, než by mohl.