

Univerzita Karlova

Filozofická fakulta

Ústav informačních studií a knihovnictví



Diplomová práce

Bc. Marek Cieslar

Osobní data jako komodita:

**Jak trh s daty funguje, jakou mají cenu a co o tom čeští uživatelé
internetu vědí**

Personal data as a commodity: How does the data market work, what is data worth
and what Czech internet users know about it

Praha 2022

Vedoucí práce: Mgr. Vít Šisler, Ph.D.

Poděkování

Rád bych poděkoval vedoucímu diplomové práce Mgr. Vítu Šislerovi, Ph.D., který mi pomohl práci nasměrovat správným směrem a který mi v průběhu jejího vzniku poskytl mnoho rad a doporučení.

Prohlášení

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů a že práce nebyla využita v rámci jiného vysokoškolského studia či získání jiného nebo stejného titulu.

V Praze, dne 26. července 2022

Bc. Marek Cieslar

Abstrakt

Nedílnou součástí současného internetu je sběr osobních dat jeho uživatelů. Cílem této práce je zmapovat trh s osobními daty; jak funguje, jaké vlivy ho zformovaly do dnešní podoby a jak ho vnímá veřejnost. Teoretická část se opírá o existující akademické zdroje, články z odborných i zpravodajských médií a dalších relevantních pramenů, které souvisejí se sběrem, prodejem a hodnotou osobních dat. Jedním ze záměrů teoretické části je také představit principy, jak lze data nacenit, protože dosud zřejmě neexistuje zdroj, který by obsáhle shrnoval existující metody, jak se cena osobních dat sbíraných na internetu může z pohledu uživatelů i trhu určit. Zjištění z teoretické části prohlubuje kvalitativní výzkum, který metodou hloubkových rozhovorů rozkrývá, nakolik jsou si pražští vysokoškolští studenti vědomí toho, co se s jejich daty děje, jakou pro ně mají hodnotu a jaký mají na současné nastavení trhu názor. I přes celospolečenskou závažnost problematiky jde zřejmě v tuzemském prostředí o jeden z prvních kvalitativních průzkumů na toto téma a zjištění, která přináší, mohou posloužit jako odrazový můstek pro budoucí výzkum.

Klíčová slova: osobní data, uživatelská data, online sběr dat, extrakce dat, soukromí na internetu, cena dat, hodnota dat, trh s daty, agregátoři dat, cílená reklama, personalizace reklam

Abstract

Personal data extraction is an inherent part of today's internet. The intention of this work is to outline the market with personal data; how it functions, which circumstances have formed it and how it is being perceived by the public. The theoretical part relies on existing academic knowledge, media articles and other relevant sources which focus on data market and aggregation. One of the theoretical part's objectives is to present existing principles on how the price of personal data can be determined by users and by the market. Findings from the theoretical part will be enriched by a qualitative research which, by the means of interviews, reveals to what extent are Prague university students aware of online data collection, how they value the data and what is their opinion on that matter. Despite the gravity that online data extraction imposes on society, the qualitative research described in this work is probably one of the first in Czechia that focuses on this topic and that can therefore serve as a starting point for future exploration.

Key words: personal data, user data, data extraction imperative, online privacy, internet privacy, data value, data market, data brokers, data aggregators, targeted ads, personalized advertising

Obsah

Úvod	7
1. Byznysový zážrak	9
1.1 Google	9
1.2 Sociální sítě	11
1.3 Těžba dat jako nutnost	13
2. Komodifikace dat	17
2.1 Kategorizace osobních dat	19
2.2 Hodnota dat pro firmy	22
2.3 Hodnota dat pro uživatele	25
3. Nejrozšířenější metody sbírání dat na internetu	30
3.1 Geolokalizace a unikátní identifikátory	30
3.2 Cookies	32
3.3 Pixels a beacons	36
3.4 Sociální sítě a aplikace	37
4. Shrnutí teoretické části a rešerše	38
5. Povědomí Čechů o sběru osobních dat na internetu	39
5.1. Cíl výzkumu a výzkumné otázky	39
5.2 Metoda, nástroje a průběh výzkumu	39
5.3 Výběr a skladba výzkumného vzorku	40
5.4 Pilotní výzkum	41
5.5 Analýza dat	41
5.6 Etika výzkumu	42
5.7 Limity výzkumu	42
5.8 Výsledky výzkumu	43
5.8.1 Průběh rozhovorů a charakteristika účastníků	43
5.8.2 Analýza hloubkových rozhovorů s participanty	46
Účel uživatelských dat	47
Domnělá cena dat	48
Kroky vedoucí k zanechávání menšího množství dat	51
Důvody, proč svá data uživatelé poskytují	53
Jsou data z pohledu uživatelů platidlo?	55
5.9 Diskuze	58
Závěr	62
Seznam obrázků	64

Seznam tabulek	64
Seznam příloh	64
Zdroje	65

Úvod

Již v roce 2000 Federal Trade Commission věděla, že 99 % internetových společností sbírá data o uživateli svých služeb (Seligman & Taylor, 2000). Subjektů, které s daty obchodují nebo je komerčně využívají, je dnes nespočet a se stále se zvyšujícím množstvím času a aktivit, které na internetu vykonáváme, o sobě zanecháváme stále detailnější data. Dostali jsme se do doby informační ekonomiky (Samuelson, 2000) nebo dokonce do doby tzv. kapitalismu dohledu, jak stávající nastavení zhodnotila Shoshana Zuboff ve své knize *The Age of Surveillance Capitalism* z roku 2019. Není jediná, kdo dospěl k závěru, že sběr a následná monetizace uživatelských dat se stala nejatraktivnějším byznys modelem dnešní doby.

Nejen technologické společnosti se snaží po vzoru Googlu, Facebooku a dalších od svých uživatelů extrahovat nejrůznější data a vydělat na nich. Výdělek skýtá zpravidla personalizovaná reklama. Google byl prvním subjektem, který tento byznys model proslavil: testovat ho začal v roce 2001 a výdělek společnosti tehdy rázem vyskočil o 400 procent na 86 milionů amerických dolarů. Díky tomu se firma poprvé ocitla v černých číslech (Zuboff, 2019). Následující rok už byl výdělek zhruba 400 milionů a tempo růstu od té doby nepřestalo zpomalovat (Statista, 2022).

Od přelomového objevu Googlu, že i „obyčejná“ data v podobě dotazů zadaných do vyhledávače lze využít ke zvyšování účinnosti reklamy, raketovým tempem rostly nejen výdělky mnohých společností operujících se zmíněným byznys modelem, sílila ale i honba za objevem nových kanálů, jejichž pomocí se dají data sbírat a reklama zpřesňovat. Populární americký publicista Bruce Schneier zaměřující se na kyberbezpečnost už v roce 2006 tvrdil, že online sledovací nástroje jsou všudypřítomné. Jak se dočtete na následujících stránkách, firmy dnes nemají zájem pouze o dnes už všeobecně známé soubory cookie, ale i o zdánlivé detaily v podobě půdorysu vašeho bytu, jak je v něm rozložený nábytek a kde se co na podlaze nachází.

Svět by byl pravděpodobně úplně jiný nebýt ekonomického zázraku webových služeb stojících na sbírání dat o uživateli. Nejenže hluboce zakořenily v životech miliard lidí, ale mnohé společnosti se díky svému extrémně silnému ekonomickému a informačnímu kapitálu staly dominantním globálním politicko-ekonomickým hráčem (Morozov, 2011).

Ve výsledku se tak společnost dostává do určité paradoxní situace – jak v roce 2021 uvedla americká nezisková organizace Electronic Frontier Foundation (EFF), která už od

90. let bojuje proti zneužívání technologií proti zájmům demokratické společnosti: největší technologické firmy jsou v natolik dominantním postavení, že na nich přirozeně závisejí i subjekty, které nesouhlasí s tím, co dělají. Například politici nezájmově upozorňují na to, že technologičtí giganti představují hrozbu pro demokracii, na druhou stranu ale těmto společnostem platí, aby prostřednictvím jejich technologií a sociálních sítí oslovili voliče. Druhým příkladem jsou podle EEF novináři, kteří sice upozorňují na rizika sběru dat, na druhou stranu pak ale jejich online články obsahují sledovací technologie, které sbírají data pro ty samé subjekty, jež v článku novináři kritizují (Cyphers & Doctorow, 2021).

Kvůli čemu jsme se do tohoto okamžiku dostali? Je to ochota lidí o sobě prozrazovat informace a dostupnost technologií, které je umožňují sbírat. Jak popsala Canellopoulou-Bottis a Bouchagiar (2018), z osobních dat se stala komodita, popřípadě slovy van Dijck (2014) platidlo, kterým hradíme využívání internetových služeb. Zuboff (2019) uživatelská data dokonce přirovnává k surovině.

Cílem této práce je zmapovat trh s osobními daty; jak funguje, jaké vlivy ho zformovaly do dnešní podoby a jak ho vnímá veřejnost. Kromě stručného výčtu dějinných událostí, které byly pro současný stav trhu s daty zásadní, teoretická část této práce obsahuje také teoretické koncepty, jak lze na problematiku sběru dat nahlížet. Důležitou součástí teoretické části je také přehled principů, jak lze data nacenit. Teoretická část končí představením základních metod a druhů osobních dat, která se v online prostředí sbírají. Následuje kvalitativní výzkum, jehož cílem je hlouběji porozumět, jak uživatelé internetu na online trekování reagují. Zvolenou metodou jsou hloubkové rozhovory a výzkumným vzorkem jsou pražští vysokoškolští studenti. Závěrem celé této práce je diskuze nad výsledky výzkumu s ohledem na existující poznatky představené v její teoretické části.

1. Byznysový zázrak

Matthew Crain ve své knize *Profit Over Privacy* z roku 2021 konstatuje, že reklama se stala vítězem komercializace internetu. Podle něj ale nejde o reklamu, jak ji známe z offline světa, ale o reklamu, která se zakládá na neomezených možnostech sběru dat o uživatelích. Následující stránky jsou věnované objevu tohoto byznys modelu a jeho vývoje do dnešních dní. Zvláštní pozornost bude věnována službám Google a Facebook, protože se běžně označují za služby, které přinejmenším postavily základy dnes už velmi etablovaného trhu s daty uživatelů internetu. Jsou to také ony, kdo na trhu s digitální reklamou i v posledních letech ještě nadále upevňuje své už beztak dominantní postavení, a to navzdory obavám o soukromí a přísnějšího dohledu ze strany regulátorů (Gray, 2021).

1.1 Google

Okamžik zázraku na poli internetových služeb se odehrál v roce 2001, a to na konci tzv. internetové horečky, známé také jako DotCom Bubble. Jak v roce 2020 připomněla datová vědkyně Jeanne E. Daniel v článku pro server Medium, šlo o dobu propadu akcií většiny internetových a technologických startupů, včetně Applu a Amazonu. Google sice tenkrát ještě nebyl veřejně obchodovaný na burze¹, nicméně ani on, byť se do té doby těšil velké důvěře mezi svými investory i uživateli, krizi neunikl. Vedení Googlu včetně zakladatelů Sergeye Brina a Larryho Page se dostalo pod velký tlak, který vyústil v požadavek přetvořit Google tak, aby se dal jeho byznys škálovat a aby výdělkův skokově narostly. Jak popisuje například reportér technologického magazínu The Verge Nick Statt ve svém článku z roku 2019, ve firmě se tenkrát už nějakou dobou vědělo, že cesta, jak docílit větších tržeb, vede skrz zapojení efektivní reklamy do jejího vyhledávače – už tehdy pay per click reklamu využíval například konkurenční vyhledávač Overture, který v Googlu samozřejmě monitorovali – jenže Brin a Page (1998) se vůči zapojování cílených reklam do podnikání Googlu dlouho vymezovali.²

¹ Google vstoupil na burzu na jaře roku 2004. (Hild, 2008)

² Zakladatelé Googlu Larry Page a Sergey Brin (1998) dlouho odmítali postavit byznys model své firmy na reklamě. Očekávali totiž, že vyhledávače financované reklamou budou ze své podstaty zaujaté směrem k inzerentům a vzdálené potřebám spotřebitelů.

Nedlouho poté ale s příchodem již zmíněné krize internetových společností zakladatelé Brin a Page nakonec otočili. Podrobnosti o přechodu Googlu na nový byznys model nastiňuje Shoshana Zuboff (2019) své knize *The Age of Surveillance Capitalism*. Píše:

„Page a Brin se zdráhali zavést inzerci. Nicméně s přibývajícím důkazy, že reklamy mohou zachránit firmu před krizí, se jejich přístup změnil. Zachránit firmu také znamenalo zachránit sebe samotné před tím, aby byli pouze dalšími velmi chytrými chlápky, kteří nepřišli na to, jak vydělat opravdové peníze... Reakcí zakladatelů Googlu na tento strach ... tak bylo vyhlášení „výjimečného stavu“, v němž se usoudilo jako nezbytné ponechat stranou zásady a hodnoty, které provázely Google obdobím svého vzniku a rané existence.“³ (str. 84)

Ve stejné knize Zuboff (2019) cituje pasáž z rozhovoru tehdejšího investora Googlu Michaela Moritze pro server *Foreign Affairs*, ve kterém Moritz shrnul, jak se Google během krize proměnil ve firmu, kterou následně čekal nevídaný ekonomický boom:

*„Prvních 12 měsíců nebyla pro Google procházka růžovým sadem, protože firma na začátku nedělala byznys, který nakonec začala využívat. Nejdřív šla jiným směrem, kterým byl prodej licencí svých vyhledávacích mechanismů větším internetovým společnostem a korporacím... Prvních šest, sedm měsíců létaly peníze divokým tempem z okna. A pak, velmi důmyslně, Larry ... a Sergey ... a ostatní se zaměřili na model, který viděli rozvíjet jinou firmu, *Overture*⁴, a tím bylo řazení inzerce. Viděli, jak tento model mohou vylepšit, zkvalitnit a osvojit si ho. To byznys transformovalo.“⁵ (str. 85)*

³ Volně přeloženo z anglického originálu: *„Page and Brin had been reluctant to embrace advertising, but as the evidence mounted that ads could save the company from crisis, their attitudes shifted. Saving the company also meant saving themselves from being just another couple of very smart guys who couldn't figure out how to make real money... The Google founders' response to the fear ... effectively declared a “state of exception” in which it was judged necessary to suspend the values and principles that had guided Google's founding and early practices.“*

⁴ Společnost *Overture* v roce 2002 zažalovala Google za ukradení jejich patentu. (Olsen, 2002)

⁵ Volně přeloženo z anglického originálu: *„The first 12 months of Google were not a cakewalk, because the company didn't start off in the business that it eventually tapped. At first it went in a different direction, which was selling licences for its search engines to larger internet properties and corporations... Cash was going out of the window at a feral rate during the first six, seven months. And then, very ingeniously, Larry... and Sergey... and others fastened on a model that they had seen this other company, *Overture*, develop, which was ranked advertisements. They saw how it could be improved and enhanced and made it their own, and that transformed the business.“*

Jinými slovy, data, o kterých se společnost nejprve domnívala, že jejich jediným využitím je zpřesňování výsledků vyhledávání, začala zpeněžovat cílením reklam. Google už rázem data nesbíral pouze k poskytování lepší personalizované služby, ale také pro zhodnocování svého reklamního prostoru. Jak potvrzují výroční zprávy Googlu, rok 2001 byl pro Google vůbec prvním rokem, který uzavřel se ziskem. Dva roky poté už vykázal čistý zisk 105 000 amerických dolarů a v roce 2005 dokonce skoro půldruhého milionu amerických dolarů.

To, že cílení reklam pomocí behaviorálních dat dokáže být oproti běžné reklamě extrémně efektivní, ukázali ve svém experimentu z roku 2009 Jun Yan et al.. Zjistili, že komerční vyhledávače dokáží na základě behaviorálního cílení na uživatele zvýšit pravděpodobnost prokliku na reklamu až o 670 %.

1.2 Sociální sítě

S podobným byznysovým modelem si pohrávala i sociální síť Facebook v čele se svým zakladatelem Markem Zuckerbergem, a to už krátce od své expanze z univerzitní půdy do mezinárodního internetového prostoru v roce 2006. Cílené reklamy nazvané Facebook Ads Zuckerberg představil v roce 2007. V tiskové zprávě z té doby se píše, že: „...zakladatel a CEO Facebooku Mark Zuckerberg dnes zavedl Facebook Ads - reklamní systém pro obchodníky, který je propojí s uživateli a který reklamní sdělení zacílí na žádoucí publikum.“ (Facebook, 2007).⁶

Jak poukazuje článek serveru The New York Times, Facebook už tehdy sledoval, co jeho uživatelé navštěvují a co na internetu nakupují, a to i v okamžiku, kdy byli z Facebooku odhlášeni (Story & Stone, 2007). Šlo tedy o aktivitu, která neměla pro uživatele sociální sítě žádný význam – jediným cílem bylo získávat data o svých uživateli pro personalizaci reklam, které se jim pak na Facebooku ukazovaly.⁷

Byznys model založený na monetizaci uživatelských dat formou personalizované reklamy Facebooku, podobně jako předtím v Googlu, dopomohl k exponenciálnímu růstu

⁶ Volně přeloženo z originální tiskové zprávy v angličtině: „Facebook founder and CEO Mark Zuckerberg today introduced Facebook Ads, an ad system for businesses to connect with users and target advertising to the exact audiences they want.“ (Facebook, 2007)

⁷ Ostatně toto dělají společnosti dodnes. Člověk nemusí být uživatelem ani Googlu ani Facebooku, přesto o vás těmto firmám internetové stránky, které prohlížíte, poskytují data o vaší návštěvnosti a chování. Společnost DuckDuckGo (2020), která se snaží Googlu konkurovat, tvrdí, že 86 % stránek světového internetu poskytuje data Googlu o svých návštěvnicích. V roce 2014 obdobnou analýzu provedl Timothy Libert, který sledoval, s kým sdílí data o svých návštěvnicích zhruba milion nejpopulárnějších webových stránek podle Alexy. Došel k závěru, že většina z nich data sdílí s třetími stranami, zejména s Googlem a pak i Facebookem

tržeb, byť v prvních letech nestačily na pokrytí enormních provozních nákladů. Web TechCrunch zaměřený na zpravodajství a publicistiku z celosvětové technologické a startupové scény v roce 2012, kdy Facebook vstupoval na burzu, na základě oficiálních finančních podkladů pro IPO uvedl, že v roce 2004 Facebook utržil 382 000 amerických dolarů, následující rok 9 milionů a v roce 2006 to bylo dokonce 48 milionů amerických dolarů. Podle TechCrunchu firma reportovala rozdíly mezi tržbami a výdaji až po roce 2006 a konečně v roce 2009 z 777 milionových tržeb Facebook vykázal čistý zisk 290 milionů amerických dolarů (Tsotsis, 2012).

Ukázalo se, že jde o nezpochybnitelně skvělý byznysový model, neboť společnosti Facebook i Google dosáhly bezprecedentního bohatství. Dle svých oficiálních statistik z roku 2021 utržil Facebook v roce 2020 skoro 86 miliard amerických dolarů, z toho 84 miliard pocházelo z inzerce. Ve stejném roce poukázala americká stanice CNBC na fakt, že Google vykázal ještě víc, a to skoro 183 miliard amerických dolarů, z nichž 80 % pocházelo z reklamy (Graham & Elias, 2021). To, že valuace firmy přímo závisí na její výdělky z inzerce, potvrzuje i největší, 26% propad ceny akcií Facebooku, který se stal začátkem února 2022 – média, včetně výše citované CNBC, informovala o tom, že dle analytiků šlo o reakci na odhad růstu společnosti v dalším kvartálu. Ten měl být mnohem nižší, než se očekávalo. Mezi důvody, které měly zapříčinit očekávaný slabší výkon Facebooku, byla nová opatření v oblasti ochrany soukromí ze strany společnosti Apple, která se negativně dotýkají výdělků z cílení reklamy (Feiner, 2022).

Právě společnost Facebook, která se v říjnu 2021 přejmenovala na Meta (Meta, 2021), a Google (ten od roku 2015 spadá pod holding Alphabet), ukázaly, jak efektivní je založit svůj byznys na sběru dat a prodeji personalizované reklamy. Publicista Cory Doctorow soustavně upozorňuje na to, že aby tento byznys model fungoval, je také nutné, aby lidé měli limitované možnosti sběr dat omezit, což podle něj potvrzuje například zmíněný dramatický pokles hodnoty Facebooku v reakci na to, že se zařízení od společnosti Apple začaly výslovně ptát na to, jestli si s Facebookem přejí svá data sdílet (Doctorow, 2022).

Podle Zuboff (2019) byl každopádně ekonomický zázrak dnes největších internetových společností natolik atraktivní, že se de facto většina více i méně významných firem chtěla na objevu přizpůsobit a nalézt způsob, jak alespoň část nového byznys modelu ve svém oboru podnikání využít.

1.3 Těžba dat jako nutnost

Jednou z prvních technologických společností, která se inspirovala Googlem a Facebookem, byl Microsoft. Software stojící za jeho internetovým vyhledávačem Bing Microsoft přetvořil podle vzoru Googlu. Využil také svého dominantního postavení na trhu s operačními systémy a určil v nich Bing výchozím vyhledávačem a asistentem. Zhruba rok od této transformace, tedy na podzim roku 2015, viceprezident Microsoftu pro reklamu Rik van der Kooi na předváděcí akci pro inzerenty prohlásil, že: „Čtyři z pěti dotazů zadaných do vyhledávače jdou za Googlem, zatímco ve vyhledávací liště Windows čtyři z pěti dotazů vyhledává Bing.“ (Marvin, 2015). Ryan Gavin, generální manažer Microsoftu pro vyhledávání, cloud a obsah, na stejné akci prohlásil, že „Vyhledávání je klíčovým komponentem k jejich monetizační strategii.“ Následující rok Microsoft odkoupil za více než 26 miliard amerických dolarů sociální síť LinkedIn (Microsoft, 2016), čímž pro změnu následoval vzor Facebooku. CEO Microsoftu Satya Nadella, který stál ze pivotem Microsoftu vstříc sběru uživatelských dat a personalizované reklamě, k odkoupení LinkedInu uvedl, že to přinese „...nové příležitosti pro monetizaci skrz individuální a institucionální členství a cílenou reklamu.“ (Tweedie, 2016) Tato rozhodnutí přispěla k prosperitě Microsoftu, jehož akcie od té doby po několikaleté stagnaci začaly každým rokem výrazně růst. Pomocí online nástroje Yahoo Finance (2022) lze dohledat, že zatímco začátkem roku 2014 se jedna akcie Microsoftu prodávala za zhruba 34 amerických dolarů, v roce 2017 už to bylo skoro dvakrát tolik. V lednu 2020 stála jedna akcie přes 158 amerických dolarů. Server VisualCapitalist v roce vydal infografiku, podle které z LinkedInu a vyhledávače Bing pocházelo zhruba 11 % tržeb Microsoftu, tedy přes 12 miliard amerických dolarů (Desjardins, 2019).

Microsoft zdaleka nebyl jedinou firmou, která si uvědomila, jak lukrativní může sběr dat a na něm založené cílení reklamy být. Od okamžiku, kdy ekonomický zázrak Googlu a Facebooku poukázal na to, jakou hodnotu data mají, podnikatelský svět experimentuje s nespočtem více či méně zajímavých a více či méně úspěšných aktivit, jak data sbírat a využít ve svůj prospěch. Siemens (2014) publikoval video, ve kterém tvrdí, že „...data jsou všude a získávají každým dnem na důležitosti.“ Ve videu také zazní tvrzení, že „...bychom z dat měli udělat aktivum a proměnit ho v hodnotu.“ Společnost IBM ve stejném roce podle Sadowskeho (2019) uvedla, že „...dnes se všechno skládá z dat“ a že „...už nejde data jen pasivně sbírat, ale je třeba je také aktivně vytvářet.“ (str. 2)

Některé firmy přišly s produkty, které byly stvořené právě s cílem sbírat od uživatelů data. O některých víme hlavně skrze právní spory, jež způsobily. Řada z nich se dostala do

povědomí skrze média. Redaktor CNN Criss (2016) například informoval o stížnosti podané americkému regulátorovi FTC na „chytrou“ panenku Cayla od společnosti Genesis Toys, která byla vybavená kamerou a mikrofonom s připojením k internetu. Hračka tak nejen že internetové připojení potřebovala k hledání odpovědi na otázky a konverzaci, kterou mohla vést s dětmi, ale podle stěžovatelů také internetové připojení prokazatelně využívala k zasílání dat softwarové společnosti Nuance Communications vyvíjející umělou inteligenci s důrazem na rozpoznávání řeči. Britský zpravodajský server Guardian v článku Oltermanna (2017) informoval, že německé úřady tuto „hračku“ označily jako nepovolené narušení soukromí občanů a rodičům doporučily panenku zničit. Podobně jako Cayla přitom funguje většina domácích asistentů připojených k internetu od Amazonu, Googlu, Applu nebo Microsoftu. Levin (2019) v článku pro Inc. vyjádřil své znepokojení, že boom IoT zařízení⁸ včetně „chytrých“ televizí osazených kamerami a mikrofony technologické firmy vyvinuly a propagují je se záměrem extrahovat data svých uživatelů. To však výrobci popírají. Například společnost Amazon (2022), která prodává populární systémy chytré domácnosti Echo a Alexa na svých webových stránkách píše:

„Možná uvažujete nad tím, jestli Alexa poslouchá, o čem se bavíte, či všechno, co řeknete. Možná chcete vědět, jestli vás Alexa poslouchá, i když přímo se zařízením Echo neinteragujete. Odpověď na tyto otázky je ne. Alexa a všechna zařízení Echo jsou navržena s důrazem na vaše soukromí.“⁹

V některých případech, na rozdíl například od zmiňované panenky Cayla, vůbec nebylo nutné vyvíjet úplně nové technologie a produkty s cílem data sbírat. Stačilo využít už existující technologii a pouze se na zužitkování uživatelských dat více zaměřit. Příkladem tohoto přístupu mohou být například mobilní operátoři, zejména ve Spojených státech. Někteří z nich ještě v roce 2018 prodávali třetím stranám data svých klientů, a to bez jejich vědomí a souhlasu. V článcích pro The Verge o tom psali Statt (2018) a Singleton (2018) s odkazem na zjištění tamní zpravodajské agentury Associated Press (AP). Asi nejvíce palčivými se ukázaly být záznamy GPS polohy téměř každého mobilního telefonu, které se

⁸ IoT je zkratkou anglických slov *Internet of things*. V češtině se můžeme setkat s českým termínem *Internet věcí*. Jako IoT zařízení se typicky považují ta fyzická zařízení, která mají schopnost vyměňovat si data s jinými zařízeními a spoluvytvářet tak vzájemně komunikující síť.

⁹ Volně přeloženo z anglického originálu: „*You may wonder if Alexa is listening to conversations, or listening to everything you say. You may want to know whether Alexa listens to you when you're not directly interacting with an Echo device. The answer to those questions is no. Alexa and all of our Echo devices are designed with your privacy in mind.*“

bez omezení obchodovaly i s velmi málo důvěryhodnými subjekty s pochybnými zájmy. AP v článku Bajaka (2018) také informovala o tom, že data odkoupená třetími stranami od mobilních operátorů se využívala například ke sledování pohybu zásilek, zaměstnanců, prevenci bankovních podvodů, ale i k cílené reklamě a sledování konkrétních osob.

Jak vyplynulo z předchozího příkladu, byznys model založený na sběru a monetizaci uživatelských dat vytvořil prostor pro firmy, které se zabývají výhradně přeprodáváním dat. Nemohou se k nim dostat jinak, nežli je nakoupit nebo získat z veřejně dostupných zdrojů, případně je třídít, doplňovat, analyzovat a nabízet dál, ať už dalšímu prostředníkovi nebo firmě, která data dokáže využít, například pro svůj marketing. Podle autora odborných publikací na téma sběru osobních dat profesora Matthew Craina se těmto sběračům a prodejcům dat říká data brokers, tedy datoví makléři (Crain, 2016). Běžně se ale také používá termín datoví agregátoři (Loshin, 2013). V roce 2022 jich v seznamu vedeném organizací Privacy Rights Clearinghouse jen v USA bylo zhruba 130. Americká společnost WebFX (2020), která se zabývá digitálním marketingem, odhadla celosvětový počet datových agregátorů dokonce na 4000. Ve stejném článku odhaduje roční světové tržby z obchodování s daty na 200 miliard amerických dolarů. Stejně číslo uvedl i Crain (2016). Z toho řádově miliardy amerických dolarů utržil každý ze tří v té době největších datových makléřů – Experian, Equifax a Transunion. Ventakadri spolu se svým týmem (2018) označil za data brokery, kteří agregují a zpeněžují data, také například firmy Facebook a Google. Podle Elsaify a Hasana (2020) z přibývajících důkazů vyplývá, že obecně firmy z výměny dat mezi sebou těží. Nutno dodat, že na trh s daty vstupují i nekomerční subjekty – poptávka o osobní data je například ze strany státních institucí (Cyphers, 2021).

Crain (2021) uvádí, že velmi brzy poté, co se z dat stal lukrativní materiál, se společnosti rozhodly sbírat nejen data, pro které měly využití, ale i ta data, se kterými zrovna nevědí, jak v daný okamžik naložit. Vědí totiž, že jejich využití se může teprve objevit až v budoucnu. To, že data sbírá, ale zatím je třetí stranám neprodává, uvedla například společnost iRobot, výrobce autonomního vysavače Roomba, u kterého se prokázalo, jak ukazuje například článek The New York Times, že sbírá data o půdorysu místností a předmětech, které se v nich nacházejí (Astor, 2017). Google v prvopočátku a pravděpodobně často ani dnes nemusí mít úplně jasno, jak některá agregovaná data využije – přesto tak dělá s tím, že mu to jeho technologie umožňuje a že se nasbírané údaje mohou v budoucnu hodit. Fourcade a Healy (2017) tento princip, kdy firmy sbírají data, protože je přinejmenším škoda tak nečinit, nazývají pojmem „data imperative“. Podobně na problematiku nahlíží i novinářky Angwin a Valentino-Devries (2012), které v článku The

Wall Street Journal píše, že uchovávání a analyzování byť i zdánlivě mondénních lidských aktivit už není výjimkou, ale normálem.

Příkladů, jak se firmy inspirují a experimentují s byznys modelem, který nejznámějším technologickým společností přinesl nevídané bohatství a moc, existuje mnohem více. Podle experta na kyberbezpečnost Bruce Schneiera (2013) je způsobů sledování tolik, že ani nemá smysl se snažit sledování vyhýbat. Pár příkladů uvedených výše má pouze naznačit základní způsoby, jak se s uživatelskými daty obchoduje, a to je jejich sběr a využití in-house, jejich sběr a prodej třetím stranám, pouhý pře prodej, a nakonec jejich hromadění pro jejich budoucí využití. Nutno dodat, že k boomu sbírání dat ruku v ruce přispěl i vývoj cen technologií k tomu nutných, zejména úložišť a jejich správy. Jejich cena začala být v jeden okamžik natolik příznivá, že ekonomicky začalo i třeba bezdůvodné sbírání dat dávat oproti jejich nesbírání nebo mazání větší smysl (Sadowski, 2019).

Uživatelská data se stala obživou či přinejmenším velmi lukrativní součástí byznysu celé řady společností, které k nim přistupují jako k majetku, či kapitálu, popřípadě surovině (The Economist, 2017). Nakolik je tato surovina cenná, a jak se její cena vyvíjela, popíše v následující kapitole diplomové práce.

2. Komodifikace dat

Podle Zuboff (2019) industriální kapitalismus přeměňoval přírodní zdroje v komodity a dnešní kapitalismus dohledu si nárokuje člověka jako zdroj pro tvorbu nové komodity. Naše životy jsou tak podle autorky zpracovávány k získávání behaviorálních dat určených k zpřesňování marketingových produktů a modelů pro předpověď lidského chování. Zuboff shrnuje, že uživatelé neznamení pro technologické společnosti nic víc než „human natural resources“, tedy volně přeloženo, „lidské přírodní zdroje“ nebo „lidské přírodní bohatství“.

Podle Sadowského (2019), který se na problematiku dívá z hlediska digitálního kapitalismu, je analogie s lidmi jakožto zdroji přírodního bohatství přespříliš zjednodušující. Tvrdí, že nelze uživatelská data chápat vždy jako komoditu, ale spíše jako kapitál. Sadowski argumentuje například tím, že data stále cirkulují a jsou cíleně vytvářena a sbírána a tento pohyb je podle něj stejný jako akumulace kapitálu v souladu s tím, jako to popsali K. Marx ve svém díle *Kapitál*. Sadowski netvrdí, že by data a peníze byly zaměnitelné, nicméně je podle něj lze postavit na stejnou úroveň. Právě imperativ akumulace dat a vytváření nových kanálů pro jejich sběr, podobně jako to je s kapitálem, Sadowskému připadá jako důkaz, že data jsou spíše kapitálem než-li komoditou. Van Dijck (2014) data, která o sobě uživatelé internetu prozrazují, přirovnává k „*platiidlu*“¹⁰, které je normální firmám poskytnout jako odměnu za možnost využívat jejich online služby.

Data, na rozdíl od kapitálu, mohou mít různou kvalitu, která ovlivňuje jejich hodnotu. Zásadním faktorem, který hraje v ceně osobních dat roli, je jejich stáří (Malgieri & Custers, 2018). Osobní data musejí být co nejaktuálnější, aby se s nimi lidé dali co nejlépe targetovat, tedy aby se na ně dalo co nejlépe a nejefektivněji cílit. Důležité je to hlavně u mladých, jejichž preference a chování se velmi rychle mění. Podle Malgieriho a Custerse jsou tedy data dynamickou komoditou. Tvrdí také, že osobní data začínají mít hodnotu až v okamžiku, kdy začínají tvořit celistvější obraz o dané osobě.

V dnešní době totiž nejde jen o základní údaje, jako je věk nebo pohlaví. Vzhledem k tomu, že data bývají zpravidla využívána k cílení personalizované reklamy, pro taková specifická sdělení je o dané osobě (nebo v praxi spíše o skupině) nutné znát mnohem více informací. Jeden z největších agregátorů dat, společnost Acxiom, například běžně sbírá demografické, ekonomické, behaviorální, zdravotní, náboženské, sexuální a zkušenostní údaje, které rozšiřuje i o informace z offline světa (Crain, 2016). Podle Craina měla Acxiom

¹⁰ Volně přeloženo z anglického termínu „*currency*“.

v roce 2014 přes 3000 dat o skoro každém dospělém Američanovi. Podle již zmíněného WebFX (2020) mají databáze CampaignGRID a ProPublica informace o politickém smýšlení a preferencích až o 80 % všech registrovaných amerických voličů.

V minulosti se prokázalo, že právě takové specifické informace o jednotlivcích, jako je například politické smýšlení, jsou pro poskytovatele online inzertního prostoru velmi cenné. Publicista Daniel Dočekal v pořadu Českého rozhlasu Dvojka Káva o čtvrté moderátorky Zuzany Vlčkové (2022) připomněl, že například Facebook poptával záznamy o zdravotním stavu lidí, neboť šlo o data, na základě kterých mohl na své uživatele velmi přesně cílit reklamu z oblasti velmi lukrativního byznysu.¹¹

Například novinářka Madhumita Murgia v roce 2017 v prezentaci pro TEDx uvedla, že makléři dat ji znají lépe než její sousedi. Vědí o ní například, že pracuje v médiích, kde žijí její rodiče, kde nakupuje, co čte, jaké jsou její zájmy, jaké má ráda filmy, kolikrát do roka cestuje a podle čeho si vybírá destinace, včetně toho s jakým předstihem kupuje letenky. Vědí také, kolik Murgia vydělává, jak často si nechává dovážet jídlo a jaké jídlo to je. Nechybí podle ní ani informace o tom, jak a s kým tráví víkendy, do které restaurace chodí po práci, nebo také že k ní do bytu chodí uklízečka. Kromě těchto faktů mají agregátoři dat také informace o jejím charakteru: například že spíše než sama pracuje raději v týmu, že je spíše optimista atd.. Jediné, co prý v záznamech data agregátorů není, je její jméno, protože datové balíčky musejí být ze zákona anonymizované.

To, že informace o sobě uživatelé nemusejí ani výslovně prozrazovat, ale technologické společnosti si dokáží mezery v datech samy doplnit, míní i Zuboff (2019). Jak naznačila Murgia ve své prezentaci na TEDx (2017), společnosti o svých uživateli neshbírají pouze fakta, jako jsou věk, etnicita a demografie, ale na bázi jejich textů a chování na síti analyzují jejich psychologii, a zařazují je tak do psychologických modelů, díky nimž dokáží doplňovat mezery v datech a hlavně chování a emoce předpovědět. To potvrzují ve svém experimentu Kosinski, Stillwell a Graepel (2013), kteří dokázali na základě lajků z Facebooku, demografie a psychometrických testů celkem přesně odhadnout intimní osobní informace, jako je sexuální orientace, etnicita, náboženská víra, politické názory, inteligence, nálada, užívání návykových látek, úplnost rodiny, věk nebo pohlaví.

¹¹ Daniel Dočekal v rozhlasovém pořadu řekl: „Facebook pravidelně dlouhá léta, ne moc v Česku, ale ve Spojených státech, nakupuje o uživateli i informace o tom, jak jsou movití, co mají za nemovitosti, v jakých jsou příjmových kategoriích a třeba jeden čas i zdravotní informace. Byl poměrně hlad po tom se pustit do zjišťování toho, jak jsou na tom lidé z hlediska zdraví, protože tam je obrovský byznys ohledně cílení další reklamy.“

Například Facebook dokáže podle Zuboff (2014) jít ještě dál. Předpovídá emoce svých uživatelů a umí podle ní dokonce regulovat svůj obsah tak, aby v uživateli kýženou emoci vyvolal, a zvýšil tak úspěšnost určitých reklam. Na svém webu odkázala autorka na výzkum Kramera, Guillorye a Hancocka z roku 2014, kteří na síti Facebook na vzorku skoro 700 000 uživatelů otestovali fenomén tzv. emocionální nakažlivosti¹². Dokázali, že emoce a chování uživatelů dokáže významně ovlivnit tak jednoduchá věc, jako je výběr postů ukázaných ve vybraných příspěvcích. Jejich kontroverzní experiment probíhal tak, že aniž by to uživatelé Facebooku tušili (stačilo, že každý uživatel dal svolení s experimenty při první registraci do sítě, svolení k experimentům totiž je jednou z podmínek užívání Facebooku), algoritmy jim začaly v newsfeedu servírovat buď pouze negativní informace, nebo naopak výhradně pozitivní sdělení. Výzkumníci Facebooku ve výsledné zprávě uvedli, že uživatelé byli ovlivňováni emocemi, kterým byli vystavováni, což mělo dopad i na jejich vlastní aktivitu. Ta buď rostla nebo klesala, přičemž tyto změny chování a emocí vykazovaly schopnost promítat se na další uživatele sociální sítě.

I proto lidé sociální sítě, jako je Facebook, přirovnávají ke gigantickým sociálním experimentům. Právě Facebook, nyní již Meta, má celé oddělení věnované výzkumu zvané Meta Research, které se specializuje na provádění studií a experimentů.¹³ Na jeho stránkách se lze například dočíst o studii zvané *Tvorba lásky*¹⁴, ve které výzkumníci Facebooku prezentují, jak se sociální síť na základě analýzy dat svých uživatelů naučila předpovídat, kdy si spolu dva konkrétní uživatelé vytvoří intimní vztah odehrávající se i mimo online prostředí (Facebook, 2014). Zpravodajský server The Guardian, který vycházel z informací deníku Australian (2017), na základě uniklých interních informací Facebooku informoval, že sociální síť dokáže pro inzerenty cílit na teenagery, kteří se zrovna cítí ve stresu, nejistotě nebo kteří se cítí hloupě, méněcenně či poraženě (Levin, 2017).

2.1 Kategorizace osobních dat

Podle toho, jak se osobní, nebo uživatelská data sbírají, se údaje rozlišují na „identifikovatelné“ (identifiable), „anonymní“ (anonymous) a „pseudonymní“ (pseudonymous). Identifikovatelná jsou ta data, která o sobě dotyčný člověk dobrovolně a vědomě odevzdal. Anonymní jsou informace sesbírané a zpracované „zpracovateli dat“, a to často bez souhlasu a vědomí dotyčných osob, které si mohou myslet, že poskytují jen

¹² Volně přeloženo z anglického originálu „*emotional contagion*“.

¹³ Stránky výzkumného oddělení včetně studií najdeme na adrese: <https://research.facebook.com>.

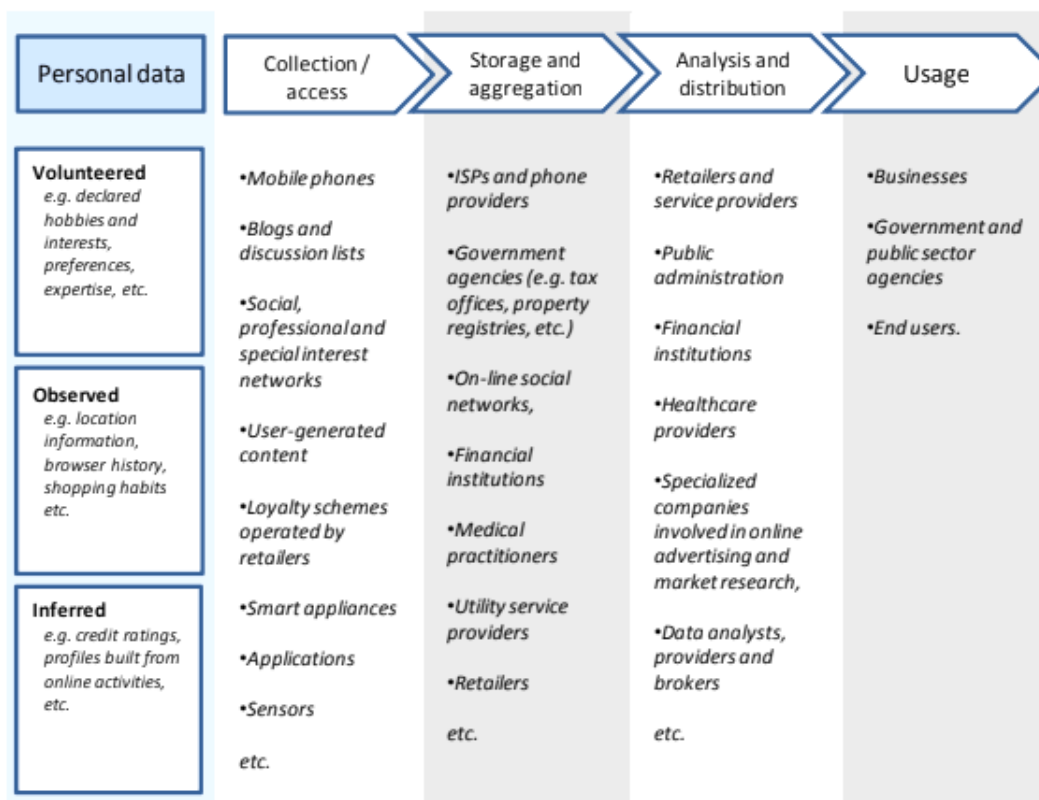
¹⁴ Volně přeloženo z názvu v angličtině *The formation of love*

anonymní údaje. Poslední kategorie, pseudonymní osobní data, se získávají od třetích stran bez účasti samotných osob (Birch, Cochrane & Ward, 2021).

OECD ve své zprávě z roku 2013 píše, že rozlišování mezi osobními a neosobními daty se stává stále obtížnější. Cituje Narayanana a Shmatikova, kteří v roce 2010 uvedli, že: „V okamžiku, kdy se jakýkoliv záznam přiřadí skutečné identitě osoby, jakákoliv spojitost mezi touto informací a virtuální identitou zabraňuje, aby osoba zůstala v anonymitě.“¹⁵ (str. 8)

OECD ve zmíněné zprávě také dodává, že techniky využívané v době vzniku studie, jako jsou fráze zadané do vyhledávačů, navštívené webové stránky, GPS poloha a IP adresa, běžně dovolují spojovat data s reálnou identitou osob. Samotná data pak rozdělují do následujících kategorií: dobrovolně prozrazené údaje, vysledované údaje a vyvozená data. Rozdělení popisuje následující graf.

Figure 1. Personal data value chain



Obrázek 1: Rozdělení dat podle zprávy OECD (2013) na základě dat z WEF (2011).

¹⁵ Volně přeloženo z anglického originálu: „Once any piece of data has been linked to a person’s real identity, any association between this data and a virtual identity breaks the anonymity of the latter.“ (Narayanan & Shmatikov, 2010, str. 8).

Druhý způsob, kterým lze data klasifikovat, vychází ze samotné informace, kterou záznam nese. Ed Busby a kolektiv (2012) z Boston Consulting Group, která se zabývá digitálním marketingem, uvedli pět kategorií dat. Zároveň ale připomínají, že žádný z těchto typů dat sám o sobě nezaručuje bezchybné targetování. Pro lepší efektivitu údajně většina marketérů kombinuje vícero z nich.

- *Demografická data: Obsahují informaci, jako je věk, pohlaví, příjem. Tato data mohou být buď dobrovolně prozrazená, nebo vyvozená, a tvoří jádro třídění spotřebitelů nebo zákazníků.*
- *Behaviorální či kontextuální data: Vypovídají o zájmech uživatelů nebo náladách. Mohou být získána dobrovolně nebo vyvozená na základě konzumovaného obsahu či pomocí cookies. Propojování behaviorálních dat s nákupním chováním je obtížné a je často třeba propojit řadu dílčích informací pro efektivní zásah reklamou.*
- *Nákupní preference: Tato data vypovídají o plánech uživatele uskutečnit nákup určitého produktu či artiklu. Mohou být dobrovolně získána (například pokud vyplní kontaktní formulář potvrzující zájem dozvědět se více o daném produktu), anebo tato vypořazovaná na bázi vyhledávání. Mohou být také vyvozená na základě uskutečněných nákupů. „Retargeters“ — tedy společnosti, které sledují, na co lidé koukali, ale nekoupili - hlásily úspěchy s tím, že díky této informaci daný nezakoupený artikl uživatelům později inzerovali.*
- *Sociální data: Popisují vztah uživatele s dalšími lidmi. Z pohledu marketingu, sociální nebo společenská data předpokládají, že lidé, kteří jsou navzájem propojeni, budou také sdílet podobné atributy a nákupní chování. Tyto informace může být získána dobrovolně například skrz sociální sítě nebo na základě vzájemné interakce, například online zasláním novinového článku známému.*
- *Polohová data: Marketéři jsou schopni určit polohu uživatelů pomocí několika postupů. Údaje o poloze byly tradičně získávané pomocí IP adresy uživatele nebo na základě internetových stránek, které si prohlíží (podle historie prohlížení lze přibližně odvodit, kde uživatel žije). Internet v mobilu zlepšuje přesnost zaměření polohy a také obchodníkům umožňuje cílit reklamu v reálném čase a místě.*

2.2 Hodnota dat pro firmy

Lane a kolektiv (2014) píše, že neexistuje jeden, ale hned několik trhů, kde se poptávka po datech setkává s nabídkou. Jedním z nich je prodávání a nakupování ve velkém od tzv. agregátorů, kde uživatelé v transakci už vůbec nefigurují.

Již zmíněná digitálně-marketingová společnost WebFX (2020) uvádí, že jedna z firem obchodující z daty prodávala balíček dat 1000 lidí, jež spojovaly zdravotní problémy typu anorexie, zneužívání léků a deprese, za 79 amerických dolarů. Pokud bychom částku vydělili počtem lidí na seznamu, vyjde nám, že data jednoho člověka „stála“ pouze necelých 0,08 amerických dolarů. Olejník s kolegy v roce 2013 uvedli, že informace z historie internetového prohlížeče jedince se inzerentům prodávají za částku 0,0005 amerických dolarů a méně. Podobně na problematiku nahlíží i autoři článku Financial Times (Steel, 2013): čtveřice novinářů uvedla, že profil jedince obsahující základní informace, jako je věk, bydliště a pohlaví, měl hodnotu přibližně 0,0005 amerických dolarů. Až v okamžiku, kdy se tyto údaje rozšíří o další konkrétnější informace, jako je rodinný stav či specifická lékařská diagnóza, tak cena údajně roste v řádech desítek centů. Podle zprávy OECD z roku 2013 měla informace o bydlišti cenu půl dolaru, datum narození dva americké dolary, číslo řidičského průkazu tři americké dolary a číslo zdravotního pojištění osm amerických dolarů. Dle OECD tak konkrétně společnost Facebook za rok za data průměrného uživatele získala v rozmezí od čtyř do sedmi amerických dolarů.

Data tedy mají nejen jednorázovou cenu při jejich přeprodeji mezi subjekty, ale mají také dlouhodobější hodnotu. Právě jak vyčíslila OECD (2013), stejná data se dají zpeněžovat delší dobu. Z ekonomického hlediska se tedy uživatelská data dají nazvat aktivem (Birch, Cochrane & Ward, 2021). Jenkins z týmu cílení reklamy společnosti Experian, tedy jednoho z největších agregátorů dat, v roce 2012 pro server The Drum napsal článek, ve kterém vyčíslil, že dlouhodobý výdělek z jedné e-mailové adresy využívané k zasílání obchodních sdělení je průměrně 84,5 britské libry, přičemž konkrétní cena může být vyšší nebo nižší v závislosti na zájmech majitele dané e-mailové adresy. Například e-mailová adresa člověka se specifickými a ideálně nákladnými zájmy za dobu své existence vydělá mnohem více. Potvrzuje to i Steel s kolegy (2013), kteří zase poukazují na skutečnost, že člověk, který si chce koupit například auto, finanční produkt či dovolenou, má z hlediska směny dat větší hodnotu. Jenkins (2012) v souvislosti s tím uvádí, že například e-mail fanouška cestování může v dlouhodobém horizontu vydělat až 236 liber. Naopak e-mail, který patří člověku s méně specificky zaměřenými zájmy, vydělá méně, třeba 79 liber. Takových adres je podle

autora mnohem více než těch „drahých“ a chodí na ně také mnohem více reklamních sdělení, avšak s menší vidinou úspěchu. Proto se průměrná částka přibližuje spíše k té nižší hranici než té vyšší.

Zajímavý postup na vyčíslení hodnoty, kterou data uživatelů největších internetových služeb generují, zvolil šéf společnosti Casebook PBC Tristan Louis ve svém komentáři pro Forbes z roku 2013: Vzhledem k tomu, že naprostá většina výtěžku společností, jako je Facebook nebo Google, pochází z prodeje reklamního prostoru, vzal si jejich výroční zprávy s výtěžky, vydělil je počtem uživatelů a číslo následně ještě poměřil s valuací firmy. Tímto výpočtem mu vyšlo, že jeden uživatel čtvrtletně firmám vydělal mezi 1,72 až 3,76 americkými dolary.

Podobný způsob využívají i samotné technologické společnosti. Například Facebook ve své výroční zprávě z roku 2020, ve které prezentuje své finanční výsledky, přepočítal své tržby z reklamy na počet uživatelů. V průměru podle této zprávy jeden uživatel Facebooku v roce 2020 vydělal něco málo přes 31 amerických dolarů. Průměrný evropský uživatel vydělal více, necelých 50 amerických dolarů. Méně peněz Facebooku plynulo z regionů s nižší kupní silou (Facebook, 2020). Podobně jako na výše uvedeném příkladu o hodnotě e-mailových adres pro marketingové účely se tak ukazuje, že cena dat uživatelů se mění v závislosti na tom, kde žijí - nebo ještě přesněji - v závislosti na tom, jakou mají kupní sílu. Tucker a Lumbrecht (2019) navíc dokázaly, že pro inzerty je dražší cílení reklam na mladé ženy.

Kromě čtvrtletní, roční nebo jakékoliv jiné hodnoty měnící se v čase, Glikman a Glady (2015) v článku pro server TechCrunch míní, že hodnota jednoho uživatele se dá také odvodit z ceny, za kterou se internetové služby pořizují. Jejich úvaha spočívá v tom, že pokud Facebook koupil WhatsApp s jeho 600 miliony uživateli za 19 miliard amerických dolarů, měl každý uživatel cenu 30 amerických dolarů. Podle těchto jednoduchých počtů měl 30 amerických dolarů cenu i každý uživatel Instagramu v momentě jeho prodeje v roce 2012. Takto podle autorů najdeme cenu uživatele mezi 15 až 40 americkými dolary v závislosti na zaměření platformy a typu dat, které od uživatelů sbírá. Pokud touto logikou vydělíme sumu 44 miliard amerických dolarů, za kterou se miliardář Elon Musk na jaře 2022 dohodl odkoupit sociální síť Twitter, necelými 300 000 uživateli, kterou tato síť v době dohody měla (Kreč, 2022), dostaneme částku okolo 190 amerických dolarů za jednoho uživatele. Nutno dodat, že Musk v létě 2022 oznámil odstoupení od rozjednaného odkupu Twitteru s argumentem, že není jasné, kolik účtů je na této sociální falešných (X Holdings, SEC, 2022).

Právě dělení hodnoty firem, které si na monetizaci dat zakládají, počtem uživatelů je metodou, která se objevuje v již zmíněné zprávě OECD (2013): V roce 2012, kdy se akcie Facebooku začínaly volně obchodovat na burze, se hodnota firmy vydělená počtem uživatelů rovnala 111 americkým dolarům. OECD poznamenává, že nejde o naprosto přesnou metodu, protože výsledek zkreslují různé investice a další kapitál, který finanční bilanci firmy ovlivňuje. Nicméně je podle autorů zprávy důležité si uvědomit, že žádná z metod na vyčíslení ceny osobních dat nedokáže opravdovou hodnotu přesně vystihnout a určit.

Základní přehled o tom, kolik data stojí, lze získat také při procházení darknetu. Například databáze e-mailů jsou podle projektu Dark Web Index serveru Privacy Affairs velmi běžně obchodovaným artiklem, v zásadě i velmi levným, protože na aktuálnost adres a jejich kvalitu se dá jen těžko spolehnout. V posledním Indexu sestaveném Ignoffem a Zoltanem (2022) je například uvedeno, že balík 600 tisíc novozélandských e-mailových adres stojí pouhých 10 amerických dolarů. Deset amerických dolarů stojí také objemný balíček 4,78 milionů mexických e-mailových adres. Podle údajů stejného Indexu vycházejí v přepočtu na jeden údaj nejdražší české e-mailové adresy, kterých bylo možné za deset amerických dolarů pořídit 350 tisíc. O poznání dražší jsou podle metriky seznamy e-mailových adres se specifitějším zaměřením. Více než 120 tisíc e-mailových adres soukromých amerických zubařů už na darknetu dohromady stálo 50 amerických dolarů.

Osobní data nebývají na černém trhu nikterak lukrativním zbožím. Profesor A&M University v Texasu Ravi Sen v článku pro The Next Web z roku 2021 napsal, že nabídka osobních dat, díky kterým lze konkrétní lidi identifikovat, na černém trhu dlouhodobě převyšuje poptávku, a cena tak v důsledku toho klesá. Sen uvádí, že zatímco v roce 2014 takovéto informace o jedné osobě mohly vyjít na čtyři americké dolary, pouhý rok poté už jen na jeden americký dolar. Oproti tomu podle Sena na černém trhu získávají oblibu zdravotní záznamy lidí. Zdravotní společnosti pak za uniklá data vyplácejí podle autora vysoké sumy. (Sen, 2021)

Právě částky, které společnosti musely firmy vynaložit na odškodnění klientů, jejichž data z nějakého důvodu unikla na veřejnost, jsou dalším pohledem, jak data nacenit. David Slížek, šéfredaktor serveru Lupa.cz, který se věnuje dění na poli technologií a internetu, v roce 2019 referoval o verdiktu v odvolacím řízení Městského soudu v Praze, který firmě Internet Mall a.s. nařídil zaplatit žalobci odškodné za data, která firmě unikla spolu s údaji dalších stovek tisíc zákazníků. Podle Slízka šlo totiž o unikátní rozsudek, který v Česku dosud neměl obdoby. A to proto, že: „*Lidem, kterým z nějaké online služby unikly osobní údaje, dává naději, že mohou podobným způsobem získat u soudu odškodnění i oni.*“ Výše

odškodného za uniklé jméno, příjmení, telefonní číslo, e-mailovou adresu a heslo k účtu, činila 10 tisíc korun plus úroky.

Na další příklad upozorňují Glikman a Glady (2015). Uvádějí, že americký nadnárodní telekomunikační a mediální konglomerát Comcast musel zaplatit odškodnění 100 amerických dolarů více než 75 000 zákazníkům za únik jejich dat. Situace byla ještě o to zajímavější, neboť šlo o klienty, kteří zaplatili Comcastu příplatek za to, aby jejich data zůstala soukromá.

Na internetu lze najít zahraniční služby, které se za provizi zabývají vymáháním odškodného za uniklá data. Avizují, že odškodné, na které mohou mít poškození nárok, se může pohybovat v řádech tisíců eur či liber.¹⁶

Další přístup, jak zjistit hodnotu a cenu dat, je skrze služby, které finančně odměňují své klienty, když o sobě dobrovolně prozrazují osobní informace. Tento přístup se stal obzvláště atraktivní v pojišťovnictví, kde mají osobní informace, jako jsou například zdravotní záznamy, přímou souvislost s rizikem, a tedy i cenou životní pojistky (Malgieri & Custers, 2018). Již zmíněná zpráva OECD z roku 2013 také připomíná, že například data broker Experian nabízí Američanům pojištění dat proti krádeži identity, a to za 155 amerických dolarů ročně. OECD to uvádí spíše jako kuriózní přístup k oceňování osobních údajů, zejména pokud podle OECD vezmeme v potaz, že Experian data jedné osoby dokáže ročně zpeněžit za zhruba šest a půl amerických dolarů. Ochota utrácet za ochranu dat nicméně spíše nastiňuje, jakou hodnotu datům přiřazují samotní uživatelé. Tomuto tématu je věnována následující kapitola.

2.3 Hodnota dat pro uživatele

Jakou mají data cenu z pohledu samotných uživatelů internetových služeb? To zkoumal například absolvent oboru Studia nových médií na Filozofické fakultě Univerzity Karlovy Ondřej Trhoň s Barborou Hyclovou a Danielem Šlahařem (2017), kteří se tématu věnovali v rámci výstavy Big Bang Data v pražské galerii DOX. Výzkumníci nechali lidi nejprve odhadnout, jakou mají jejich data, která o nich Facebook nasbíral, cenu, a následně jim na základě několika již existujících „kalkulaček cen dat“¹⁷ vyčíslil, na kolik si jejich data Facebook reálně cení. Zpravidla jen málokdo ze zúčastněných čekal, že tržní hodnota jejich dat je tak nízká, a to oproti původním odhadům mnohonásobně.

¹⁶ Jde například o firmy DataClaim (<https://dataclaim.co/en-eu>) nebo Data Breach (<https://data-breach.com/>).

¹⁷ Mezi kalkulačkami cen dat byly například výpočty Financial Times a Totally Money.

K obdobným závěrům došel i průzkum české výzkumné agentury STEM/MARK zveřejněný na webu Seznam zprávy z roku 2021, který se tisícovky respondentů ptal na to, kolik peněz vygenerují na základě jejich dat například společnosti Facebook, Google či Twitter. Podle autorů studie jde řádově o stovky až pár tisíc korun ročně, což správně odhadla pouze zhruba pětina dotazovaných. (Jindrová, 2021)

Výzkum pod vedením A. Collis z roku 2022 mezi účastníky zjišťoval, kolik podle nich stojí jejich data ze sociálních sítí. Výzkumníci následně zjistili, že odhadní cena se napříč výzkumnými vzorky nejen velmi lišila (od jednotek po desetitisíce amerických dolarů), ale do odhadů respondenti promítali své socio-ekonomické zázemí. Ti s nižšími příjmy přiřkli svým datům také nižší cenu. Stejně tak levnější cenovku svým datům dávali také lidé tmavé pleti – medián pro respondenty bílé pleti činí 1000 amerických dolarů, pro tmavé pleti polovinu.

Malgieri a Custers (2018) upozorňují na to, že pokud lidé přisuzují svým osobním datům hodnotu, mají větší tendenci si je chránit a bránit jejich volnému pohybu. Že se tak snaží aktivně činit tvrdí Acquisti a kolektiv (2020). Přesto podle Johnsona (2020) lidé sice hlásají, že jim na online soukromí sejde, doopravdy jim ale na něm spíše nezáleží. Ovšem podle Craina (2016) to není tak jednoduché. Dle jeho slov je člověk v dnešní západní společnosti postaven před těžko řešitelné rozhodnutí: a to smířit se se sdílením svých dat nebo upustit od života v moderní společnosti, která se dnes bez digitálních technologií neobejde.

Nedávná studie pod vedením Kluiterse z pražské Vysoké školy ekonomické v Praze (Kluiters a kol., 2022) dokázala, že spotřebitelé při nákupu či hledání placených služeb přinejmenším zvažují, jak se společnosti k jejich datům chovají, a to pomocí jednoduché logiky: při porovnání výkonnosti a nefinančních zpráv více než 2800 amerických společností zjistili, že ty, které se snaží být zodpovědnými správci osobních dat, jsou v dlouhodobějším horizontu také výkonnějšími.

Jednou z metod, kterou vědci využívají k určení ceny osobních dat z pohledu uživatelů, je nechat je zaplatit za nesbírání nebo nezveřejňování jejich osobních údajů. Například Tsai a kolektiv (2011) pomocí experimentu vyčíslili, že pakliže je lidem jasně deklarováno, co se s jejich daty může stát, při nákupu různých druhů zboží v hodnotě okolo 15 amerických dolarů lidé preferují e-shopy, které jejich soukromí chrání víc, a to i za cenu, že prodávají zboží zhruba o americký dolar draž. Savage a Waldman (2013) přišli se závěrem, že lidé jsou ochotni zaplatit jednorázově 2,28 amerických dolarů za skrytí své historie v internetovém prohlížeči, 4,05 amerických dolarů za uchování kontaktů, 1,19

amerických dolarů za nesdílení polohy, 1,75 amerických dolarů za skrytí identifikačního čísla svého telefonu a 3,58 amerických dolarů za skrytí obsahu svých textových zpráv. Jiný experiment zjistil, že polovina účastníků nebyla ochotná vyměnit anonymizovanou dárkovou kartu v hodnotě 10 amerických dolarů za hodnotnější s 12 americkými dolary, která by sledovala jejich nákupní chování (Acquisti a kolektiv., 2013). V novějším online experimentu výzkumníci účastníkům nabídli vyšší peněžní bonus za účast ve výzkumu, pokud by po vyplnění dotazníku s výzkumníky ještě sdíleli svůj Facebookový profil. Za dva a půl amerického dolaru to většina odmítla (Svirsky, 2019).

Již zmíněný průzkum agentury STEM/MARK (Jindrová, 2021) poukázal také na to, že velká část české společnosti nerada data s technologickými firmami sdílí. Například 63 % Čechů údajně odmítá se společnostmi sdílet svou historii prohlížení internetu. Stejný počet respondentů (63 %) také uvedl, že odmítá s technologickými společnostmi, jako je Facebook, Google či Twitter sdílet obsah svých textových zpráv nebo e-mailů. Je samozřejmě otázkou, jaká část těchto respondentů je ochotná podniknout proti sbírání těchto dat nějaké kroky (například využíváním alternativních služeb, které si na nesbírání uživatelských dat zakládají, nebo nejrůznějšími sofistikovanými plug-iny do prohlížečů), popřípadě jaká část z nich se používání služeb vyhýbá. Vzhledem k tomu, že server MediaGuru (2017) informoval o tom, že v Česku mělo aplikaci Facebook Messenger v telefonu staženou 64 % majitelů chytrého telefonu, značná část z respondentů, kteří pro STEM/MARK uvedli, že odmítají s technologickými společnostmi sdílet obsah svých textových zpráv, si zřejmě neuvědomuje, že tak činí bez svého vědomí. Byť například Facebook uvádí, že k obsahu zpráv zasílaných pomocí aplikací Messenger nebo Whatsapp nemá kvůli end-to-end kódování přístup, jak informoval v článku pro server ProPublica Elkind, Gillum a Silverman (2021), Facebook ke zprávám zaslaným skrz jeho aplikace přístup má a s jejich obsahem pracuje.

To, že lidé nakonec stejně raději obětují něco ze svého soukromí, aby ušetřili, dokazuje výzkum Preibusche, Kubler a Beresforda (2013). V něm vyšlo najevo, že pouze třetina dotazovaných si byla ochotná při zakupování filmu na DVD připlatit jedno euro za to, aby o nich obchodník sbíral méně dat. K podobnému závěru došel výzkum European Network Information and Security Agency (ENISA) z roku 2012 (Jentzsch a kol.). Vyšlo v něm najevo, že méně než třetina jeho účastníků je ochotna si připlatit za příslib toho, že by o nich provozovatel služby nesbíral data pro marketingové účely. Na neochotu platit za ochranu dat poukázal také experiment Beresforda a jeho kolegů (2012).

Výzkum Brynjolfssona, Collise a Eggerse (2019) se zase respondentů ptal, jakou hodnotu pro ně hojně využívané a zdarma poskytované digitální pomůcky a služby mají,

respektive jaká finanční kompenzace by jim přišla adekvátní, kdyby je nemohli používat. Výzkumníci zjistili, že median finanční kompenzace, za kterou by byli uživatelé Facebooku ochotní se této sociální síti na měsíc vzdát, činí 48 amerických dolarů. Byť nejde o hodnotu, kterou respondenti přiřazovali datům, ale tomu, nakolik ji ve svých životech potřebují, trojice vědců zjistila, že se tato hodnota v čase mění s tím, jak se mění oblíbenost a užitek jednotlivých služeb ve společnosti.

Podle Acquistiho představuje vyčíslení hodnoty osobních dat z pohledu lidí problém. Do výsledků podle něj zasahují behaviorální a kognitivní heuristiky, na což poukázal společně s Johnem a Loewensteinem (2013). Cena, kterou dotazovaní svým datům dávali, se údajně lišila na základě rámování až pětinasobně. Tento výzkum kromě jiných zmiňuje i již zmíněná studie OECD (2013), která dosavadní výzkum ceny dat z pohledu uživatelů označila jako „extrémně závislý na kontextu a nepřesný“. Jako příklady studií, které se zabývají hodnotou osobních dat a soukromí, OECD uvádí Spiekermanna a kol. (2002), Chellappa a Sin (2005), Wathieu a Friedman (2005), Huberman a kol. (2006), Cvrceka a kol. (2006), Hui a kol. (2007.) a Acquistiho a kol. (2009). Obecně se ale podle OECD dá tvrdit, že:

„...poměr uživatelů, kteří odmítají nabídku peněžní kompenzace za snížené soukromí, je větší než poměr těch, kteří jsou za větší míru soukromí ochotni zaplatit“. (str. 30)¹⁸

Na to, že osobní data mají pro uživatele hodnotu, kterou lze vyčíslit penězi, poukázal projekt Vetri (2022)¹⁹, který uživatele dokonce finančními odměnami motivuje k tomu, aby svá data vycházející z chování na internetu sdíleli s komerčními i nekomerčními subjekty. Dle informací ze svého webu se Vetri snaží být jakýmsi prostředníkem, který je schopný na tržišti agregovat nabídku a poptávku po konkrétních datech konkrétních skupin lidí. Ti se zároveň sami rozhodují, s kým a za jakým účelem chtějí data sdílet, přičemž Vetri jim za poskytnutí dat svěří odměnu, kterou je zpravidla konkrétní slevová hodnota na nákup u vybraných obchodníků. Vetri na svém webu uvádí následující:

„Věříme, že vaše data jsou jen a jen vaše a ne k užítku a zneužívání třetími stranami ... Věříme, že lidé by měli být za svá data férově odměněni. Cokoliv od

¹⁸ Volně přeloženo z anglického originálu: „Generally, the fraction of consumers who will reject an offer to obtain money in exchange for reduced privacy is larger than the fraction of consumers who will accept an economically equivalent offer to pay money in exchange for protection of privacy“. (OECD, 2013, str. 30)

¹⁹ Informace jsou dostupné na webové adrese <https://vetri.global/>.

behaviorálních dat po ty finanční či nákupní, případně názory, tato všechna data mohou generovat stabilní příjem pro své právoplatné vlastníky.“²⁰

²⁰ Volně přeloženo z anglického textu: „*We believe that your data is yours and yours only and not for online third parties to use and abuse ... We believe that individuals should be fairly remunerated for their data. Anything for behavioural data to financial and purchasing data and opinions ... can generate a steady revenue stream for their rightful owners.*“

3. Nejrozšířenější metody sbírání dat na internetu

Jak už bylo uvedeno v první kapitole diplomové práce, data o sobě můžeme zanechávat i při používání chytrého vysavače, televize či jiného zařízení s připojením na internet. Data jsou o nás sbírána i v řadě aktivit v offline světě. Výše popsany úspěch největších technologických společností se dostavil na základě sběru dat, která o sobě návštěvníci zanechávají při používání internetu a na něm postavených online službách. V této kapitole bude následovat přehled nejběžnějších metod, které se ke sběru dat využívají.

3.1 Geolokalizace a unikátní identifikátory

Jednou z prvních informací, kterou o sobě návštěvníci webu jejich provozovatelům zanechávají, je tzv. digital fingerprint, neboli digitální otisk. Jde o na první pohled nedůležité údaje, jako je rozlišení obrazovky, zařízení uživatele, ze kterého se připojuje, nebo typ a nastavení jeho prohlížeče, včetně nainstalovaných fontů. Electronic Frontier Foundation (2022) na svém webu vysvětluje, že byt' by tyto informace jednotlivě nemohly coby identifikátor uživatele obstát, v kombinaci vytvářejí natolik individuální mix, díky kterému lze relativně přesně odlišit jednoho návštěvníka od jiného. Zejména pak v kombinaci s IP adresou.²¹

Administrátoři webových stránek a služeb vědí, které IP adresy s tou jejich komunikují nebo komunikovaly. Ačkoli mezi IP adresou a geografickou polohou není přímá souvislost, Muir a van Oorschot (2006) vyjmenovávají celou řadu možností, jak lze díky IP adrese určit polohu zdroje síťové komunikace, a to s přesností až na metry, pokud se jedinec proti geolokaci pomocí IP adresy nebrání.

Zatímco geolokace na bázi IP adresy může na jednu stranu přinést řadu uživatelsky přívětivých funkcí, jako je automatické nastavení správného jazyka či zkratka personalizace obsahu na základě přibližné polohy, tou nejlukrativnější je cílení reklamy a analytika pro marketingové účely (Muir, van Oorschot, 2006). Řada firem se také zabývá sbíráním IP adres do databází, které následně monetizuje.²²

²¹ Každé zařízení, ať už počítač, telefon, nebo router, se v síti identifikuje IP adresou. Jde o identifikátor v podobě kódu, který zařízením s připojením na internet umožňuje spolu efektivně komunikovat. Slovo adresa není náhodné – jde o důležitý identifikační údaj, díky němuž je jednotlivým článkům sítě jasné, odkud příkazy pocházejí a kam odpovědi na ně posílat.

²² Muir a van Oorschot (2006) vyjmenovávají firmy Akamai, Digital Envoy, MaxMind, Quova, and Verifia.

Google (2022) ve své poradně na webu píše, že mobilní zařízení s operačním systémem Android nebo Apple dokáže pomocí GPS geolokalizovat s odchylkou zhruba 20 metrů. Uvádí také, že pro geolokalizaci využívá také signály z dostupných vysílačů Wi-Fi a také věží mobilního signálu. Dodává, že uživatelé zařízení s Android mají možnost volby mezi přibližnou a přesnou geolokací.

Jaký je mezi nimi rozdíl specifikoval redaktor webu Svět Androida Filip Houser. V roce 2021 uživatelům mobilních zařízení fungujících na operačním systému Android vysvětloval, co jim nadcházející nová verze operačního systému (Android 12) dovolí v souvislosti s polohovými službami nastavit a proč. Píše:

„V novém Androidu se objeví pojmy „přesná“ (precise) a „přibližná“ (approximate) poloha, přičemž volbu mezi nimi bude možné provést u každé aplikace. Jaký je mezi nimi rozdíl? Zatímco tu první určuje primárně GPS satelit a je přesná i na několik desítek centimetrů, přibližnou většinou určují GSM či WiFi sítě s přesností na desítky nebo stovky metrů. Je logické, že některým aplikacím bude k fungování stačit třeba jen ta druhá, čehož chtějí tvůrci Androida využít pro další ochranu soukromí uživatelů.“

Informace z GPS lokalizátoru nebo dostupných Wi-Fi hotspotů sice mohou zajišťovat řadu pro uživatele praktických funkcí, jakou je například velmi přesná navigace, využívají se ale také za účelem sběru dat o chování lidí a k marketingovým účelům (Taylor, Devlin & Curran, 2012).

Dalšími údaji, díky nimž lze uživatele do určité míry identifikovat a lokalizovat, je znalost umístění Wi-Fi routerů v reálném prostředí. Na praktiku, že technologické firmy mapují, kde se se které přístupové body s bezdrátovým připojením k internetu nacházejí, upozornil zejména skandál kolem projektu Street View společnosti Google. Ta v roce 2007 začala vysílat do ulic auta či lidi vybavené technologií pro skenování okolí. Online mapy Googlu tak dnes umožňují si ve virtuálním prostředí projít a prohlédnout ohromné množství reálných míst. Pár let od startu tohoto projektu však vyšlo najevo, že cílem nebylo pouze vylepšit mapy Googlu, ale také pomocí Wi-Fi přijímače putujícího reálným světem zmapovat fyzické umístění Wi-Fi přístupových bodů k internetu a sbírat data z nich pro marketingové účely, zejména pak na cílení reklamy (Burdon & McKillop, 2015). Nutno dodat, že Google nebyl první a ani jedinou společností, která tato data sbírala. V článku Kiss (2010) pro The Guardian například stojí, že společnost Google ve své obhajobě před veřejností, ale i soudy,

keré po odhalení této praktiky následovaly, argumentovala mimo jiné tím, že i před ní údaje z Wi-Fi routerů sbírala společnost Skyhook či Microsoft.

Existují metody, jak geolokalizaci a identifikaci svých zařízení blokovat nebo ztěžovat. V případě zařízení s GPS a polohovými službami dnes víceméně lze nastavit, kdy tyto údaje sledují a s kým je sdílejí. IP adresy hardwaru, který se využívá k prohlížení internetu, se mohou manuálně nahradit za nové. V evropském prostředí také lze na seriózních webových stránkách zvolit, jestli jim uživatel umožní identifikátory, jako je IP adresa, sbírat a sdílet s dalšími subjekty. Toto nastavení lze také plošně přednastavit v některých internetových prohlížečích, například v Safari a dalších. Podle Muira a Oorschota (2006) je nejsnazší a nejlevnější se proti geolokalizaci a identifikaci uživatele pomocí IP adresy bránit pomocí tzv. proxy. Ta funguje jako jakýsi prostředník v síťové komunikaci, který zabraňuje vysledovat koncové zařízení. Muir a Oorschot (2006) nicméně dodávají, že i proxy lze obejít a uživatele vystopovat.

Co se identifikátorů svého Wi-Fi připojení týče, lze podniknout kroky, jak je před vnějším světem skrýt. Stoprocentní ochranu před odhalením právě fungujícího vysílače je nicméně zřejmě pouze přechod na kabelové připojení k internetu. Velmi omezená je také obrana proti zanechávání digitálního otisku. Podle Electronic Frontier Foundation (2022) je dosud nejlepším řešením používání alternativních internetových prohlížečů, jako je například Tor, popřípadě využívat blokátoři domén, které digitální otisk sbírají.

3.2 Cookies

Jednou ze stěžejních metod, jak webové stránky identifikují návštěvníky a sledují jejich chování, jsou soubory cookie, kterým se v Česku někdy říká také sušenky nebo koláčky. Jde o soubory, které stránky ukládají do prohlížečů a následně v nich čtou, díky čemuž mohou pamatovat, co uživatel na dané stránce dělal, a navázat tak v relaci.

Za vynálezce cookies se považuje Lou Montulli. Reportér serveru Quartz Rivero (2021) s Montullim zpracoval rozhovor. Stojí v něm, že Montulli coby zaměstnanec společnosti Netscape, která vyvíjela internetové prohlížeče, v roce 1994 řešil problém, jak dát internetovým stránkám vědět, že na nich uživatel už v minulosti byl. Publicista Daniel Dočekal v již zmíněném pořadu Káva o čtvrté (Vlčková, 2022) posluchačům vysvětlil, že z dnešního pohledu tak cookies efektivně řeší velmi praktické záležitosti, jako je například odpadnutí nutnosti pokaždé se přihlašovat do stejné webové stránky nebo schopnost e-shopu pamatovat si, co už má člověk při prohlížení stránek internetového obchodu v nákupním

košíku. Inzerenti ale už po pár letech existence cookies dokázali tyto soubory využívat k cílení reklam a ke sledování chování a návštěvní historie uživatele. Podle Dočekala se díky cookies a jejich schopnosti ukládat o uživateli velmi podrobné informace napříč internetem rozvinul celý reklamní průmysl, který umožňuje na lidi cílit velmi přesnou reklamu.

Sám Montulli se v rozhovoru pro Quartz (Rivero, 2021) nechal slyšet, že cookies tvořili s důrazem na soukromí lidí a jejich následné využití pro reklamu prý nečekali. Podle dalšího článku na serveru Quartz (2022) v roce 1995 firma DoubleClick odhalila v cookies potenciál sledovat uživatele napříč internetem. V roce 2008 pak tuto společnost koupil za 3,1 miliardy amerických dolarů Google se záměrem rozšířit své byznysové aktivity i na poskytování cílené reklamy na základě cookies.

Jak už bylo uvedeno výše, tzv. cookies třetích stran - tedy cookies, které zjednodušeně řečeno nezajišťují základní funkce jednotlivých internetových stránek, nýbrž sbírají informace o historii prohlížení a chování napříč internetem, - jsou pro tzv. big tech firmy, jako je Google nebo Facebook, jedním ze zásadních prostředků pro sběr dat a jejich následnou monetizaci. Již zmíněné studie DuckDuckGo (2020) a Liberta (2014) poukázaly na to, že jsou to právě ony, kdo technologie pro cookies třetích stran nejčastěji zprostředkovávají.

Množství informací, které cookies umožňují o daném uživateli nasbírat, a také způsob, kterou se tak děje – tedy automaticky a v naprosté většině případů mimo vědomí návštěvníků – vytvořil zejména u evropských úřadů předpoklad, že dochází k manipulaci občanů, které firmy zneužívají k agregaci co největšího množství dat, a tudíž také zisků. Evropský parlament tedy v roce 2016 přijal směrnici General Data Protection Regulation, známou pod zkratkou GDPR. V ní mimo jiné stanovil, že provozovatelé internetových stránek jsou povinni vyžadovat souhlas či nesouhlas návštěvníka se sběrem cookies a se zpracováním jeho dat (Nařízení 2016/ 679). Mezi důvody vzniku regulace mimo jiné stojí:

„Rozsah shromažďování a sdílení osobních údajů významně vzrostl. Technologie umožňují jak soukromým společnostem, tak orgánům veřejné moci využívat při provádění jejich činností osobní údaje v nebyvalém rozsahu. Fyzické osoby stále častěji své osobní údaje zveřejňují, a to i v globálním měřítku. Technologie změnily ekonomiku i společenský život...“ (Nařízení 2016/ 679, str. 2)

Směrnici provozovatelé webů přijali s různou mírou důslednosti. Jak napsal Vincent (2020) pro server The Verge, GDPR se mnohým podařilo obejít tak, že uživatelé jsou

nejrůznějšími způsoby nucení sběr a zpracování cookies přijmout, aby se k obsahu na stránce dostali. Autor článku používá termín „cookie wall“, tedy bariéru, kterou člověk musí odkliknout, aby ho stránka pustila dál nebo aby ho v pohybu po webu neomezovala.

Nastavení soukromí rázem také začaly provázet manipulativní mechanismy, které pro uživatele zesložitují zákaz sbírání cookies na dané webové stránce. Tyto mechanismy se anglicky nazývají „dark patterns“. Ze studie provedené výzkumníky z univerzity v Aarhusu vyplynulo, že naprostá většina nejnavštěvovanějších britských stránek dark patterns manipulujících uživatele k souhlasu se sběrem uživatelských dat využívala (Nouwens a kol., 2020). Lomas (2021) pro TechCrunch dokonce konstatovala, že většina internetových stránek pravidla GDPR záměrně nedodržuje a nezdělaná uživatelům možnost nesouhlasit se sběrem cookies ani nenabízí.

Francouzský úřad pro ochranu osobních údajů v roce 2022 informoval o udělení pokuty 150 milionů eur americké společnosti Google a 60 milionů eur americké společnosti Facebook za to, že návštěvníky svých webů naváděly k udělení souhlasu se sběrem cookies, a to tím, že zatímco k udělení souhlasu stačil jeden klik, pro zvolení nesouhlasu jich bylo nutné udělat více. (CNIL, 2022)

Od ledna také 2022 platí přísnější pravidla souhlasu s ukládáním cookies i v Česku. Novelu Zákona o elektronických komunikacích Česko schválilo během roku 2021. Upravilo se v něm, jak se cookies sbírají. Z pohledu uživatele se změnilo zejména to, že na něj musejí weby požadovat souhlas či nesouhlas s jejich sběrem a možnost určit, které cookies je ochotný využívat. (Zákon 374/2021, 2021)

Regulace, které provozovatelům webových stránek a služeb komplikují sběr osobních údajů, motivují firmy k hledání alternativních monetizačních strategií. Uvědomili si to samotní technologičtí giganti. Například aktualizace webového prohlížeče Safari od společnosti Apple začátkem roku 2020 začala ve svém výchozím nastavení blokovat cookies třetích stran. Další velmi rozšířený prohlížeč Firefox od společnosti Mozilla se k tomuto kroku odhodlal ještě dříve. (Statt, 2020) Ve výchozím nastavení to z nejrozšířenějších prohlížečů západního světa nemá pouze Chrome, jehož autorem je Google. Jde o celosvětově nejpoužívanější prohlížeč: podle Statcounter (2022) má zhruba 65% podíl na trhu. Jak informoval mezi mnohými server BBC, i Chrome ale oznámil odchod od využívání cookies do roku 2023 (BBC, 2021). Pozdější přechod k zákazu cookies třetích stran vysvětlil Privacy Engineer Director Vinay Goel (2021) v blogovém článku. Píše, že je potřeba dostatek času na to, aby se provozovatelé webových stránek, kteří se spoléhali na výdělcích spojených

s používáním cookies mohli zařídit a přešli na nový model monetizace, který Google hodlá v souladu s ochranou soukromí uživatelů internetu do tohoto data vyvinout a zavést.

Skoncovat s cookies třetích stran chce i český ekvivalent Googlu, Seznam, který je v tuzemsku velmi silným hráčem v online vyhledávání a provozování cílené reklamy. Ve svém tištěném informačním letáku z prosince 2021 doslova píše, že:

„V důsledku změn, které se chystají v souvislosti s koncem cookies třetích stran, budou ohroženy příjmy z cílené reklamy. Řešením je přihlašování uživatelů přes identitu Seznam.cz, které dokáže identifikovat uživatele napříč doménami. Pokud bude mít web našeho klienta přihlášené uživatele, můžeme i nadále pracovat s cílenou reklamou, a zpeněžit tak svůj obsah.“ (Seznam, 2021, str. 1)

Nutno podotknout, že být se může na první pohled zdát, že se provozovatelé internetových stránek a služeb vzdávají cookies třetích stran neradi, opak může být pravdou. Největší hráči v tlaku na konec cookies třetích stran zjevně vidí nové příležitosti. Podle mnohých expertů i z toho důvodu, že mohou přejít na nový model cílení reklamy, který ještě více upevní jejich dominanci na trhu online marketingu.

Novinář, publicista a vysokoškolský pedagog Petr Koubský, který se mimo jiné soustředí na technologická a novomediální témata, v podcastu Online Plus, který vysílá Český rozhlas Plus, k očekávanému konci cookies třetích stran a jejich nahrazení přihlašováním přes Seznam identitu řekl, že vidí snahu splnit literu zákona, který zakáže cookies třetích stran, a zároveň si ponechat i nadále funkčnost cílit reklamu nejen na svém webu, ale i napříč internetem. Podle něj je zákaz cookies třetích stran poměrně tvrdý vůči menším inzerentům a menším reklamním systémům a zvýhodňuje největší digitální firmy. Podle něj se také ty největší firmy, jako je Facebook a Google, bez cookies třetích stran docela dobře obejdou, protože existuje spousta cest, jak se k informacím o uživatelích dostat. Podle Koubského je navíc patrné už nějakou dobu, že se provozovatelé webů snaží motivovat uživatele, aby si u nich založili účet, přihlašovali se do něj, a tudíž dávali možnost být sledováni, podobně jako je to od začátku na Facebooku nebo Googlu. Koubský dokonce v pořadu tvrdí, že soukromí na internetu je ztracené navždy, neboť provozovatelé internetových stránek a služeb vždy donutí uživatele, aby souhlas se sběrem informací dali. Dodává také, že co se týče cookies třetích stran, většina běžných uživatelů vůbec netuší, co dělají. Zdůraznil také, že v dnešní době už by mělo být každému jasné, že na webech a službách, kam je nutné se přihlásit, dáváme své soukromí k dispozici a že vše, co tam

děláme, je pro provozovatele naprosto průhledné a viditelné. Navíc existují další velmi rozšířené způsoby sběru dat, které jsou cookies velmi podobné a které se zatím do hledáčku regulátorů nedostaly. (Slížek, 2021)

3.3 Pixels a beacons

Jedním z nich je například tzv. pixel, jiným označením beacon či web bug. Nabízí ho třeba Facebook pod názvem Facebook pixel (nebo Meta pixel) či Google. Využívá se ke sledování chování návštěvníků online, které se následně zužitkovává primárně k marketingovým účelům (Facebook, 2022a). Jde o podobný mechanismus jako v případě cookie – rozdíl je v tom, že data z pixelu se neukládají v počítači uživatele, ale na uložištích svého provozovatele. Podle International Association of Privacy Professionals (2022) je pixel z pohledu návštěvníka neviditelný – jde o kus kódu, který provozovatel nahraje na své webové stránky či vloží do e-mailů. Díky tomu pak může sbírat a sdílet data o chování návštěvníků.

Už před lety Smith (1999) pro Electronic Frontier Foundation uvedl, že skryté technologie typu web bug neboli beacon umožňují provozovateli webu a technologie sledovat IP adresu přístroje návštěvníka, URL stránky s web bugem, kterou navštívuje, URL daného web bugu, beaconu či pixelu, čas, kdy k otevření stránky došlo, typ použitého internetového prohlížeče, a v neposlední řadě i údaje z cookies v zařízení návštěvníka.

Beacon nebo pixel bývá schovaný také v e-mailech, zejména pak v těch, které obsahují marketingová sdělení prostřednictvím k tomu určených softwarů, například službou Mailchimp. V článku Krasnoff (2021) pro server The Verge se lze dočíst, že cílem obchodníků a služeb typu Mailchimp není pouze marketing, ale také extrakce dat, neboť v okamžiku otevření e-mailu obsažený beacon ihned zasílá odesílateli tohoto sdělení polohové a časové údaje. Podle autorky článku se proti těmto sledovacím technologiím dá bránit jen omezeně, a to zákazem automatického stahování obrazových dokumentů v nastavení e-mailových aplikací.

Koubský v jiné epizodě pořadu Online Plus (Slížek & Skácel, 2022) komentuje, že jako uživatelé máme pouze velmi chabý přehled o tom, co pixels sledují a jak vlastně fungují. Existují podle něj jen rozsáhlé domněnky, co vše je například Facebook pomocí této technologie schopen sledovat. V únoru 2022, jak připomíná, proto internetový vyhledávač Mozilla (2022) oznámil spuštění studie zvané Mozilla Pixel hunt, po jejímž vyhodnocení by bylo možné uceleněji potvrdit, co vše Facebook pixel o lidech sbírá za data.

3.4 Sociální sítě a aplikace

Výše zmíněné příklady extrakce dat platí zejména pro prohlížení na internetu, jinými slovy, navštěvování webových stránek, e-shopů apod. Zejména v případě užívání sociálních sítí a podobných aplikací ale bývá pravidlem, že se sbírá a analyzuje vše, co na nich uživatel dělá. Například Meta ve svých podmínkách užívání služeb z jara 2022 uvádí, že sbírá a analyzuje nejen data, která o sobě uživatelé aktivně prozrazují v podrobnostech svého profilu či ve statusech nebo osobních zprávách, ale také data plynoucí z interakce s obsahem, například v podobě času stráveného nad jednotlivými příspěvky. Sbírá tedy nejen lajky či komentáře, ale také pouhou pozornost jednotlivým obsahům. V podmínkách se dále píše, že Facebook sleduje i pohyb kurzoru, či identifikátory, jako typ a nastavení zařízení, včetně stavu baterie. V zásadě využívá všechny výše popsané prostředky, od souborů cookie po geolokaci v podobě dostupných přístupových bodů s připojením k internetu. Meta deklaruje, že to vše využívá nejen pro personalizaci obsahu na svých sítích a vylepšování svých produktů, ale také k lepšímu cílení reklam. Souhlasem s těmito podmínkami uživatelé souhlasí také s výzkumnými záměry společnosti a se sdílením těchto dat svým partnerům. (Facebook, 2022b)

Zatímco proti řadě metod sbírání a analýzy dat se lze bránit, jediným zaručeným způsobem jak aplikacím typu Facebook zabránit sběru dat na svých službách, je tyto služby nepoužívat. Pokud si člověk nepřeje pouze to, aby se o něm data sbírat přestala, ale také aby došlo k výmazu těch, které o něm společnost již nasbíraly, v evropském prostoru je díky evropským nařízením GDPR možné požádat o výmaz v souladu s „právem být zapomenut“. Mělo by to být možné u všech online služeb, jejichž sběr dat závisí na výslovném souhlasu uživatele. (Nařízení 2016/ 679)

Jak už bylo nastíněno v předchozích částech této práce, provozovatelé online služeb a webů s data poskytují třetím stranám, tudíž i v případě, že člověk docílí výmazu ze všech online služeb, které používal, nemusí to znamenat, že se jeho profil z trhu s uživatelskými daty vytratí. Pro smazání digitální stopy by teoreticky bylo nutné docílit výmazu naráz u všech subjektů, které o dané osobě mohou data vlastnit, včetně celé řady agregátorů a prodejců dat, se kterými běžný návštěvník internetu do přímého kontaktu nepřichází. I z toho důvodu se na trhu objevily služby, které zprostředkovávají hledání digitální stopy, včetně jejího mazání. Jednou z takových je například DeleteMe. Ta na svých webových stránkách (JoinDeleteMe, 2022) uvádí, že je schopna za roční poplatek 10,75 amerických dolarů průběžně hledat a mazat digitální stopu u stovek agregátorů osobních dat a služeb.

4. Shrnutí teoretické části a řešení

Z výše uvedených informací plyne, že množství dat, které provozovatelé webových stránek a služeb o svých návštěvnicích a uživateli sbírají, se v průběhu zhruba dvacet let zvětšilo natolik, že je prakticky nemožné vyjmenovat, s čím vším společnosti pracují a co vše o jedincích znají. O některých metodách sběru dat a jejich analýze a obecně osudu veřejnost, včetně té odborné, dosud ani nemá hlubší znalosti. (Elsaify & Hasan, 2020) Nejasné je také to, s nakolik sofistikovanými informacemi o uživateli firmy pracují. Najde se řada autorů, která tyto mezery v poznání pomocí dedukce doplňuje. S ohledem na zájmy technologických společností a s ohledem na informace, které jsme dosud měli možnost zaznamenat, se tito experti domnívají, že firmy jsou schopny mít precizní obrázek o svých uživateli, za který jsou také schopné vydělat nemalé peníze. (Zuboff, 2019; Crain, 2021)

I z několika výše uvedených oblastí sběru dat vyplývá, že možnosti obrany proti sběru osobních dat je v online světě omezený, ne-li nemožný. (Acquisti a kolektiv, 2020; Schneier, 2013) Omezování či ztěžování sběru dat pro uživatele zpravidla znamená vynaložit úsilí a nezdarma také akceptovat horší funkce a komfort. Ve výsledku se tak můžeme setkat s tvrzeními, že soukromí na internetu je definitivně ztraceno (Sarat, 2014) či že jediný způsob, jak data s internetovými firmami data nesdílet, je internet nepoužívat a dobrovolně se vyčlenit z účasti v moderní masové společnosti. (Crain, 2016)

Přesto je zatím nejasné, jestli a jak lze cenu dat určit. (Collis, 2022; OECD, 2013). Profesor Carnegie Mellon University Acquisti, který se ve svých výzkumech a článcích na cenu soukromí a dat zaměřuje, v reakci na jednu ze svých studií v roce 2019 pro *The Wall Street Journal* řekl, že „ekosystém online reklam je komplexní a neprůhledný.“²³ (Hagey, 2019). Například Collis (2022) se ve svém výzkumu, během něžž měli lidé odhadovat cenu svých dat z Facebooku, se spolu s kolegy sami neodvážili cenu stanovit.

²³ Volně přeloženo z anglického originálu: „*The online ad ecosystem is complex and opaque.*”

5. Povědomí Čechů o sběru osobních dat na internetu

V této části se bude práce věnovat kvalitativnímu výzkumu. Ten se opírá o hloubkové rozhovory s pražskými vysokoškoláky na téma hodnoty dat, která o sobě v online prostoru zanechávají. Nejprve představím metodologii a cíle výzkumu, posléze výsledky a zjištění hloubkových rozhovorů a nakonec závěry výzkumu, díky nimž dokážeme lépe pochopit problematiku, které se tato práce věnuje.

5.1. Cíl výzkumu a výzkumné otázky

Při rešerši existujících výzkumů na téma ceny soukromí na internetu z pohledu uživatelů vyšlo najevo, že v mezinárodním měřítku lze najít práce, které se touto problematikou zabývají (Collis, 2022; Johnson, 2020; Acquisti, 2020, 2013; Svirsky, 2019; Malgieri & Custers, 2018; Savage & Waldman, 2013; Tsai et al. 2011). V Česku dosud vznikaly spíše kvantitativní průzkumy, např. zmíněný průzkum agentury STEM/MARK (Jindrová, 2021). Záměrem výzkumné části této práce je proto provést první sondu do kvalitativních výzkumů na toto téma. Výzkumné otázky se zaměřují na význam soukromí na internetu pro respondenty, na to, jaký mají uživatelé internetu přehled nad tím, co se o nich v online prostředí sbírá, a jakou si myslí, že mají jejich data hodnotu. Pro praktickou část diplomové práce jsou vytyčeny celkem tři výzkumné otázky:

VO1: Uvědomují si lidé aktivní v online prostředí, že jejich data mají hodnotu?

VO2: Nakolik si lidé aktivní v online prostředí cení své soukromí?

VO3: Podnikají lidé nějaké aktivní kroky k tomu, aby o sobě zanechávali méně dat?

5.2 Metoda, nástroje a průběh výzkumu

Pro zodpovězení navržených výzkumných otázek jsem zvolil metodu kvalitativního výzkumu. Ten je založený na logickém procesu indukce, tedy na vyvozování obecných závěrů na základě dílčích poznatků. Kvalitativní výzkum testuje existující teorie a vytváří nové. (Sebera, 2012) Klíčovým výzkumným prostředkem je podle Cresswella (2009) sám výzkumník. Mezi výhody kvalitativního výzkumu patří: podrobný vhled do jedince či

skupiny, zkoumání v přirozeném prostředí, možnost navrhování nových teorií, snadná reakce na místní podmínky a situace a pomoc při počáteční exploraci fenoménů. (Sebera, 2012)

V souladu se záměry této práce byla zvolena výzkumná metoda hloubkových rozhovorů, a to technikou tzv. rozhovoru pomocí návodu. Ta hlídá, aby se rozhovory dotkly všech relevantních témat. Jak uvádí Hendl (2005), jde o techniku, která dotazovaným dopřává určitou svobodu ve svých výpovědích a která výzkumníkovi zjednodušuje analýzu a tzv. kódování dat.

Návod rozhovoru má tři tematické celky (viz Příloha č. 1: Návod k hloubkovému rozhovoru). První celek má za cíl hlavně účastníky výzkumu rozmluvit a dát jim čas se aklimatizovat před před pokládáním pro výzkum stěžejních otázek. Ty se nacházejí v následných dvou částech návodu rozhovorů, tvoří jejich hlavní náplň a přímo se týkají výzkumných otázek. V průběhu samotných rozhovorů byl kladen důraz na to respondenty nehodnotit a věnovat pozornost sdělením, které předávají, aby se nezapomnělo na případnou sondáž pomocí doplňujících otázek. Veškeré hloubkové rozhovory byly nahrávány pomocí nástroje Diktafon na mobilním telefonu iPhone.

5.3 Výběr a skladba výzkumného vzorku

Co se výběru respondentů týče, pro cíle tohoto výzkumu, kterým je provedení jednoho z prvních kvalitativních výzkumů mezi Čechy na téma sběru osobních dat online, se zdálo nejvhodnější najít respondenty, kteří výzkumníkovi důvěřují, aby byli ve svých výpovědích otevření a upřímní. Výzkumný vzorek je tedy omezený na skupinu pražských vysokoškoláků do 30 let, kteří se s výzkumníkem už předem znali, s jednou výjimkou.

Záměrem následně bylo zajistit, aby byl vzorek co nejvíce heterogenní (co se týče pohlaví, oboru i času stráveného online) a poskytl tak různorodé pohledy na zkoumané téma. Do vzorku byly zahrnuty pouze osoby, které jsou v online světě aktivní a které vedle u nás nejpoužívanějších internetových vyhledávačů, jako je Seznam nebo Google, využívají také alespoň jednu z nejpoužívanějších sociálních sítí, jako je Facebook, Instagram, Twitter či TikTok.

Výzkumný vzorek se skládá celkem z 10 respondentů. Není v něm zahrnuta osoba, která se zúčastnila pilotního výzkumu. Průměrný věk respondentů je 25 let a pro svou osobní potřebu na internetu v průměru každodenně stráví 3,5 hodiny. Všichni respondenti momentálně bydlí v Praze. Podrobnější charakteristiky výzkumného vzorku lze zhlédnout v Tabulce č. 1.

Jméno	Věk	Odhad času stráveného na webu (denně)	Univerzita a fakulta
Petr	29	4h	FCHI VŠCHT
Martin	25	6h	FSV UK
Lucie	23	1h	FPH VŠE
Jaroslav	27	2h	FCHI VŠCHT
Markéta	23	2h	FF UK
Jan	25	5h	FSV UK
Tereza	24	4h	1 LF UK
Klára	23	3h	PRF UK
Marek	27	4h	FSV UK
Marie	25	4h	FŽP ČZU

Tabulka č. 1: Věk, odhad stráveného času na internetu, navštěvovaná univerzita a fakulta účastníků hloubkových rozhovorů. Jména nejsou shodná s opravdovými jmény účastníků výzkumu.

5.4 Pilotní výzkum

Pilotního výzkumu se zúčastnila jedna osoba, a to muž. Šlo o velmi důležitou část výzkumu, která nejenže ověřila srozumitelnost otázek, ale také pomohla s praktickým nácvikem vedení rozhovorů podle návodu. Pilotní výzkum také poskytl odhad množství času, který bude účast ve výzkumu respondenty stát. Pilotní rozhovor trval přibližně 30 minut. V návodu rozhovorů neodhalil žádné zásadní nedostatky nebo nesrozumitelnosti. Následné výzkumné rozhovory trvaly zhruba od 20 do 35 minut.

5.5 Analýza dat

Zvukové záznamy rozhovorů byly nejprve doslovně přepsány a pak byly jazykově očištěny. Následoval proces kódování s cílem roztřídit získaná data. Kódování se skládalo ze tří fází: V první fázi bylo použito tzv. otevřené kódování. Během něj se na nižší abstraktní úrovni hledají a označují témata, která se většinou vztahují k výzkumným otázkám, případně

k teoretickému základu práce. Otevřené kódování může také poukázat na nová témata, kterým se bude v následných krocích ještě potřeba věnovat.

Po otevřeném kódování přichází fáze tzv. axiálního kódování. To už mezi témata označených v otevřeném kódování hledá souvislosti, včetně příčin, důsledků, podmínek a interakcí. Tvoří také jednotlivé kategorie a hledá další, které souvisejí. (Hendl, 2005)

Třetí fází je tzv. selektivní kódování. Během něj výzkumník identifikuje pro výzkum stěžejní témata. K selektivnímu kódování může dojít až v okamžiku, kdy jsou na základě axiálního kódování všechna relevantní témata a kategorie rozkryté. Pracuje se také s jejich porovnáváním a v datech se také hledají názorné příklady. (Hendl, 2005)

Hendl (2005) uvádí, že jednotlivé fáze kódování „...se nemusí používat zcela odděleně. Představují spíše různé způsoby, jak zacházet s textovým materiálem, a výzkumník může mezi nimi podle potřeby přecházet. V každém případě začíná analýza dat otevřeným kódováním a ke konci analýzy se provádí spíše selektivní kódování.“ (str. 247)

5.6 Etika výzkumu

Na začátku každého setkání s účastníkem výzkumu byla rámcově představena výzkumná metoda hloubkových rozhovorů a jaké téma bude jejich náplní. Následovala prosba o možnost pořizování hlasového záznamu, obeznámení se se způsobem využití nasbíraných dat a také seznámení s právy účastníka výzkumu, jako je například přerušování účasti ve výzkumu bez nutnosti udat důvod. Všichni účastníci výzkumu následně vlastnoručně podepsali informovaný souhlas účastníka výzkumu (Příloha č. 2), který měli možnost předem nerušeně prostudovat. Po skončení rozhovoru naprostá většina setkání ještě pokračovala neformálním hovorem o zkoumaném tématu.

5.7 Limity výzkumu

Jak uvádí Hendl (2005), je potřeba také upozornit na skutečnost, že závěry z kvalitativního výzkumu nemusí být zobecnitelné na širší populaci a do jiného prostředí. Vzhledem k tomu, že výzkumný vzorek výzkumu popsaného v této práci tvořili pražští vysokoškoláci pod 30 let věku, je Hendlovo upozornění držet v patrnosti. Kvůli použitému výzkumnému vzorku, ve kterém řada pražských vysokých škol chybí, mohlo také dojít socio-demografické limitaci výzkumu. Nakonec je nutno uvést, že limitace představuje ze své podstaty samotný kvalitativní výzkum, protože jeho průběh i zpracování se odvíjí od

subjektivních rozhodnutí osoby výzkumníka, který tak do výzkumu nevyhnutelně promítá vlastní preference. (Hendl 2005)

5.8 Výsledky výzkumu

5.8.1 Průběh rozhovorů a charakteristika účastníků

Všechny rozhovory byly realizované osobně v prostředí, které vybrali sami respondenti, aby jim bylo příjemně. Ve všech případech šlo klidný koutek ve veřejně přístupném prostoru. Každý rozhovor následoval návod (viz Příloha 1: Návod k hloubkovému rozhovoru). Po skončení rozhovoru a vypnutí zvukového nahrávání ve všech případech navazovala neformální konverzace, zpravidla k tématu výzkumu.

Petr, 29 let, VŠCHT, odhadovaný denní průměr času na internetu: 4 hodiny

Petr používá internet hlavně pro zábavu. K tomu využívá zejména Youtube, Instagram. Účet má ale i na dalších sociálních sítích, jako je Facebook či LinkedIn s tím, že Facebook a jeho chatovací platformu Messenger využívá hlavně pro komunikaci s přáteli. Byť má k internetu kladný vztah - během rozhovoru řekl, že na internetu musel být jako jeden z prvních - při prohlížení internetu využívá jak na počítači tak mobilu blokátory a činí aktivní kroky k tomu, aby nejen omezoval počet zobrazovaných reklam, ale aby také minimalizoval digitální stopu, která by vedla k jeho osobě. Je ochotný si za to zaplatit. O soukromí na internetu začal dbát v okamžiku, kdy stahoval torrenty či prozkoumával dark web. Zároveň nechce dávat svá opravdová data obchodníkům.

Martin, 25 let, FSV UK, odhadovaný denní průměr času na internetu: 6 hodin

Martin při studiu vysoké školy pracuje jako novinář a internet je proto pro něj zásadním prostředím, ve kterém se hranice mezi prací a soukromím do určité míry stírá. Na internetu pravděpodobně tráví ze všech účastníků výzkumu nejvíce času. K internetu má velmi kladný vztah - je to pro něj nejen zdroj informací a užitečný pomocník v každodenním životě, ale také zdroj zábavy. Účet má na většině sociálních sítí, včetně Twitteru a TikToku. O tom, že se o něm na internetu sbírají data pro komerční účely, z určité části věděl, vadí mu to ale jen velmi málo. Naopak reklama ho obtěžuje a to natolik, že je ochoten si v krajním případě i zaplatit za to, aby se jí zbavil.

Lucie, 23 let, FPH VŠE, odhadovaný denní průměr času na internetu: 1 hodina

Lucie na internetu (pro soukromé účely) tráví ze zúčastněných odhadem nejméně času. Z velké většiny podle ní jde o čas strávený na Instagramu a síti Youtube. Jde u ní hlavně o zábavu, pak následuje komunikace s přáteli. Reklama ji obtěžuje, nicméně ne natolik, aby proti ní podnikala nějaké restriktivní kroky. O žádných blokátorech reklam dosud nevěděla. O tom, že online služby a weby sbírají data o svých uživatelích, možná tušila, ale více do hloubky o tom nikdy nepřemýšlela. Nijak jí to ale nevadí. Na druhou stranu si vzpomíná na okamžik, kdy se jí telefon zeptal, jestli chce data sdílet s aplikacemi. To zvolila zákaz. Internet vnímá velmi pozitivně, jako neoddělitelnou součást svého života. Podle Lucie to tak je, protože je součástí generace, která na internetu vyrostla.

Jaroslav, 27 let, VŠCHT, odhadovaný denní průměr času na internetu: 2 hodiny

Jaroslav podle svých slov internet spíše než ke komunikaci nebo sociálním sítím používá jako zdroj informací o vývoji na investičním trhu a ke sledování aukcí investičního alkoholu. Podle něj je to na jednu stranu zábava, ale na druhou stranu se investicím doopravdy věnuje, takže jde o zábavu, která může přerůst v přivýdělek. Jaroslav není na sociálních sítích příliš aktivní, na Facebooku využívá pouze Messenger a na Instagramu sleduje jen dění na univerzitě. Čistě zábavu Jaroslav nachází na Youtube, který má schopnost jeho pozornost pohltit, což Jaroslav vnímá negativně a během rozhovoru zdůraznil, že má tendenci čas strávený na sociálních sítích omezovat, včetně komunikace, kterou raději s lidmi uskutečňuje v reálném světě, třeba u piva.

Markéta, 23 let, FF UK, odhadovaný denní průměr času na internetu: 2 hodiny

Markéta označila čas, který na internetu tráví, jako lelkování, během nějž se věnuje sociálním sítím – hlavně Instagramu a Facebooku. K internetu má pozitivní vztah. I její brigády v oblasti marketingu by se bez internetu neobešly. Reklamu na internetu vnímá, vadí jí hlavně ve videích. Některé cílené reklamy v ní spíš vyvolávají smích, protože jsou prý targetované špatně. O tom, že se na internetu data sbírají, určité povědomí měla, a to díky svému příteli.

Jan, 25 let, FSV UK, odhadovaný denní průměr času na internetu: 5 hodin

Pro Jana je internet hlavně zdrojem informací. Čte zprávy nebo se věnuje pasivnímu sledování Youtube či jiných platforem s mluveným slovem. Má Facebook a Instagram a obecně internet vnímá pozitivně, protože podle něj jde o všeobecně dostupnou studnici informací. Co naopak vnímá negativně či z čeho má obavy, jsou sociální sítě, které z pohledu

Jana lidi zotročují. Čas strávený na sociálních sítích se proto snaží omezovat. Jan si nedělá iluze o pravých záměrech technologických společností, včetně toho, jak zpracovávají a využívají uživatelská data. Jejich sběr mu vadí, stejně tak reklama na internetu.

Tereza, 24 let, 1. LF UK, odhadovaný denní průměr času na internetu: 4 hodiny

Pro Terezu internet znamená hlavně sociální síť. Na rozdíl od ostatních účastníků výzkumu na nich nejen obsah konzumuje, ale také ho hojně produkuje. Instagram je pro ni navíc hlavní komunikační platformou pro kontakt s přáteli. Dle jejích slov je scrollování Instagramem a Facebookem relax, ale zároveň vymývání mozku. Byť ví, že se o ní data na internetu sbírají, nemá přesné ponětí o tom, co vše s nimi společnosti dělají. Soukromí na internetu chápe hlavně jako to, co o sobě vědomě na sociální síť neprozradí ostatním.

Klára, 23 let, PrfUK, odhadovaný denní průměr času na internetu: 3 hodiny

Klára je jediný účastník výzkumu, který se s výzkumníkem předem neznal. Internet využívá hlavně pro orientaci ve zprávách a aktualitách, které konzumuje prostřednictvím Twitteru. Jde o sociální síť, kterou prý využívá v současné době nejvíce ze všech. Nicméně sama na ní obsah neprodukuje, pouze ho konzumuje. Stejně tak činí v případě Instagramu. Druhou aktivitou je komunikování s přáteli, a to zejména prostřednictvím Messengeru. Klára si uvědomuje, že online služby sbírají uživatelská data, a i z toho důvodu je třeba na onen Twitter přihlášená pod jiným jménem a nevytváří tam vlastní obsah. Jejimi slovy: „*Jsem paranoidní. To, že o sobě všichni všechno sdílejí, nemá evokovat pocit, že jednotlivá osobní data nejsou důležitá nebo zneužitelná.*“

Marek, 27 let, FSV UK, odhadovaný denní průměr času na internetu: 4 hodiny

Marek se označuje za pasivního konzumenta internetového obsahu, který občas něco dělá i aktivně, třeba hraní online šachů. Sleduje hlavně Facebook a Youtube. Relativně krátce má také založený Instagram. Uvědomuje si, že používáním internetu a online služeb o sobě firmám prozrazuje data - sám říká, že internet není zadarmo, že on sám je produktem, který provozovatelé webových stránek nabízejí inzerentům. Nevidí v tom ale problém, dokud si to uvědomuje. Reklamy mu vadí více a snaží se je blokovat. Největší výhodu ale spatřuje v tom, že je vše zadarmo.

Marie, 25 let, ČZU, odhadovaný denní průměr času na internetu: 4 hodiny

Marie dělá na internetu mix aktivit - od komunikace s přáteli, vyhledávání informací a zábavy, například ve formě poslechu hudby. Ze sociálních sítí používá nejvíc Instagram, ale skoro výhradně jen jako sledující, který obsah nepřidává. Uvědomuje si, že webové stránky sbírají data o uživateli. I z toho důvodu nemá zájem na sdílení vlastního obsahu na sociálních sítích. Obzvláště dbá o soukromí své čtyřleté dcery. Podle Marie mohou data, která o sobě na sociálních sítích zanecháváme či která o nás jejich provozovatele znají, znamenat v budoucnu problém.

5.8.2 Analýza hloubkových rozhovorů s účastníky

Na základě následného kódování byly vytvořeny kategorie, které byly rozděleny do podkategorií (Tabulka č. 2). Aby bylo zabráněno případným nesprávným interpretacím, budou v průběhu prezentování výsledků uváděny celé výroky či pasáže zúčastněných hloubkových rozhovorů.

Kategorie	Podkategorie
Účel uživatelských dat	<ul style="list-style-type: none">• Data jako jeden ze způsobů zisku• Personalizace reklamy
Domnělá cena dat	<ul style="list-style-type: none">• Odhad hodnoty• Jaká data mají z pohledu firem hodnotu?• Hodnota se mění
Kroky vedoucí k zanechávání menšího množství dat	<ul style="list-style-type: none">• Nesouhlas s cookies• Blokátory a další metody omezování digitální stopy
Důvody, proč svá data uživatelé poskytují	<ul style="list-style-type: none">• Ostatní to dělají také• Nevadí mi to• Nemá to smysl
Jsou data z pohledu uživatelů platidlo?	<ul style="list-style-type: none">• Raději data než peníze• Lidé nechtějí za online služby a obsah platit penězi• Čas jsou peníze

Tabulka č. 2: Kategorie plynoucí z analýzy hloubkových rozhovorů.

Účel uživatelských dat

Data jako jeden ze způsobů zisku pro internetové stránky a služby

Většina účastníků výzkumu (n=7) uvedla, že data slouží provozovatelům webových stránek nebo služeb, jako zdroj příjmu. Rozlišují, že je lze používat nejen k personalizaci reklamy či služby, ale také jako artikl, se kterým se obchoduje se třetími stranami.

Jasně to artikuloval například Petr v reakci na otázku, jaký mají internetové stránky nebo služby byznys model: „*A pak je třetí varianta stránek, která v podstatě vydělává jen na tom, že sbírá statistiky o tobě: návštěvnost, které reklamy tě zajímají, na co klikneš, překlíky a tyto věci. Prodává se to třetím stranám.*“

Jaroslav si částečně protřečil, když řekl, že: „*Mám za to, že nějak placené tyto služby musejí být. A tak je podle mě lepší, že společnosti prodávají pozornost uživatele, a ne jeho data,*“ aby v odpovědi na následující otázku odpověděl, že: „*...všechny z nich si asi nějaké informace vzájemně poskytují. A je otázka, za jaké peníze,*“ a „*Hodně se ale také mluví o tom, jestli se tyto data prodávají třeba zaměstnavatelům, pojišťovnám nebo specializovaným agenturám.*“

Podobně to vidí i Martin. Jeho odpověď na otázku, kdo všechno se k jeho datům dostane, odpověděl: „*No tak v první řadě ten, kdo je sbírá a v druhý pak ten, komu je prodají.*“ Jan to shrnul takto: „*... [společnosti] počítali s tím, že to bude vždycky zadarmo a ty příjmy půjdou odjinud. Z reklamy nebo z prodeje dat.*“ Podle Kláry je zase prodej dat pro účel reklamy jen jedním z možných způsobů, jak provozovatelé online služeb s daty nakládají: „*Z toho, že získávají informace o uživateli, které potom můžou dál prodávat, třeba pro účely reklamy.*“

Co se Marka týče, z jeho odpovědi vyplývá, že společnosti mohou data využívat jako prodejní artikl. To potvrzuje i jeho předpoklad, že: „*[s] daty asi nemůže nakládat každý tak, jak se mu zlíbí.*“ A nakonec Marie na otázku, z čeho vydělává například Instagram peníze, uvedla, že: „*z reklamy a z informací, které berou o lidech.*“

Personalizace reklamy

Všichni účastníci výzkumu (n=10) si uvědomují, že zdrojem příjmu webů a online služeb, jako je Facebook nebo Instagram, je reklama. Uvědomují si také, že jde o reklamu, která je cílená na základě jejich vlastní digitální stopy. Ve všech rozhovorech tato znalost zazněla explicitně, pouze v případě Martina je nutné ji částečně hledat mezi řádky. Řekl

například: „...dokáží propojovat moje chování z jiných stránek s tím, jak mi nastavují to prostředí u sebe. Cílem je, abych vlastně chodil víc tam, kam oni chtějí, tedy tam, kde moje chování dokáží lépe monetizovat. A posléze na přímou otázku „Co třeba cílené reklamy?“ odpověděl: „No, tak ty na mě asi cílí.“

Domnělá cena dat

Odhad hodnoty

V každém z uskutečněných hloubkovém rozhovoru byla také část zaměřená na odhad finanční hodnoty, kterou uživatelská data mají. Jak se ukázalo v předešlé části, většina účastníků výzkumu si uvědomuje, že data firmy sbírají, aby je mohly dále prodávat. Všichni zúčastnění hloubkových rozhovorů vědí, že se data využívají k cílení reklamy. Jakou ale mají data podle respondentů hodnotu v penězích?

Petr, který se ze všech účastníků výzkumu údajně nejvíc snaží chránit své soukromí, odhaduje, že Facebook na průměrném evropském uživateli za měsíc nebo rok vydělá maximálně 10 centů. Hodnota dat v řádech centů byl také odhad Martina: „No, tak to se asi dá nějakým způsobem vyčíslit penězi, ale já nevím, jakou hodnotu v penězích to přesně má.“ Následně dodal: „A jestli si někdo o mně chce koupit všechna data, tak to může mít hodnotu dolarových centů?“.

Jan ani Marie se neodvážili hodnotu dat odhadovat. Stejně tak Lucie nevěděla, kolik by mohla uživatelská data stát, nicméně podle ní bude jejich hodnota „Asi dost velká“. Podobně reagovala i Markéta, podle které mají data „extrémní cenu“. Markéta posléze ještě odhadla cenu dat, která o ní Facebook za poslední rok nasbíral, na 10 tisíc korun. Odhad Kláry je ještě vyšší. Podle ní může cenovka dat, která má Facebook o jednom svém uživateli nasbíraná, dosahovat „řádů stovek tisíc dolarů“. Také podle Terezy mají data velkou hodnotu, která se ani „nedá v penězích vyčíslit“. Podle jejích slov mají osobní data dokonce „hodnotu lidských životů. Dají se strašně snadno zneužít.“

Jaroslav v odpovědích naznačil, že se by se hodnota dat mohla dát odhadnout podle výročních zpráv Facebooku, jehož absolutní většina příjmů plyne právě z reklamy. Řekl: „Já si myslím - ale je to jenom hrubý odhad a nevím to jistě, že v Evropě to bylo nějakých 15-20 dolarů ročně na jednoho uživatele a v Americe to byl dvojnásobek nebo trojnásobek...“. Marek si pak podobným způsobem odhad ceny dat, která o svých uživateli sbírá Facebook, vytvořil rychle, a to na základě jeho výdělku a počtu jeho uživatelů: „Asi to nebude málo vzhledem k tomu, že Facebook je velmi výdělečná firma. Ale kolik přesně, to nevím. Facebook

má dejme tomu dvě miliardy uživatelů, takže to vlastně může být kdekoliv mezi desítkami a stovkami dolarů za osobu?“

Vzhledem k tomu, že se už v teoretické části této práce ukázalo, že cenu dat nelze přesně stanovit, nelze ani hodnotit, jestli měli respondenti ve svých výpovědích pravdu. Co se nicméně hodnotit dá, je jistota, se kterou účastníci výzkumu na otázky týkající se ceny dat reagovali. Tři účastníci výzkumu (n=3) si ani netroufli cenu dat odhadnout. Další čtyři (n=4) respondenti uvedli, že nemají představu či nevědí, kolik by to mohlo být, byť následně nějaký odhad provedli. Dva respondenti udali odhad bez přílišného přemýšlení (Petr, Tereza), jeden po dlouhé pauze ve své promluvě (Klára).

Účastníkům výzkumu byl následně sdělen přibližný finanční obnos, který Facebook prostřednictvím reklamy utržil za průměrného uživatele v Evropě v roce 2020. Dva respondenti (n=2) usoudili, že jde o adekvátní sumu. Mezi nimi je Jaroslav, který na otázku, jestli mu to přijde moc, nebo málo, odpověděl: *„Jak se to vezme. Když se to vezme na tu masu lidí, která ho používá, tak to je úplně krásné.“* Druhým respondentem byl Marek. Toho částka nepřekvapila a sám ji odhadl na podobnou sumu.

Polovina účastníků výzkumu (n=5) označila něco málo přes 1000 korun, které v roce 2020 Facebook utržil za průměrného evropského uživatele pomocí reklamy, za překvapivě nízkou částku. Například Lucie prohlásila: *„1000 korun? Tak to jsem čekala mnohem víc ... Čekala jsem, že to budou alespoň stovky tisíc.“* Markéta řekla: *„Nepřijde mi to moc. Asi jsem čekala mnohem víc.“* Větší sumu čekala i Tereza: *„Přijde mi to jako strašně malá částka na to, jak moc nás může síť ovlivnit. Instagram a Facebook o lidech mohou vědět hodně choulostivé věci, a pokud by data někam unikla nebo je někdo zneužil, tak to může zkazit lidský život.“* A nakonec i Marie řekla, že je to *„málo, hodně málo.“*

Janova reakce je někde mezi oběma póly. Na jednu stranu mu hodnota připadá nízká, na tu druhou ale racionální a v součtu ohromná: *„Takhle najednou mi to zní, jako že je to dost málo, ale když si spočítám, kolik je těch uživatelů, tak jsou to vlastně šílené peníze... Jsou to velké miliardy. Ale takhle na jednoho mi to zní docela málo. To bych čekal, že dělají i víc, třeba tisícovku na měsíc. Ale asi to je dost.“*

Jediný, kdo čekal nižší sumu, byl Martin. Dodal ale, že se domnívá, že za něj konkrétně Facebook tolik neutřil: *„Já si myslím, že patřím do skupiny, na kterou cílení z hlediska nějaké demografie může být silné, ale moje osobní chování vybočuje z průměru této demografické skupiny.“*

Jaká data mají podle účastníků výzkumu pro firmy hodnotu?

Všichni respondenti (n=10) mají do určité míry nějakou představu o tom, co všechno se o nich na internetu sbírá. Rozlišují také, jestli jde o webové stránky, na kterých nejsou přihlášení svým loginem, nebo jestli jde o služby, kde mají svůj profil, typicky sociální sítě. Všichni účastníci výzkumu (n=10) shrnuli, že sociální sítě typu Facebook, Instagram anebo jiné nejpoužívanější služby o svých uživateliích vědí mnohem více, než běžné webové stránky. Polovina respondentů (n=5) ani množství dat nebo konkrétní informace, které sociální sítě o svých uživateliích shromažďují, nevyjmenovávala a shrnula to tak, že tyto stránky sbírají vše, co lze sledovat.

Jaroslav například řekl: „*Je otázka, jaké všechny informace o nás společnosti mají, ale říká se, že Google o nás ví skoro všechno.*“ Domnívá se také, že pravdu o tom, co vše se sbírá a co se s daty následně děje, nelze přesně určit: „*To by asi musel říct někdo, kdo dělal ten kód. A ten není úplně běžně přístupný. Myslím si, že to bude hodně spleť a že firmy se budou hodně snažit o to, abychom určité věci nevěděli.*“ Markéta v podobném duchu uvedla: „*...úplně všechno, co děláš na Facebooku, i to, co tam neděláš úplně úmyslně, se sbírá.*“ Markéta si také myslí, že se o ní sbírají data skrz mikrofon jejího mobilního telefonu. Tereza, která podle svých slov hodně používá Instagram, prohlásila: „*Myslím si, že Instagram o mně ví tak asi všechno.*“ Obdobně na otázku, co platformy, do kterých je člověk přihlášený, sbírají, Marek uvedl: „*Vlastně úplně všechno.*“ Podobně to vidí i Marie: „*No myslím si, že klidně všechno, všechno, co se dá nějak zaznamenat.*“

Ostatní respondenti (n=5) sice nepoužívali slova „vše“ nebo „všechno“, přesto i oni uvádějí, že nejpoužívanější online služby sbírají informací velmi mnoho. Petr například řekl, že sociální sítě toho o svých uživateliích zaznamenávají „*strašně moc*“ a následně své tvrzení doplnil výčtem: „*...všechno, co si tam o sobě napíšeš, o tobě Facebook ví. Druhá věc: existují různé analytické postupy, jak vytáhnout nějaké zájmy ze statusů, které děláš, lajků, které dáváš, fotografií, které nahráváš. V podstatě Facebook umí dělat sítě mezi lidmi, a to buď na základě toho, jaké tam máš přátele, ale on to kolikrát umí vytvořit i v závislosti na tom, kde olajkuješ restauraci nebo dáš sdílení polohy. On umí poznat, jací lidi jsou kolem tebe, on to umí pospojovat a na základě toho ti vyhodnocuje nějaká data.*“

Martin také vyjmenoval řadu informací, které weby sledují, aby následně doplnil, že „*...z těchto dat obecných se dají dělat užší závěry.*“ Pak také dodal, že chování uživatelů se sleduje napříč internetovými stránkami a závěry se vzájemně propojují.

Lucie se domnívá, že množství informací, které o svých uživateliích ví například Google, může být až děsivé: „*...lidé říkali, že toho o tobě ví hrozně moc, třeba to, jaké jsou*

tvaje zájmy, kolik ti je, co děláš. A to je děsivé.“ Jan zase na otázku, co vše mohou sociální sítě o svých uživateli vědět, odpověděl slovy: *„Já si myslím, že si dokážou udělat nějaký celkový obraz o tom, co nás zajímá.“* Myslí si také, že *„...teoreticky o nás mají veškeré informace, které píšeme do Messengeru, což můžou být citlivé informace.“* Klára sdělila, že sociální sítě *„...toho zpracovávají mnohonásobně víc, než si dokážeme představit.“* Dodala také, že *„...sociální sítě mají více méně celý soubor dat o daném uživateli.“*

Hodnota dat se mění

Několik respondentů (n=3) také uvedlo, že hodnota dat se může měnit v čase a v závislosti na okolnostech. Jedním z nich byl Jaroslav, který řekl: *„[data] mohou mít nějakou hodnotu teď a nějakou v budoucnu. Třeba zatím není monetizovaná informace, že někdo měl v rodině mnohem větší riziko, dejme tomu rakoviny, ale časem - pokud se budou moct tyto informace prodávat pojišťovněm, - tak najednou budou mít mnohem větší cenu.“* Dále Jaroslav řekl: *„Když se časem zjistí, že by bylo velmi lukrativní určovat polohu člověka a na základě toho třeba nabízet produkty, což už se podle mě z části určitě děje, tak se to určitě začne prosazovat víc.“*

Marie si myslí, že data, která o sobě dnes v online prostředí zanecháváme, mohou mít v budoucnu mnohem vyšší hodnotu, popřípadě do určité míry odlišnou. *„...třeba za dvacet let se můžeme dostat do situace, kdy pohovory budou vypadat úplně jinak a tohle [možnost o sobě na internetu nezanechávat data] může mít velkou hodnotu.“* Martin se domnívá, že data jednotlivých uživatelů mění hodnotu z hlediska toho, jaké je nákupní chování jednotlivce, které se s časem proměňuje. V hloubkovém rozhovoru řekl: *„...v ten moment, kdy si na základě cílené reklamy koupím auto za milion, tak najednou mají moje údaje asi cenu milionu.“*

Kroky vedoucí k zanechávání menšího množství dat

Nesouhlas s ukládáním cookies

Většina účastníků výzkumu (n=6) se snaží do určité míry využívat kroky vedoucí k tomu, aby o sobě na internetu zanechávají méně dat prostřednictvím cookies. Například Petr na internet nepřistupuje jinak než se zapnutými programy na blokování sledovacích technologií, včetně VPN a dalších. Svou digitální stopu shrnuje takto: *„Myslím si, že mám významnou digitální stopu, která ale vede někam úplně do háje... Mám placenou blokadu reklam a trackerů na všech zařízeních...“*

Oproti němu Lucie neví, že nějaké technologie proti sběru online dat existují. Zná ale cookies, které občas odmítne. „...*když je to něco, co jsem potřebovala jednou vyhledat, tak tam jim ty cookiny prostě odmítnu.*“ Jaroslav říká, že se snaží cookies odmítat, ale že není tak důsledný, jak by měl být: „*Já se snažím odmítat všechny nezbytné cookies. Samozřejmě se někdy ukliknu a dám ‚povolit vše‘.*“ Markétě naopak podle jejích slov sběr cookies až tak nevadí. Když na internet přistupuje z počítače, tak se snaží blokovat polohové služby.

Jan sebe označuje za toho „*otravného uživatele*“, který poctivě prochází nastavení cookies a zakazuje jejich sběr. Doslova říká: „*Já jsem ten otravný uživatel, který má Adblock a všechny cookies odmítá, i když je to furt klikání.*“ Specifikoval také důvod, proč si s tím dává práci: „... *cítím se pak aspoň trochu pohodlněji, že o mě nemají možná úplně všechny informace, když to odmítám.*“ Tereza cookies také omezuje, byť si nemyslí, že by to mělo na její internetovou stopu velký efekt. Řekla, že: „*Definovat je [cookies] asi neumím, ale sama si je omezují, i když si nemyslím, že je to úplně funkční.*“

Blokátory a další metody omezování digitální stopy

Zřejmě jen jediný účastník výzkumu stěžuje internetovým společnostem honbu za osobními daty využíváním blokátorů nebo prostředků sloužících k odvádění stopy od reálného uživatele, jako je třeba VPN. Jak už bylo uvedeno o několik odstavců výše, nejvíce aktivní je v tomto ohledu Petr. Za programy, které blokují anebo komplikují sběr dat, je ochoten i zaplatit, a to zhruba kolem 1000 korun ročně.

Několik respondentů (n=3) nicméně některé metody pro ochranu soukromí na internetu zná (Klára, Marek). Jaroslav je i někdy v minulosti využíval.

„Nějakou dobu, docela dlouho, jsem si platil VPNku, ale tam šlo přesně o to, že ji někde máš blokovanou - třeba v našich školních sítích je z velké části blokována. U mobilů mi to pak někdy dost zlobilo nebo snižovalo rychlost. A strašně často tě právě přes VPNku ani někam nechtějí pustit, protože se loguješ z jiného zařízení. Bohužel to komplikovalo život, ale snažil jsem se to používat, ale teď jsem si to neprodloužil. Stálo to nějakých 20, 30 dolarů na rok, což není nějaká extrémní cena...“.

Klára k vědomému ztěžování přístupu k osobním údajům volí ještě jinou strategii: je jedinou z výzkumného vzorku, kdo uvedl, že na sociální síti Twitter vystupuje pod falešným jménem a že má účet navázaný na separátní e-mailovou adresu, a to z toho důvodu, aby nebylo možné její aktivitu tak snadno spojit s její reálnou identitou. Twitter přitom dle svých

slov využívá jen ke sledování zpráv a ostatních uživatelů, sama na něj nic nepřidává. V rozhovoru řekla: „*Dříve jsem měla uživatelský Twitter a něco jsem tam i přidávala. Teď ho mám registrovaný na takovém mailu, který používám jen k přihlašování, a mám tam anonymní jméno. Jsem v tom hodně paranoidní. Podle mě to v budoucnu bude problém...*“.

Stejně tak jako Klára, se i Marie snaží svou digitální stopu a soukromí do budoucna chránit tím, že zejména na sociálních sítích omezuje produkci svého vlastního obsahu a stává se jen konzumentem obsahu ostatních. „*Já jsem třeba na Instagramu jenom sledovatel a nemám tam moc svých příspěvků. Většina mých příspěvků jsou stories, které jsou na síti jenom na chvíli, takže jsem spíše takový zevrubný stalker.*“

Důvody, proč svá data uživatelé poskytují

Ostatní to dělají také

Významná většina účastníků výzkumu (n=8) uvedla, že jedním z důvodů, proč se na sociálních sítích vyskytují, byť na nich zanechávají data, je nechuť odstříhnout se anebo jít proti zbytku společnosti. Čtyři respondenti podle svých slov nechtějí ztratit kontakt s ostatními uživateli. Například podle Petra je Facebook a Messenger služba, jejíž hodnota pro uživatele závisí na jejím rozšíření ve společnosti: „*Je to služba, která je dost užitečná. Asi neznám mladého člověka, který by nebyl k zastižení na Messengeru.*“ Jaroslav to vnímá podobně. Ztráta kontaktu s lidmi mu brání se od sociálních sítí zcela odstříhnout: „*Já to i někdy dělám, že to odstříhnu, ale je problém, že často tam máš vytvořené sociální skupiny, třeba lidi ze střední, a nechci to úplně tvořit někde jinde.*“ Také podle Markéty je komunikace ohromné plus, které stojí za to být předmětem cílené reklamy a sbírání dat. Řekla, že: „*Tak pro tebe jako uživatele je to určitě komunikační médium, pomocí kterého můžeš být v kontaktu s lidmi kdykoliv, odkudkoliv, pokud máš přístup k internetu. Takže to je určitě obrovské plus.*“ Tereza má podobný pohled na věc, jako výše zmíněný Jaroslav: „*Já se všemi kamarády komunikuji přes Instagram, spíš než přes Messenger, takže to v dnešní době nejde. V bublině je člověk už zapletený, takže vyplést se z toho by znamenalo udělat velký, složitý krok.*“

Čtyři respondenti hloubkových rozhovorů uvedli, že internet a sociální sítě i přes některé nevýhody využívají, protože tak činí ostatní. Například Lucie uvedla, že nemá tendenci se proti současnému byznys modelu nijak vymezovat, protože jí její osoba v kontextu masy lidí nepřijde důležitá: „*...mě uklidňuje to, že jsem jeden člověk z extrémně hodně, hodně, hodně, hodně lidí a že vlastně nikoho nezajímám.*“

Podle Jana se pak společnost dostala do bezvýchodné a podřadné situace. Společnost označil za otroka internetových služeb, které se prosadily v životě moderního člověka. Doslova řekl, že: „...*kdyby [užívání sociálních sítí] zpoplatnili, tak by ztratili půlku uživatelů a vliv na společnost. To otroctví, o kterém mluvím, to je taková metafora. Tím, že je to tak dobře přístupné, je to součást našich životů, na které jsme závislí. Je to součást moderního života.*“

Podle Kláry proto není možné, aby byly internetové služby zpoplatněné. Podle ní by to znamenalo, stejně jako to naznačil Jan, odliv významné části uživatelů. Klára si zároveň myslí, že lidé současné nastavení monetizace na internetu nezpochybňují, protože je k tomu společnost lhostejná: „*Mně přijde, že jsou dneska lidé lhostejní. I když si uvědomují, že společnosti zpracovávají hodně dat a že je to zneužitelné, tak je to nechává lhostejnými, protože když to dělají všichni, tak si myslí, že je to normální, že je to tak v pořádku.*“ Podobně jako Klára to vidí i Marie: „*Myslím si, že mladí lidé to řeší málo a starší na to nadávají. Myslím si, že pokrok nezastavíš.*“

Nevadí mi to

Zhruba polovině účastníků výzkumu (n=5) nevadí, že se o nich na internetu sbírají data. Vyhovuje jim, že za online obsah nemusejí přímo platit. Například Martin nemá problém s tím, že konzumování obsahu, který je na internetu volně dostupný, můžeme chápat jako obchod za uživatelská data: „...*stojí mi to za to využívat služby a konzumovat obsah za tuhle cenu, kterou si nějakým způsobem asi uvědomuji - nevím, jestli zcela ve veškerý šíři a každým okamžikem, kdy ty služby používám, ale nevadí mi to. Stojí mi za to chodit na sociální sítě a na weby s tím, že oni pak ta data můžou nějak využívat.*“ Lucie to vnímá velmi podobně. Prohlásila: „...*mně to asi nevadí, že si ta data berou.*“ Doplnila také, že využívání webů a online služeb nebere jako obchod. „*Z mého pohledu to беру tak, že to je zadarmo, tečka.*“ Poté, co se dozvěděla, kolik přibližně Facebook vydělal v roce 2020 za průměrného uživatele v Evropě, dodala, že jí to připadá jako velmi málo a že „...*v tom případě jí to je ještě víc jedno.*“ Nutno ale dodat, že v rozhovoru také uvedla, že zakazuje některé cookies a že v telefon některým aplikacím zablokovala sdílení dat.

Markétě současný stav také spíše vyhovuje. Spíše než to, že dává svá data svou přítomností na internetu k dispozici, jí vadí reklamy. Jejými slovy: „...*cena za to, že tam potom na tebe cílí tolik reklam – asi mi to nepřijde nic špatného, se přiznám.*“ Dodala také, že pokud to není na popud jejího přítele, tak se firmám nesnaží moc komplikovat extrakci jejích dat, a fakt, že o ní data společnosti sbírají, jí nijak nevadí. Přesto má údajně ve svém

počítači zablokované sdílení polohy. Pro další respondentku, Terezu, nikdy nebyl sběr dat online téma, nad kterým by uvažovala. Zároveň uvedla, že by za weby a sociální sítě ani nebyla ochotná platit, a to ani za to, že by se jí nezobrazovaly reklamy, které ji obtěžují. Stejně tak Markovi vyhovuje, pokud má přístup k online obsahu zadarmo, a to i za tu cenu, že se o něm sbírají data. Jeho slovy: „*Kromě toho, že mi firmy budou cpát svůj produkt, tak nevidím žádné velké nebezpečí v tom, že moje data o aktivitě budou ty firmy mít.*“

Nemá to smysl

Tři účastníci výzkumu vyjádřili pochybnosti nad tím, jestli mají aktivní kroky, které podnikají s cílem ztěžovat agregaci jejich dat, kýžený efekt. Jaroslav k udělování souhlasu či nesouhlasu se sběrem cookies řekl: „*Je tedy otázka, jestli oni ty informace někde mají s tím, že je nevyužijí, nebo že je stejně využijí, a člověk o tom akorát neví, ať se snaží, jak se snaží.*“ Podobnou myšlenku v souvislosti s cookies vyjádřila i Tereza, když řekla, že cookies omezuje, byť není přesvědčena o tom, že její volba má reálný dopad.

Klára uvedla, že v blokování sběru dat nevidí tolik smyslu, protože stejně o sobě vždy nějakou stopu, která vede k její osobě, zanechává: „*... stejně myslím, že přes tu nejvyšší úroveň, třeba přes IP adresu nebo e-maily, se to (se mnou) dá spojit vždycky.*“

Jsou data z pohledu uživatelů platidlo?

Raději data než peníze

Někteří účastníci výzkumu (n=3) vyslovili ochotu platit online službám za to, aby o nich nesbíraly data. Jednou z nich byl Petr, který řekl: „*...líbilo by se mi, kdyby tam byla možnost platit třeba stovku ročně za přístup k Facebooku, tak, aby tam nebyly reklamy, aby tam nebylo trasování a trackování...*“. Ochota platit plyne i z hojného užívání Messengeru: „*Messenger je hodně užitečná věc, za kterou bych 25 korun měsíčně úplně v klidu dal.*“ Druhou účastnicí, která by obětovala peníze namísto svých dat, je Marie. Ta by dle svých slov byla ochotná měsíčně platit 1000 korun za to, aby o ní data nesbíral Instagram. „*Jo, asi jo, asi bych to dala. Brala bych to jako investici,*“ dodala. Markétě současný byznys model internetových společností nevádí, na otázku, kolik by byla ochotná měsíčně platit za Facebook bez sběru dat a reklam, odpověděla, že by takovou možnost využila, pokud by to nestálo více než 300 korun měsíčně.

Jaroslav, byť má tendenci ztěžovat sběr svých osobních dat, uvedl, že není schopen posoudit, jestli je, anebo není ochoten platit za hypotetickou variantu online služeb, které by

data o svých uživatelích nesbíraly. „*No, to by bylo dobré, kdybych měl křišťálovou kouli, a dokázal vidět, jaké všechny důsledky to na mě doopravdy má,*“ uvedl a doplnil: „*Když bych si dokázal spočítat, že mě to donutí si koupit o tolik a tolik více zboží, které nepotřebuju, tak bych to dokázal nějak ohodnotit. Ale teďka to nedokážu ohodnotit.*“ Nakonec nicméně uvedl, že obecně by lidé za ochranu svého soukromí online měli být ochotní nějaké finance vynakládat: „*...myslím, že určitě by člověk měl být ochoten nějaká procenta ročně, která vydělá, do toho investovat...*“.

Pokud by existovala možnost platit za verzi online služeb, které by nesbíraly o svých uživatelích data, většina respondentů (n=6) by tuto možnost nevyužila. Pro Martina, Lucii, Terezu a Marka není sběr dat výměnou za bezplatný obsah na internetu problém, a proto by nebyli ochotní platit za ekvivalent, který by data nesbíral. Marek navíc vyslovil pochybnost nad tím, jestli by síť typu Facebook data o svých uživatelích nesbírala, i kdyby si za takovou možnost platil. Řekl: „*...i kdyby teď Facebook zavedl placené členství, tak ta data by stejně pořád sbíral. Viděli jsme v Kongresu - nebo kde to bylo - Marka Zuckerberga, jak odpovídá na otázky. Ptali se ho, jestli by teď v daném okamžiku mohl zpřístupnit data všech svých uživatelů, a vyšlo najevo, že ano. Takže já si myslím, že Facebook by ta data sbírat nepřestal. Možná by přestal ukazovat reklamu, ale nemohl by potom být tím produktem, který si platí reklamní agentury...*“.

To samé jako Marek si myslí i Klára: „*Vy byste věřil tomu, že by to Facebook nedělal? Ta data stejně budou někde v prostoru,*“ říká. Klára sice není úplně lhostejná k tomu, co o sobě na internetu zanechává za data, přesto platit přímo penězi by za obsah nebo služby, které na internetu využívá, platit nechtěla. Podobně to vnímá Jan, kterým sice sběr uživatelských dat nevyhovuje, přesto by za online obsah či služby platit penězi nechtěl. Janovi přijde „*na hlavu*“ platit si za soukromí. Řekl: „*To já nevím, jestli bych za právo na soukromí ochotný platit. Jako chápu, že kdybychom se o tom bavili z toho ekonomického hlediska, tak se o tom dá uvažovat, ale když se o tom zamyslíš obecně, tak to je úplně na hlavu, že by sis měl platit za to, abys měl soukromí. Z tohoto úhlu pohledu bych nechtěl platit nic.*“ Jan navíc posléze dodal, že pokud by se možnost platit si na internetu za soukromí zavedla, byl by to pro něj impuls přestat tyto služby využívat. Řekl: „*Ale vlastně bych byl rád, kdyby to zavedli, byl by to pro mě ještě víc bič to omezit.*“ Doplnil, že jedinou službou, u které by uvažoval o placení, je Messenger, nicméně stejně nakonec věc uzavřel tím, že by i tuto službu raději opustil.

Lidé nechtějí za online služby a obsah platit penězi

Pět z deseti účastníků výzkumu také zmínilo, že pokud by se současný způsob monetizace obsahů a služeb na internetu změnil na placený penězi, značná část návštěvníků by je přestala využívat. Podle Lucie jsou například sociální sítě zadarmo, aby přilákaly co nejvíce uživatelů. To samé si myslí Markéta, podle které jsou weby a služby zdarma, „...*aby tam bylo co nejvíce lidí*“. Dodala také, že „...*lidé, možná obecně Češi ... neradi utrácejí*.“

Jan uvedl, že kdyby byly sociální sítě zpoplatněné, „...*tak by ztratili polovinu uživatelů a vliv na společnost*“. Jaroslav konstatoval, že „*lidé si myslí, že je to zadarmo. Proto firmy dokážou sehnat masy*“. Nakonec Klára řekla, že sociální sítě by jako placené služby v masové společnosti neobstály: „...*myslím, že by to asi nefungovalo, kdyby to nebylo zadarmo, protože si neumím představit nikoho, kdo by si platil za sociální síť*.“

Platíme svým časem

V hloubkových rozhovorech s respondenty vyšlo dále najevo, že spíše než data účastníci výzkumu vnímají svůj čas jako platidlo za bezplatný online obsah. Do určité míry to zaznělo ve čtyřech rozhovorech. Dva respondenti explicitně řekli, že z jejich pohledu je čas online návštěvníků platidlem za bezplatné online obsahy. Jednou z nich je Tereza, která uvedla, že za používání bezplatných online služeb „*platíme svým časem*“. Upřesnila, že sociální sítě „...*sbírají data a pomocí toho pracují a snaží se vytvářet obsah. Tak, aby nás to zabavilo a zároveň nám pořád měli co nabídnout*“. Druhým respondentem, který odpověděl podobně, je Jaroslav: „...*tak sítě vždycky budou nabízet podobného obsahu víc, aby tam člověk trávil víc času. Díky tomu pak dokážou prodat více reklam, které monetizují. Ve zkratce je to tak, že sítě mají zájem na tom, aby tam člověk trávil co nejvíc času. Protože pro sítě čas jsou peníze*.“

Oba dva se možná i z toho důvodu také snaží čas, který na online službách tráví, omezovat, a to zejména poté, co si uvědomí, že na nich času tráví příliš mnoho. „*Někdy dělám to, že si sítě deaktivuji, třeba když je zkouškové, aby mi to nežralo čas*“, uvedla Tereza. Jarda zase řekl: „*Já bych se toho klidně asi vzdal i na nějakou dobu úplně zadarmo. S tím není problém*.“ Na otázku, kdy má tendenci svůj čas na online službách omezovat, pak odpověděl: „*Když si náhodou všimnu, že na tom trávím víc času a zbytečně. Mně se to víceméně stává jenom na Youtubu, protože ten bohužel používám hodně... tam skáče jedno video za druhým a to je zabiják času. Někdy tam sice něco zajímavého nebo vtipného vyskočí,*

ale často je to jenom hromada balastu. To klidně rád omezím. A já sám vidím, kolik času někteří lidé tráví času na Instagramu, a toho se strašně bojím.“

Stejně tak Klára i Jan mají tendenci omezovat čas, který věnuje online aktivitám. Jan na otázku, jestli měl už někdy chuť ze sociálních sítí odejít, řekl: „*Mám to střídavě. Mám ten pocit, že mi to otupuje myšlení.*“ Na začátku rozhovoru navíc také řekl, že si obsah na sociálních sítích sám omezuje a že kdyby úplně skončily, tak by to jeho život nijak negativně nepoznamenalo.

5.9 Diskuze

Ve výzkumné části se tato práce zaměřila na to, jak část pražských vysokoškoláků nahlíží na sběr osobních dat v online prostředí. Do určité míry lze najít odpovědi na všechny tři předem stanovené výzkumné otázky:

VO1: Uvědomují si lidé aktivní v online prostředí, že jejich data mají hodnotu?

Z analýzy hloubkových rozhovorů vyplývá, že všichni účastníci výzkumu si uvědomují, že se o nich v online prostředí data sbírají, a to jako prostředek, který internetovým společnostem zajišťuje finanční zisk. Všichni si uvědomují, že data se v online prostředí využívají k cílení reklam, ze kterých plyne zisk. Značná část účastníků navíc vnímá data jako prodejní artikl, se kterým firmy obchodují. Všichni si proto uvědomují, že osobní data mají hodnotu.

Specifikovat ji ale pro řadu respondentů představuje problém a musejí se uchýlit k náhodě či emocím. V první řadě málokdo o konkrétní hodnotě dat z pohledu poskytovatelů online obsahů a služeb uvažoval. V druhé řadě lidé ani nemají všeobecně dostupné informace, které by jim dovolovaly opravdovou cenu dat v daném čase znát. Zdá se, že jde o stejnou situaci, jaké docílil ve svém výzkumu Collis se svými kolegy (2022). Ti nechali také účastníky výzkumu odhadovat cenu svých dat a dostávali hodnoty založené na heuristice, což podle nich potvrzuje hypotézu, že nemají prostředky, jak zjistit, kolik data mohou stát. Byť někteří účastníci výzkumu obsaženého v této práci mají konkrétní představu o tom, kolik osobní data stojí, není ani dobře možné zhodnotit, jestli jde o představu správnou nebo mylnou.

VO2: Nakolik si lidé aktivní v online prostředí cení své soukromí?

I když si účastníci výzkumu uvědomují, že jsou o nich data na internetu agregována s vidinou zisku, naprostá většina z nich je nepovažuje jako platidlo, kterým bezplatné online obsahy nebo služby hradí. Na problematiku tedy nenahlíží tak jako van Dijck (2014), která osobní data nazývá platidlem v online světě. Za platidlo považují spíše pozornost nebo čas, jež uživatelé na internetu věnují reklamám. Zjištění, že lidé svá data nepovažují za protihodnotu vynaloženou za bezplatné konzumování online obsahu, do značné míry souhlasí se Strandburg (2013). Podle ní je zavádějící označovat data za platidlo online obsahu. Jednak zákazníci provozovatelů online obsahu a služeb jsou v praxi inzerenti, a v druhé řadě je podle Strandburg transakce data-obsah natolik nekonkrétní a skrytá, že není důvod si myslet, že by uživatelé data jako platidlo chápali. Mimochodem, reklama na internetu s výjimkami všechny účastníky výzkumu obsaženého v této práci obtěžuje. Někteří z nich se proti ní i brání programy a pluginy typu Adblock. Menšina je také ochotná se reklamy zbavovat za peníze, což potvrzují i jiné průzkumy – například An (2016) shledala, že vůle platit je malá i mezi lidmi, které online inzerce obtěžuje natolik, že využívají blokátory reklam. Za předpokladu, že by jejich blokátory přestaly proti reklamám fungovat, by za obsah bez reklam zaplatilo jen zhruba 15 %.

Ochota platit za ochranu svých dat je ale zřejmě ještě nižší: pokud by existovala možnost placené varianty nepoužívanějších online služeb, většina účastníků výzkumu by ji nevyužila. Nutno dodat, že jde o odpovědi na hypotetickou situaci, kterou dle slov respondentů provází celá řada neznámých, včetně například toho, jestli by se technologickým společenstvem dalo věřit, že o uživatelích data sbírat nebudou, i když by si to zaplatili. Nicméně toto zjištění, že ochota platit za nesbírání dat je nízká, je v souladu s již uvedenými závěry Jentzsch a kolektivu (2012), Preibusch a kolektivu (2013) či Beresford a kolektivu (2012).

Co je ale z pohledu soukromí podstatnější, je spokojenost účastníků výzkumu s aktuálním nastavením monetizace bezplatných online služeb a sběrem dat. Zhruba polovina jich je spokojená nebo jí dosavadní nastavení včetně toho, že se o nich sbírají data, nevadí. Benefity převažují, a to i pro ty, kterým sběr dat úplně nevyhovuje. To, že online obsah a služby přinášejí uživatelům *ohromný* přínos, je podle existujících poznatků podle Strandburg (2013) nezpochybnitelné. To naznačuje mimo jiné i experiment Brynjolfssona a kolektivu (2019), ze kterého vyšlo najevo, že za nepoužívání online služeb by si lidé přáli obdržet nemalé peněžní kompenzace.

Navíc, jak vyplynulo z řady hloubkových rozhovorů, zásadním důvodem, proč lidé nepoužívanější online služby využívají i přesto, že zasahují do jejich soukromí, je to, že nechtějí zprerhávat kontakt se svými blízkými. Někteří respondenti se nechali slyšet, že jde o model, na který masová společnost přistoupila a ze kterého je pro jedince těžké vystoupit. Na to, že v tomto ohledu lidé nemají příliš na výběr, upozornil například Crain (2016). To, že jsou lidé ohledně míry soukromí ochotní se podřídít chování okolí, ukázal ve svém jiném výzkumu také Acquisti (2012).

Někteří soukromí na internetu řeší spíše ze společenské roviny – tedy do jaké míry své soukromí otevírají ne firmám a provozovatelům online služeb, ale také ostatním uživatelům. Několik respondentů již nechce o sobě ostatním uživatelům o sobě nic prozrazovat, ale síť používá hlavně ke sledování ostatních nebo k vzájemné komunikaci. Jde o praktiku, kterou už v roce 2013 popsal Vitak & Ellison jako self-censorship, což můžeme do češtiny volně přeložit jako sebecenzuru. Podle Acquistiho a kolektivu (2020) toto chování lze chápat také jako jednu z metod, jak uživatelé online služeb firmám ztěžují extrakci osobních dat, což souvisí s následující výzkumnou otázkou.

VO3: Podnikají lidé nějaké aktivní kroky k tomu, aby o sobě zanechávali méně dat?

Naprostá většina účastníků výzkumu činí alespoň nějaké aktivní kroky pro to, aby o sobě provozovatelům webů a online služeb poskytovalo méně dat. Paradoxně jsou mezi nimi i ti, kterým sběr dat údajně tolik nevádí. To je nicméně v souladu s teorií Malgieri a Custers (2018), podle nichž ztěžování sběru dat souvisí s uvědoměním, že mají data hodnotu. A to si uvědomuje každý participant výzkumu. Jak bylo uvedeno v analýze výzkumu této práce, většina výzkumného vzorku upravuje sběr informací, které se zahrnují či by se daly zahrnout mezi soubory cookie. Nad jejich rámec sběr dat pro internetové společnosti komplikuje jen pár dotazovaných. Odpovědí na tuto výzkumnou otázku tedy je, že v případě pražských vysokoškoláků, na které se výzkum této práce zaměřil, většina činí nějaké kroky s cílem omezit přístup k jejich datům.

Liší se ale množství kroků, které respondenti proti extrakci dat podnikají. Při porovnávání výsledků analýzy si můžeme pokládat otázku, proč někteří dotazovaní podnikají kroků pro ochranu svých dat více či méně než ostatní. Nutno dodat, že ve výzkumném vzorku převládali ti, kteří projeví jen omezenou ochotu svá data chránit.

Respondenti zmínili celou řadu důvodů, proč nemají motivaci komplikovat extrakci jejich dat. Zaznělo, že sofistikovanější blokace trackerů představují snížení uživatelské

přívětivosti online obsahů a služeb. To, že pohyb na internetu mohou ztěžovat natolik, že člověk některá data společnostem raději předá, uvádí například Acquisti s kolektivem (2020). Důvodem může být také nevíra v to, jestli má vůbec smysl se snažit svá data chránit. Že jsou tyto pochybnosti namístě, potvrzuje například Sarat (2014). Častým motivem, který v hloubkových rozhovorech zazněl, je ale prostý fakt, že jim sběr dat nevadí natolik, aby využívali možnosti ho omezovat. To dokázal například Johnson (2020). Jak již bylo několikrát uvedeno, i ti nejsnaživější ochránci svého soukromí využívají některé online služby, které narušují jejich soukromí, protože chtějí zůstat součástí kolektivu. Na to, že jde o zásadní argument, upozornil například Crain (2016) či Acquisti (2012).

Závěr

Od okamžiku, kdy první technologické firmy (a nejen firmy) začaly ve velkém data o návštěvnicích internetu či uživatelů svých online služeb sbírat, uběhla přibližně dvě desetiletí. Online technologie, které si díky svým nesporným přínosům vysloužily prominentní místo ve fungování masové společnosti, v poslední době častěji ukazují svou stinnou tvář, ať už jsou to fake news, polarizace společnosti, uzavírání se do názorových bublin, negativní dopady na psychický stav lidí, či narušování jejich soukromí, například v podobě extrakce jejich dat. Společným jmenovatelem všech těchto jevů je podle řady odborníků byznys model, který ovládl online svět, a tím je cílení reklam na uživatele na základě znalostí, které o nich firmy sbírají.

Cílem této diplomové práce bylo zmapovat trh s osobními daty, které se na internetu sbírají. Jinými slovy, přinést co nejcelistvější obraz o tom, jak trh s daty funguje, jak se v čase vyvíjel a nakonec jakou mají data hodnotu z pohledu soukromých společností i samotných uživatelů internetu. K lepšímu porozumění toho, jakou hodnotu lidé aktivní v online prostředí svým datům přikládají, slouží kvalitativní výzkum zaměřený na generaci pražských vysokoškolských studentů, který je součástí této práce.

Z teoretické i praktické části práce vychází najevo, že kontroverze týkající se sběru uživatelských dat na webu pro technologické firmy dosud nejsou hrozbou. Byť si lidé uvědomují, že jsou jejich údaje na internetu sbírány a jejich chování sledováno, většina to vnímá jako normu, se kterou souhlasí. Těm, kterým výměna dat za bezplatný online obsah a služby nevyhovuje, tyto služby stejně nadále využívá, protože dle jejich hodnotového soudu dosud nejde o něco, čemu by jim stálo za to se úplně vyhnout. Výhody, které online služby mají, stále převládají nad nevýhodou v podobě sběru uživatelských dat.

Na základě dostupných informací lze dokonce předpokládat, že v blízké budoucnosti o sobě budou lidé firmám prozrazovat informací ještě více. I přes přísnější regulace, kritiku ze stran běžné i odborné veřejnosti mezi firmami stále převládá snaha zavádět ještě efektivnější metody sběru dat a také rozšiřovat škálu služeb a produktů, které jsou uživatele schopné sledovat.

Navíc i na základě výzkumu obsaženého v této práci je nanejvýš pravděpodobné, že lidé nemají úplné či bezchybné znalosti o tom, jak sběr dat na internetu funguje, a proto mohou na problematiku nahlížet méně kriticky, než kdyby si vše uvědomovali do důsledků. Právě nízké povědomí o tom, co se s daty děje, se zdá být jedním z pilířů, na kterém se technologickým firmám podařilo tento byznys model postavit.

To, že některé informace, které se týkají sběru uživatelských dat, i přes vynaloženou snahu nelze běžně dostupnými prostředky získat, potvrzuje i teoretická část této práce, které se sice daří řadu procesů a témat vyložit a objasnit, některé ale nadále zůstávají nejasné, a to zejména kvůli samotným účastníkům trhu s daty, kteří tyto údaje nesdílejí, pokud nemusejí. Například co všechno firmy o uživateli sledují a k čemu všemu data využívají. To jsou velmi závažné otázky, jejichž zodpovězení by mohlo dosavadní smířlivý přístup veřejnosti ke sběru dat v online prostředí ovlivnit. Nutno dodat, že zřejmě v neprospěch technologických firem, neboť z oněch zhruba 20 let jejich existence už víme, že pokud se odpovédím na některé otázky vyhýbají nebo pokud na ně odpovídají neurčitě, obávají se negativní reakce veřejnosti či regulátorů.

Jak totiž kromě zmíněných existujících zjištění plyne i z výše provedeného kvalitativního průzkumu zaměřeného na české jedince, jednotliví uživatelé by své dosavadní chování měli vůli zásadněji upravovat až na základě nálady svého okolí. Další výzkum by se proto mohl pečlivěji zaměřit na to, co jednotlivé společnosti, potažmo masová společnost, vnímala jako dostatečný impuls pro hledání alternativ za byznysový model, který dnešnímu online světu dominuje. Přitom nemusí být vyloučené, že největší technologické společnosti na tuto otázku odpovědi mají, protože je v jejich zájmu a pravděpodobně i schopnostech být o krok napřed.

Jako cenný kontext může posloužit odpověď na otázku, kdy si lidé začali uvědomovat, že jejich data mají pro internetové společnosti hodnotu – zda-li si to uvědomovali od prvního okamžiku, kdy internet začali využívat, či jestli se o tom dozvěděli až později, případně v jaké souvislosti. Výzkum obsažený v této práci s touto rovinou nepracoval. Navazující kvalitativní výzkumy na téma soukromí či sběru osobních dat by se mohly zaměřit na jednotlivé online služby, případně jiné užší sociodemografické skupiny.

Seznam obrázků

Obrázek č. 1: Rozdělení dat podle zprávy OECD (2013) na základě dat z WEF (2011)...str. 20

Seznam tabulek

Tabulka č. 1: Věk, odhad stráveného času na internetu, navštěvovaná univerzita a fakulta participantů hloubkových rozhovorů.....str. 41

Tabulka č. 2: Kategorie plynoucí z analýzy hloubkových rozhovorů.....str. 46

Seznam příloh

Příloha č. 1: Návod k hloubkovým rozhovorům

Příloha č. 2: Informovaný souhlas účastníka výzkumu

Příloha č. 3: Hloubkové rozhovory

Zdroje

Acquisti, A., John, L. & Loewenstein, G. (2009). What is Privacy Worth?, *The Journal of Legal Studies*, 42(2), str. 249-274.

Acquisti, A., John, L. K., & Loewenstein, G. (2012). The Impact of Relative Standards on the Propensity to Disclose, *Journal of Marketing Research*, 49(2), str. 160–174.

Acquisti, A., Loewenstein, G., Brandimarte, L. (2020). Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving It in the Digital Age, *Journal of Consumer Psychology*, 30(4), str. 736-758.

An, M. (2016). Why People Block Ads and What It Means for Marketers and Advertisers. *Hubspot Research*. Dostupné z: <https://www.upa.it/static/upload/why/why-people-block-ads/why-people-block-ads.pdf>

Angwin, J. & Valentino-Devries, J. (29. září 2012). New Tracking Frontier: Your License Plates. *The Wall Street Journal*. Dostupné z: <https://www.wsj.com/articles/SB10000872396390443995604578004723603576296>.

Amazon. (2022). When is Alexa Listening?. *Amazon.com*. Dostupné z: <https://www.amazon.com/is-alexa-always-listening/b?ie=UTF8&node=21137869011>.

Astor, M. (25. července 2017). Your Roomba May Be Mapping Your Home, Collecting Data That Could Be Shared. *The New York Times*. Dostupné z: <https://www.nytimes.com/2017/07/25/technology/roomba-irobot-data-privacy.html>.

Australian. (2017). Facebook Targets ‘Insecure’ Kids. *Theaustralian.com.au*. Dostupné z: <http://www.theaustralian.com.au/business/media/digital/facebook-targets-insecure-young-people-to-sell-ads/news-story/a89949ad016eee7d7a61c3c30c909fa6>.

Bajak, F. (20. června 2018). Mobile Carriers Cut Off Flow of Location Data to Brokers. *AP News*. Dostupné z: <https://apnews.com/article/north-america-technology-business-ron-wyden-california-8582857aff8146f8ac81d247533b2177>.

BBC. (25. června 2021). Google Tracking Cookies Ban Delayed Until 2023. *BBC.com*. Dostupné z: <https://www.bbc.com/news/technology-57611701>.

Beresford, A. R., Kučbler, D., & Preibusch, S. (2012). Unwillingness to Pay for Privacy: A Field Experiment, *Economics Letters*, 117(1), str. 25–27.

Birch, K., Cochrane, D. T. & Ward, C. (2021). Data as Asset? The Measurement,

Governance, and Valuation of Digital Personal Data by Big Tech, *Big Data & Society*, 8(1), str. 1-15.

Brin, S. & Page, L. (1998). The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems* 3. Dostupné z: <http://infolab.stanford.edu/~backrub/google.html>.

Brynjolfsson, E., Collins, A. & Eggers, F. (2019). Using massive online choice experiments to measure changes in well-being, *Proceedings of the National Academy of Sciences*, 116(15), str. 7250-7255.

Burdon, M. & McKillop, A. (2015). The Google Street View Wi-Fi Scandal and Its Repercussions for Privacy Regulation, *Monash University Law Review*, 39(3), str. 702-738.

Busby, E., Hammoud, T., Rose, J. & Prashad, R. (10. ledna 2012). The Evolution of Online-User Data. *BCG.com*. Dostupné z: <https://www.bcg.com/publications/2012/marketing-technology-evolution-of-online-user-data>.

Canellopoulou-Bottis, M. & Bouchagiar, G. (2018). Personal Data v. Big Data: Challenges of Commodification of Personal Data, *Open Journal of Philosophy*, 08(3), str. 206-215.

CNIL. (6. ledna 2022). Cookies: The Cnil Fines Google A Total of 150 Million Euros and Facebook 60 Million Euros for Non-Compliance with French Legislation. *CNIL*. Dostupné z: <https://www.cnil.fr/en/cookies-cnil-fines-google-total-150-million-euros-and-facebook-60-million-euros-non-compliance>.

Collis, A., Moehring, A., Sen, A. & Acquisti, A. (2022). Information Frictions and Heterogeneity in Valuations of Personal Data. *SSRN*. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3974826.

Crain, M. (2016). The Limits of Transparency: Data Brokers and Commodification, *New Media & Society*, 20(1), str. 88-104.

Crain, M. (2021). *Profit Over Privacy: How Surveillance Advertising Conquered the Internet*. University of Minnesota Press.

Creswell, J. W. (2009). *Research design: Qualitative, Quantitative, and Mixed Method Approaches*. London: Sage.

Criss, D. (8. prosince 2016). These Dolls Are Spying on Your Kids, Consumer Groups Say. *CNN.com*. Dostupné z: <https://edition.cnn.com/2016/12/08/health/cayla-ique-ftc-complaint-trnd/index.html>.

Cyphers, B. & Doctorow, C. (12. února 2021). Privacy Without Monopoly: Data Protection and Interoperability. *Electronic Frontier Foundation (EFF)*. Dostupné z: <https://www.eff.org/wp/interoperability-and-privacy>.

Cyphers, B. (10. března 2021). App Stores Have Kicked Out Some Location Data Brokers. Good, Now Kick Them All Out. *Electronic Frontier Foundation (EFF)*. Dostupné z: <https://www.eff.org/deeplinks/2021/03/apple-and-google-kicked-two-location-data-brokers-out-their-app-stores-good-now>

Daniel, J. E. (21. července 2021). The DotCom Bubble. *Medium*. Dostupné z: <https://medium.datadriveninvestor.com/the-dotcom-bubble-d33d429addcc>.

Desjardins, J. (29. března 2019). How the Tech Giants Make Their Billions. *Visual Capitalist*. Dostupné z: <https://www.visualcapitalist.com/how-tech-giants-make-billions/>.

Doctorow, C. (17. června 2022). Facebook Is Right, Apple Is Too Powerful. *Pluralistic*. Dostupné z: <https://pluralistic.net/2022/06/17/no-gods-no-masters/#oh-the-irony>.

DuckDuckGo. (5. března 2020). DuckDuckGo Tracker Radar Exposes Hidden Tracking. *DuckDuckGo*. Dostupné z: <https://spreadprivacy.com/duckduckgo-tracker-radar/>.

The Economist. (6. května 2017). The World's Most Valuable Resource Is No Longer Oil, But Data. *The Economist*. Dostupné z: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.

Electronic Frontier Foundation. (2022). Cover Your Tracks: See How Trackers View Your Browser. *Eff.org*. Dostupné z: <https://coveryourtracks.eff.org/learn>.

Elkind, P., Gillum, J. & Silverman, C. (7. září 2021). How Facebook Undermines Privacy Protections for Its 2 Billion WhatsApp Users. *ProPublica*. Dostupné z: <https://www.propublica.org/article/how-facebook-undermines-privacy-protections-for-its-2-billion-whatsapp-users>.

Elsaify, M. & Hassan, S. (5. dubna 2020). Some Data on the Market for Data. *SSRN*. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3568817.

Facebook. (6. listopadu 2007). Facebook Unveils Facebook Ads. *Meta*. Dostupné z: <https://about.fb.com/news/2007/11/facebook-unveils-facebook-ads/>

Facebook. (2014). The Formation of Love. *Meta*. Dostupné z: <https://www.facebook.com/notes/10158928005273415/>

Facebook. (2020). *FB Earnings Presentation Q4 2020*. Dostupné z: https://s21.q4cdn.com/399680738/files/doc_financials/2020/q4/FB-Earnings-Presentation-Q4-2020.pdf.

Facebook. (27. ledna 2021). Facebook Reports Fourth Quarter and Full Year 2020 Results. *Meta*. Dostupné z: <https://investor.fb.com/investor-news/press-release-details/2021/Facebook-Reports-Fourth-Quarter-and-Full-Year-2020-Results/default.aspx>.

Facebook. (2022a). Facebook pixel. *Facebook.com*. Dostupné z: <https://cs-cz.facebook.com/business/learn/facebook-ads-pixel>.

Facebook. (2022b). Zásady ochrany osobních údajů. *Facebook.com*. Dostupné z: https://www.facebook.com/privacy/policy/?entry_point=data_policy_redirect&entry=0.

Feiner, L. (3. února 2022). Facebook Stock Plummets 26% in Its Biggest One-day Drop Ever. *CNBC.com*. Dostupné z: <https://www.cnbc.com/2022/02/03/facebook-shares-plummet-22percent-after-reporting-weak-guidance.html>.

Fourcade, M. & Healy, K. (2017). Seeing Like a Market, *Socio-Economic Review*, 15(1), str. 9-29.

Goel, V. (24. června 2021). An Updated Timeline for Privacy Sandbox Milestones. *Blog.google*. Dostupné z: <https://blog.google/products/chrome/updated-timeline-privacy-sandbox-milestones/>.

Google. (18. srpna 2004). Prospectus. *Sec.gov*. Dostupné z: <https://www.sec.gov/Archives/edgar/data/1288776/000119312504143377/d424b4.htm>.

Google. (31. prosince 2005). Form 10-K For the Fiscal Year Ended December 31, 2005. *Sec.gov*. Dostupné z: <https://www.sec.gov/Archives/edgar/data/0001288776/000119312506056598/d10k.htm>.

Google. (2022). Find & improve your location's accuracy. *Support.google.com*. Dostupné z: <https://support.google.com/maps/answer/2839911?hl=en&co=GENIE.Platform%3DAndroid#zippy=%2Cturn-on-wi-fi%2Corient-yourself-with-live-view%2Cimprove-accuracy-of-your-location-with-live-view>.

Glikman, P. & Glady, N. (14. října 2015). What's The Value Of Your Data?. *TechCrunch*. Dostupné z: <https://techcrunch.com/2015/10/13/whats-the-value-of-your-data/>.

Graham, M. & Elias, J. (18. května 2021). How Google's \$150 Billion Advertising Business Works. *CNBC.com*. Dostupné z:

<https://www.cnbc.com/2021/05/18/how-does-google-make-money-advertising-business-breakdown-.html>.

Gray, A. (6. prosince 2021). Three tech giants control half of advertising outside China. *Financial Times*. Dostupné z: <https://www.ft.com/content/bcbc8674-060f-4298-aab8-91e40e00c3f2>

Hagey, K. (29. května 2019) Behavioral Ad Targeting Not Paying Off for Publishers, Study Suggests. *The Wall Street Journal*. Dostupné z: <https://www.wsj.com/articles/behavioral-ad-targeting-not-paying-off-for-publishers-study-suggests-11559167195>

Hendl, J. (2005). *Kvalitativní výzkum: Základní metody a aplikace*. Praha: Portál.

Hild, M. (2008). The Google IPO, *Journal of Business & Technology Law*, 3(1), str. 41-59.

Houser, F. (16. srpna 2021). Android 12 překope stávající praxi pro určování polohy. *Svět Androida*. Dostupné z: <https://www.svetandroida.cz/android-12-urcovani-polohy-presna-priblizna/>.

Ignoffo, Z. & Zoltan, M. (15. června 2022). Dark Web Price Index 2021. *Privacy Affairs*. Dostupné z: <https://www.privacyaffairs.com/dark-web-price-index-2021/>.

International Association of Privacy Professionals. (2022). Web Beacon. *Iapp.org*. Dostupné z: <https://iapp.org/resources/article/web-beacon/>.

Jenkins, R. (4. dubna 2012). How Much Is Your Email Address Worth?. *The Drum*. Dostupné z: <https://www.thedrum.com/opinion/2012/04/04/how-much-your-email-address-worth>.

Jentzsch, N., Preibusch, S. Harasser, A., Ikonomou, D. & Tirtea, R. (2012). Study On Monetising Privacy. An Economic Model for Pricing Personal Information. *ENISA*. Dostupné z: <https://www.enisa.europa.eu/publications/monetising-privacy>.

Jindrová, T. (11. srpna 2021). Průzkum: Češi jsou opatrní se sdílením soukromých dat. Ochota klesá s věkem. *Seznam zprávy*. Dostupné z: <https://www.seznamzpravy.cz/clanek/pruzkum-cesi-jsou-opatrní-se-sdilenim-soukromych-dat-ochota-klesa-s-vekem-171688>.

Johnson, G. A., Shriver, S. K., & Du, S. (2020). Consumer Privacy Choice in Online Advertising: Who Optes Out and At What Cost to Industry?, *Marketing Science*, 39(1), str. 33– 51.

JoinDeleteMe. (2022). Privacy Is Priceless. Stop Data Brokers from Selling Your Data. *JoinDeleteMe*. Dostupné z: <https://joindeleteme.com/privacy-protection-plans/>.

Kiss, J. (15. května 2010). Google Admits Collecting Wi-Fi Data Through Street View Cars. *The Guardian*. Dostupné z: <https://www.theguardian.com/technology/2010/may/15/google-admits-storing-private-data>.

Kosinski, M., Stillwell, D. & Graepel, T. (2013). Private Traits and Attributes Are Predictable from Digital Records of Human Behavior, *Proceedings of the National Academy of Sciences*, 110(15), str. 5802-5805.

Kluiters, L., Srivastava, M. & Tyll, L. (2022). The Impact of Digital Trust on Firm Value and Governance: An Empirical Investigation of US firms, *Society and Business Review*. Dostupné z: <https://www.emerald.com/insight/content/doi/10.1108/SBR-07-2021-0119/full/html>.

Krasnoff, B. (11. června 2021). How to Stop your Emails from Being Tracked. *The Verge*. Dostupné z: <https://www.theverge.com/22288190/email-pixel-trackers-how-to-stop-images-automatic-download>.

Kramer, A. D. I., Guillory, J. E., Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks, *Proceedings of the National Academy of Sciences*, 111(24), str. 8788-8790.

Kreč, L. (28. dubna 2022). Twitteru letos výrazně stoupl počet uživatelů, nahoru šel i zisk. Zaostal ovšem v tržbách. *CzechCrunch*. Dostupné z: <https://cc.cz/twitteru-letos-vyrazne-stoupl-pocet-uzivatelu-nahoru-sel-i-zisk-zaostal-ovsem-v-trzbach/>

Lambrecht, A., Tucker, C. (9. března 2018). Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2852260.

Lane, J. Stodden, V. Bender, S. & Nissenbaum, H. (Eds.). (2014). *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. Cambridge: Cambridge University Press.

Levin, A. (5. prosince 2019). Selling Privacy: The Next Big Thing for Entrepreneurs. *INC.com*. Dostupné z: <https://www.inc.com/adam-levin/selling-privacy-next-big-thing-for-entrepreneurs.html>.

Levin, S. (1. května 2017). Facebook Told Advertisers It Can Identify Teens Feeling 'Insecure' and 'Worthless'. *The Guardian*. Dostupné z: <https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens>.

Libert, T. (2015). Exposing the Hidden Web: An Analysis of Third-Party HTTP Requests on One Million Websites, *International Journal of Communication*, 9, str. 3544–3561.

Lomas, N. (31. května 2021). Europe's cookie consent reckoning is coming. *TechCrunch*.

Dostupné z:

<https://techcrunch.com/2021/05/30/europes-cookie-consent-reckoning-is-coming/>.

Loshin, D. (2013). *Business Intelligence: The Savvy Manager's Guide*. Burlington: Morgan Kaufmann Publishers.

Louis, T. (31. srpna 2013). How Much Is A User Worth?. *Forbes.com*. Dostupné z:

<https://www.forbes.com/sites/tristanlouis/2013/08/31/how-much-is-a-user-worth/?sh=78f12ad1c514>.

Malgieri, G. & Custers, B. (2018). Pricing Privacy – The Right to Know the Value of Your Personal Data, *Computer Law & Security Review*, 34, str. 289-303.

Marvin, G. (23. listopadu 2015). After A Year Of Transition, Microsoft Execs Say, “We’re All In On Search”. *Search Engine Land*. Dostupné z:

<https://searchengineland.com/microsoft-execs-all-in-on-search-bing-ads-next-236746>.

MediaGuru. (12. dubna 2017). Nejrozšířenější aplikací u Čechů je FB Messenger.

MediaGuru. Dostupné z:

<https://www.mediaguru.cz/clanky/2017/04/nejrozsi-renejsi-aplikaci-u-cechu-je-fb-messenger/>.

Meta. (28. října 2021). Introducing Meta: A Social Technology Company. *Meta*. Dostupné z:

<https://about.fb.com/news/2021/10/facebook-company-is-now-meta/>.

Microsoft. (13. června 2016). Microsoft to acquire LinkedIn. *News.Microsoft.com*. Dostupné z:

<https://news.microsoft.com/2016/06/13/microsoft-to-acquire-linkedin/>.

Morozov, E. (2011). *The Net Delusion: The Dark Side of Internet Freedom*. New York: PublicAffairs.

Mozilla. (2022). Facebook Pixel Hunt. *Rally.mozilla.org*. Dostupné z:

<https://rally.mozilla.org/current-studies/facebook-pixel-hunt/>.

Muir, J. A. & van Oorschot, P. C. (2006). Internet Geolocation and Evasion, *School of Computer Science*. Dostupné z:

<https://www.ccsf.carleton.ca/paper-archive/muir-computingsurveys-09.pdf>.

Narayanan, A. & Shmatikov, V. (2010). Privacy and Security Myths and Fallacies of “Personally identifiable information,” *Communications of the ACM*, 53(6).

Nařízení 2016/ 679. (26. dubna 2016). Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů), *Úřední věstník Evropské unie*. Dostupné z: <https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=CELEX:32016R0679&from=EN>.

Nouwens, M., Liccardi, I., Veale, M. Karger, D. & Kagal, L. (2020). Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence, *CHI '20 CHI Conference on Human Factors in Computing Systems*. Dostupné z: <https://arxiv.org/pdf/2001.02479.pdf>.

OECD. (2. dubna 2013). Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value, *OECD Digital Economy Papers*, 220. Paříž: OECD Publishing. Dostupné z: <https://doi.org/10.1787/5k486qtxldmq-en>.

Olejnik, S. & Tran, M. & Castelluccia, C. (2013). Selling Off Privacy at Auction, *Internet Society*. Dostupné z: <https://hal.inria.fr/hal-00915249/PDF/SellingOffPrivacyAtAuction.pdf>.

Olsen, S. (27. května 2002). Overture Sues Google over Search Patent. *CNET*. Dostupné z: <https://www.cnet.com/tech/services-and-software/overture-sues-google-over-search-patent/>.

Oltermann, P. (17. února 2017). German Parents Told to Destroy Doll That Can Spy on Children. *The Guardian*. Dostupné z: <https://www.theguardian.com/world/2017/feb/17/german-parents-told-to-destroy-my-friend-ayla-doll-spy-on-children>.

Preibusch, S., Kübler, D. & Beresford, A. (2013). Price versus privacy: An experiment into the competitive advantage of collecting less personal information, *Electronic Commerce Research*, 13, str. 423-455.

PrivacyRights.org. (2022). Data Brokers. *PrivacyRights.org*. Dostupné z: <https://privacyrights.org/data-brokers?terms=&page=1>.

Quartz. (2022). The End of Third-Party Cookies. *Quartz*. Dostupné z: <https://qz.com/guide/the-end-of-third-party-cookies/>.

Rivero, N. (25. dubna 2021). The inventor of the digital cookie has some regrets. *Quartz*. Dostupné z: <https://qz.com/2000350/the-inventor-of-the-digital-cookie-has-some-regrets/>.

Samuelson, P. (2000). Privacy As Intellectual Property?, *Stanford Law Review*, 52, str. 1125-1173.

Sadowski, J. (2019). When Data Is Capital: Datafication, Accumulation, and Extraction. *Big Data & Society*. 6(1). Dostupné z: <https://doi.org/10.1177/2053951718820549>.

Sarat, A. (Ed.). (2014). *A World without Privacy: What Law Can and Should Do?* Cambridge: Cambridge University Press.

Savage, S. & Waldman, D. M. (2013). The Value of Online Privacy, *SSRN Electronic Journal*. Dostupné z: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2341311.

Schneier, B. (2013). The Internet Is a Surveillance State. *Schneier on security*. Dostupné z: https://www.schneier.com/essays/archives/2013/03/the_internet_is_a_su.html

Sebera, M. (2012). *Vybrané kapitoly z metodologie*. Brno: Masarykova Univerzita.

Seligman, T. J. & Taylor, J. D. (19. června 2000). FTC Reverses Privacy Policy, *New York Law Journal*.

Sen, R. (19. května 2021). Here's How Much Your Stolen Personal Data Is Worth on the Dark Web. *The Next Web*. Dostupné z: <https://thenextweb.com/news/how-much-your-stolen-personal-data-is-worth-on-the-dark-web-syndication>.

Seznam. (2021). Zprávy ze Seznamu: Přihlášení přes Seznam má zelenou. *Seznam Identita*.

Siemens. (5. září 2014). *Siemens Smart Data*. [Video]. Youtube. Dostupné z: <https://www.youtube.com/watch?v=ZxoO-DvHQRw>.

Singleton, M. (19. června 2018). Verizon will stop selling real-time location data to third-party brokers. *The Verge*. Dostupné z: <https://www.theverge.com/2018/6/19/17478934/verizon-selling-real-time-location-data-third-party-securus-wyden>.

Slížek, D. (23. října 2019). Unikátní rozsudek: Mall.cz musí zaplatit uživateli odškodnění za únik hesla. *Lupa.cz*. Dostupné z: <https://www.lupa.cz/clanky/unikatni-rozsudek-mall-cz-musi-zaplatit-uzivateli-odskodneni-za-unik-hesla/>.

Slížek, D. (16. října 2021). *Magazín Online Plus se zamýšlí nad tím, jak budou zadavatelé cílit reklamu na spotřebitele po zrušení takzvaných cookies*. [Rádiové vysílání]. Rozhlas Plus. Dostupné z: https://plus.rozhlas.cz/cro_soundmanager/files/8599009/field_main_audio.

Statista (únor 2022). *Annual Revenue of Google from 2002 to 2021*. Dostupné z: <https://www.statista.com/statistics/266206/googles-annual-global-revenue/>.

Slížek, D. & Sklácel, O. (14. ledna 2022). Mozilla zjistí, co o nás sbírá Facebook. „Objevujeme, jak je internet prošmírovaný,“ hodnotí Koubský. [Rádiové vysílání]. Rozhlas Plus. Dostupné z: <https://plus.rozhlas.cz/mozilla-zjisti-co-o-nas-sbira-facebook-objevujeme-jak-je-internet-prosmirovany-8660088>.

Smith, R. M. (11. listopadu 1999). The Web Bug FAQ. *Electronic Frontier Foundation*. Dostupné z: http://web.archive.org/web/20010729060646/www.eff.org/Privacy/Marketing/web_bug.html.

Statcounter. (2022). Browser Market Share Worldwide: June 2021 - June 2022. *Statcounter*. Dostupné z: <https://gs.statcounter.com/browser-market-share>.

Statt, N. (19. června 2018). AT&T and Sprint to follow Verizon in ending its sale of user location data to third-party brokers. *The Verge*. Dostupné z: <https://www.theverge.com/2018/6/19/17479490/att-follows-verizon-user-location-data-sale-brokers>.

Statt, N. (4. prosince 2019). The Rise, Disappearance, and Retirement of Google Co-Founders Larry Page and Sergey Brin. *The Verge*. Dostupné z: <https://www.theverge.com/2019/12/4/20994361/google-alphabet-larry-page-sergey-brin-sundar-pichai-co-founders-ceo-timeline>.

Statt, N. (24. března 2020). Apple updates Safari's anti-tracking tech with full third-party cookie blocking. *The Verge*. Dostupné z: <https://www.theverge.com/2020/3/24/21192830/apple-safari-intelligent-tracking-privacy-full-third-party-cookie-blocking>.

Steel, E., Locke, C., Cadman, E. & Freese, B. (12. června 2013). How Much Is Your Personal Data Worth? *Financial Times*. Dostupné z: http://www.ft.com/cms/s/2/927ca86e-d29b-11e2-88ed-00144feab7de.html?ft_site=falcon.

Story, L. & Stone, B. (30. listopadu 2007). Facebook Retreats on Online Tracking. *The New York Times*. Dostupné z: <https://www.nytimes.com/2007/11/30/technology/30face.html>.

Taylor, J., Devlin, J. & Curran, K. (2012). Bringing Location to IP Adresses with IP Geolocation, *The Journal of Emerging Technologies in Web Intelligence*, 4(3).

TED. (23. května 2017). How Data Brokers Sold My Identity | Madhumita Murgia [Video]. YouTube. Dostupné z: <https://www.youtube.com/watch?v=AU66C6HePfg>.

Trhoň, O., Hyclová, B. & Šlahaň, D. (2017). Personality Price. Dostupné z: <http://www.ondrejtrhon.cz/personalityprice/vysledky.html>.

Tsai, J., Egelman, S., Cranon, L. & Acquisti, A. (2011). The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study, *Information Systems Research*, 22, str. 254-268.

Tsotsis, A. (2. února 2012). Facebook's IPO: An End To All The Revenue Speculation. *TechCrunch*. Dostupné z: <https://techcrunch.com/2012/02/01/facebook-ipo-facebook-ipo-facebook-ipo/?guccounter=1>.

Tweedie, S. (13. červen 2016). Microsoft Buys LinkedIn For \$26.2 Billion. *Insider*. Dostupné z: <https://www.businessinsider.com/microsoft-buys-linkedin-2016-6>.

van Dijk, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology, *Surveillance & Society*, 12(2), str. 197-208.

G. Venkatadri et. al. (2018). "Privacy Risks with Facebook's PII-Based Targeting: Auditing a Data Broker's Advertising Interface," *2018 IEEE Symposium on Security and Privacy (SP)*, str. 89-107.

Vincent, J. (7. května 2020). The EU is trying to fix its abysmal cookie consent policy. *The Verge*. Dostupné z: <https://www.theverge.com/2020/5/7/21250300/eu-cookie-consent-policy-updated-guidelines-cookie-wall>.

Vítak, J. & Ellison, N. B. (2013). "There's a Network Out There You Might as Well Tap": Exploring the Benefits of and Barriers to Exchanging Informational and Support-based Resources on Facebook, *New Media & Society*, 15(2), str. 243– 259.

Vlčková, Z. (3. února 2022). *Jsou cookies bezpečné? Máme přijmout všechna? Radí publicista Daniel Dočekal*. [Rádiové vysílání]. Mujrozhlas.cz. Dostupné z: <https://www.mujrozhlas.cz/kava-o-ctvrte/Jsou-cookies-bezpecne-mame-prijmout-vsechna-rad-i-publicista-daniel-docekal>.

WebFX (16. března 2020). What Are Data Brokers – And What Is Your Data Worth?. *WebFX*. Dostupné z: <https://www.webfx.com/blog/internet/what-are-data-brokers-and-what-is-your-data-worth-infographic/>.

WEF. (17. února 2011). Personal Data: The Emergence of a New Asset Class. *World Economic Forum*. Dostupné z: <https://www.weforum.org/reports/personal-data-emergence-new-asset-class/>.

X Holding I, Inc. (8. července 2022) *SEC*. Dostupné z: https://www.sec.gov/Archives/edgar/data/1418091/000110465922078413/tm2220599d1_ex99-p.htm

Yahoo Finance (2022). Microsoft Corporation (MSFT). *Yahoo Finance*. Dostupné z: https://finance.yahoo.com/quote/MSFT/?guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2x1LmNvbS8&guce_referrer_sig=AQAAAFASAgFpSCEUv71CLbFY27z5ilvprhnIOY-ay-nlXOfU70xBTmKNizFga5nV0HCHDFOO4L6JgtfMtKZMe18gSWucrOqxDy_DKWgMGpRI2EB3FtIFTH-FKXpd9klAg640YgNO3ibemodM_866XIVHYoAiEzgUMXVeRBDQYqcYYt43&gucounter=2.

Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y. & Chen, Z. (2009). How Much Can Behavioral Targeting Help Online Advertising?, In Proceedings of the 18th international conference on World wide web (WWW '09). *Association for Computing Machinery*, New York, NY, USA, str. 261–270.

Zákon 374/2021. (11. října 2021). Novela z. o elektronických komunikacích - EU. *Společná česko-slovenská digitální parlamentní knihovna*. Dostupné z: <https://www.psp.cz/sqw/historie.sqw?o=8&t=1084>.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.

Zuboff, S. (2014). Dark Facebook: Facebook's Secret Experiment in Emotional Manipulation Provides a Fresh Glimpse of its Radical Politics and Absolutist Ambitions. *Shoshanazuboff.com*. Dostupné z: <https://shoshanazuboff.com/book/dark-facebook/>.