

UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut sociologických studií

Katedra sociologie

Diplomová práce

2020

Tereza Svobodová

UNIVERZITA KARLOVA

FAKULTA SOCIÁLNÍCH VĚD

Institut sociologických studií

Katedra sociologie

**Sdílená ekonomika v kontextu postmateriálních hodnot:
případ segmentu ubytování v Praze**

Diplomová práce

Autor práce: Bc. Tereza Svobodová

Studijní program: Sociologie

Vedoucí práce: PhDr. Mgr. Jan Balon, Ph.D.

Rok obhajoby: 2020

Prohlášení

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

V Praze dne 21.5.2020

Tereza Svobodová

Bibliografický záznam

Svobodová, T. 2020. *Sdílená ekonomika v kontextu postmateriálních hodnot: případ segmentu ubytování v Praze*. Diplomová práce (Mgr.). Praha: Univerzita Karlova, Fakulta sociálních věd, Institut sociologických studií. Vedoucí diplomové práce PhDr. Mgr. Jan Balon, Ph.D.

Rozsah práce: 167 594 znaků s mezerami

Abstrakt

Diplomová práce se zabývá úspěchem sdílené ekonomiky v segmentu ubytování v Praze. Východiskem jsou teorie chápající rozmach sdílené ekonomiky jako produkt vývoje sociálního a hodnotového, nikoliv pouze technologického. S využitím tzv. velkých dat v podobě všech volně dostupných recenzí sdíleného ubytování přes Airbnb a tradičního ubytování přes Booking za poslední dva roky jsou komparovány zkušenosti jejich uživatelů z hlediska uspokojení potřeb, potažmo naplnění hodnot. Ke zpracování jsou užity techniky pro dolování z textů, a sice strukturní modelování témat a analýza sentimentu. Hlavním zjištěním je, že modely sdíleného ubytování uspokojují společensky rostoucí potřebu naplňovat postmateriální hodnoty na trhu mnohem lépe než modely tradičního ubytování (hotely, hostely, penziony) v Praze. Uživatelé Airbnb při popisu svých zkušeností mnohem častěji reflektují hodnoty sociální a emoční, i přestože ve většině sdílených ubytování v Praze nedochází ke sdílení fyzickému. Práce tak přináší v českém kontextu unikátní perspektivu na fenomén Airbnb a přispívá k diskusi nad tím, proč podíl sdílené ekonomiky na trhu s ubytováním v Praze roste, zatímco tradiční modely spíše stagnují.

Abstract

This master's thesis is about the success of sharing economy in the accommodation segment in Prague. The thesis is based on theories conceptualizing sharing economy as a result of social and value change, not only as technological. Using online review data, the user experience of shared accommodation via Airbnb and traditional via Booking are compared. Analysis is conducted with focus on users' satisfied needs and fulfilled values. For processing the data, text mining techniques (topic modelling and sentiment analysis) were employed. The major result is that in Prague the models of sharing economy accommodation meets the growing need in society to fulfil post-material values in the market much better than the models of traditional accommodation (hotels, hostels, boarding houses). In their experiences, Airbnb users reflect social and emotional values more often, even though most sharing economy accommodations in Prague do not involve any physical sharing with the host. The thesis thus brings a unique perspective on the Airbnb phenomenon in the Czech context and contributes to the discussion of why the market share of the sharing economy in the accommodation segment in Prague has been growing, while traditional models stagnated.

Klíčová slova

sdílená ekonomika, Airbnb, postmateriální hodnoty, big data, text mining, zpracování přirozeného jazyka, modelování témat, analýza sentimentu

Keywords

sharing economy, Airbnb, post-material values, big data, text mining, natural language processing, topic modeling, sentiment analysis

Title

Sharing Economy in the Post-Material Values Context: The Case of Accommodation Segment in Prague

Poděkování

Ráda bych tento prostor využila k poděkování svému vedoucímu, PhDr. Janu Balonovi, Ph.D., za vstřícnost i veškeré připomínky a zpětnou vazbu v průběhu přípravy i psaní práce.

Obsah

Úvod.....	1
1 Teoretická část	3
1.1 Vymezení konceptu sdílené ekonomiky – široká a úzká definice.....	3
1.2 Sdílená ekonomika jako produkt post-industriální společnosti.....	7
1.3 Sdílená ekonomika v segmentu ubytování – projekt Airbnb	11
1.3.1 Principy a hodnoty Airbnb	11
1.3.2 Kontroverznost Airbnb a situace v České republice	14
2 Metodologie a plán výzkumu	17
2.1 Vymezení cílů výzkumné části.....	18
2.2 Data	20
2.2.1 Zdroj dat, jejich výhody a alternativy	20
2.2.2 Sběr dat.....	22
2.3 Metody a postupy.....	25
2.3.1 Text mining	25
2.3.2 Příprava dat	26
2.3.3 Metoda Structural topic modeling (STM).....	29
2.3.4 Analýza sentimentu	30
2.4 Etické hledisko	32
3 Analytická část	33
3.1 Popis dat	33
3.2 Analýzy frekvencí	36
3.2.1 Komparace z hlediska nejčetněji užívaných výrazů.....	36
3.2.2 Explorace kompletních souborů z hlediska proporčních frekvencí	42
3.3 Modelování témat	48
3.3.1 Trénování, evaluace a interpretace modelů – postup	48
3.3.2 Modelování témat v korpusu recenzí z Airbnb	50
3.3.3 Modelování témat v korpusu recenzí tradičních ubytování	57
3.3.4 Komparace obsahu zkušeností se sdíleným a tradičním ubytováním	63
3.4 Porovnání obsahu recenzí z hlediska sentimentu	67
3.5 Diskuse nad výsledky.....	72
Závěr.....	77
Summary	80
Použité zdroje	81

Seznam příloh	99
----------------------------	-----------

Úvod

„Kolaborativní spotřeba (...) není jen prchavá myšlenka nebo krátkodobý trend, ale silný kulturní a ekonomický proud měnící nejen co konzumujeme, ale jak konzumujeme.“ – Rachel Botsman [2010]¹

Kolaborativní spotřeba, sdílená ekonomika či peer-to-peer produkce. Jakkoliv je tento nový trend nazýván, vyvolal silnou vlnu zájmu v akademických i společenských debatách. V jeho počátcích kolem roku 2008, kdy Airbnb a Uber, průkopníci sdílené ekonomiky, byli ještě v plenkách, se myšlenka přespát u někoho cizího doma či se jím nechat svést v jeho autě zdála jako bláznivá nebo vhodná možná tak pro pár nadšenců. Dnes je již nicméně jasné, že se nejedná jen o idealistickou představu a pomíjivý trend. Stále více projektů na obdobném principu získává své místo na trhu. Filozofie sdílení, ať už věcí, znalostí, schopností, zdrojů, kapacit k přepravě či bydlení se šíří napříč segmenty a dává vzniknout novým projektům, novým formám práce, produkce i spotřeby. To vzbuzuje mnoho otázek a zájem celé řady oborů. Od polemik nad tím, co vlastně sdílená ekonomika je a co už není, jak se liší od tradičních modelů, jestli se jedná o obyčejný byznys bez jakýchkoliv pravidel nebo naopak naději, že trh se může stát efektivnějším, ekologicky šetrnějším a heterogennějším. Podle databáze Web of Science počet publikací na téma sdílené ekonomiky rok od roku roste. Největší podíl je z oblasti ekonomie [Bruno a Faggini 2020; Blal et al. 2018; a další], managementu [Dogru et al. 2020; Akbar a Tracogna 2018; a další] a byznysu [Akhmedova et al. 2020; Benoit et al. 2017 a další].

V předložené práci nabízím perspektivu společenskovědní či axiologickou. Vycházím z předpokladu, že za úspěšností sdílené ekonomiky nestojí pouze změny technologické, nýbrž především sociální a hodnotové, které technologie urychlují poskytnutím platformy pro jejich realizaci [Botsman 2018; Hyman 2018; David 2017; Sundararajan 2016; Mason 2016; Zuboff 2010]. V návaznosti pak na případu segmentu ubytování porovnávám spotřebitelskou zkušenost s tradičními modely a modely sdílené ekonomiky právě z hlediska uspokojení potřeb, potažmo naplnění hodnot. Práci standardně dělím do tří částí – teoretické, metodologické a analytické. V teoretické části věnuji pozornost vymezení pojmu sdílená ekonomika a variabilitě

¹ Originální znění v angličtině je: *„Collaborative consumption (...) is not a flimsy idea or a short-term trend, but a powerful cultural and economic force reinventing not just what we consume, but how we consume.“* [Botsman 2010]

užívaných významů i označení. Dále popisuji souvislosti růstu tohoto nového sektoru v širším kontextu společenského vývoje, vývoje lidských hodnot i dalších faktorů. Pro komplexitu zkoumaného fenoménu se dále zaměřuji na jeden vybraný projekt, a sice Airbnb ze segmentu ubytování. Projekt představuji s důrazem na způsob jeho komunikace a filozofii. Předkládám úvahy nad tím, jaké hodnoty slibuje nabízet a jak by toto mohlo souviset s jeho úspěšností. V další části se nicméně zaměřuji také na důvody jeho kontroverze (nejen) v českém kontextu. Ve výzkumné části navazuji na nastolenou otázku hodnot, jež jsou naplňovány v rámci sdílené ekonomiky a zkoumám zkušenosti lidí ubytovaných přes Airbnb v Praze. Pro pochopení, jak se tato zkušenost liší od ubytování v tradičním sektoru, volím přístup komparativní a zkušenosti uživatelů Airbnb porovnávám se zkušenostmi uživatelů hromadných ubytovacích zařízení v Praze. Využívám tzv. velká data v podobě slovních recenzí z webů Airbnb a Booking. Pro analýzu obsahu zkušeností aplikuji techniky pro dolování v textech a výsledky mezi korpusy porovnávám. Užité postupy detailně představuji v metodologické části práce. Účelem je na případu segmentu ubytování přispět porozumění tomu, jak se liší projekty sdílené ekonomiky od té tradiční z hlediska uspokojování potřeb konzumenta. Přeneseně tedy také přispět k diskusi nad tím, proč podíl sdílené ekonomiky na trhu roste a těší se takové úspěšnosti.

1 Teoretická část

Koncept sdílené ekonomiky zdaleka není jednoznačný, zahrnuje mnoho různých projektů a užívána je celá řada definic. První oddíl teoretické části proto věnuji vymezení, jak je možné jej chápat, jak s ním pracuji v této práci a proč vznikají třecí plochy při výkladu významu. Druhou část poté zaměřuji na faktory úspěšnosti modelu sdílené ekonomiky, zejména ze sociálněvědního a axiologického hlediska. Poslední, třetí část vyhrazuji již konkrétně Airbnb, vybranému projektu ze segmentu ubytování. Představuji jeho principy, rétoriku i kritiku, včetně popsání situace v České republice.

1.1 Vymezení konceptu sdílené ekonomiky – široká a úzká definice

V debatách ohledně sdílené ekonomiky často vznikají třecí plochy pramenící z různého chápání významu celého konceptu. Nejen že neexistuje společně užívaná definice, ale dokonce není ustálené ani samotné označení. Kromě pojmu sdílená ekonomika [David 2017; Mason 2016; Sundararajan 2016; Stephany 2015] je možné setkat se také s pojmy kolaborativní spotřeba [Botsman a Rogers 2010], peer-to-peer produkce [Botsman 2018] nebo zakázková ekonomika [Hyman 2019], ale i mnohými dalšími. Velmi často jsou přitom užívány jako vzájemně zastupitelné, případně je mezi nimi jen mírně rozlišováno. Sdílená ekonomika je nicméně pojmem, který se uchytil nejvíce a který byl v roce 2015 dokonce zanesen i do Oxfordského slovníku angličtiny, proto jej užívám i v této práci. Různá označení ovšem zdaleka nejsou největším problémem. Skutečně nelehký úkol totiž spočívá v najetí hranice mezi tím, co lze a co již nikoliv označit za sdílenou ekonomiku. Uvádím proto přehled různých projektů sdílené ekonomiky a snažím se sumarizovat jejich společné znaky. Představuji dva možné přístupy, jak lze sdílenou ekonomiku chápat, a sice tzv. úzkou a širokou definici. Cílem je také ujasnit, jak sdílenou ekonomiku chápu v této práci.

Za posledních deset let se zrodilo celé množství projektů, které bývají označovány jako sdílená ekonomika, přičemž jsou ale značně různorodé. Odvětvími, kde se tyto projekty prosazují, jsou zejména ubytování, přeprava, lidská síla, finance, vzdělávání, pronájem věcí a média. Stručný přehled konkrétních projektů, globální i domácích, je k nahlédnutí v Tabulce 1 (str. 6). Základním charakteristickým znakem je provádění transakcí napřímo – lidé s lidmi

[Botsman 2018; David 2017; Mason 2016]. A co je důležité, děje se tak ve velkém měřítku, právě díky propojení strany nabídky a poptávky přes technologické platformy. Lidé již nejsou odkázáni jen na své přímé okolí, pokud si například chtějí půjčit kolo nebo u někoho přespat, ale mohou se s druhým člověkem propojit přes internet. Na straně poptávky to tedy znamená možnost uspokojit své potřeby rychlým a přímým propojením s někým, kdo disponuje poptávanými zdroji, bez využití tradičních zprostředkovatelů (banky, školy, hotely, agentury a další), často i finančně výhodněji. Na straně nabídky potom příležitost efektivnějšího využití svých zdrojů, a to velice flexibilně, bez nutnosti závazků nebo splňování náročných předpokladů.² Na těchto základních principech – **přímé transakce, využití technologií, efektivnější využití zdrojů** – jakožto charakteristikách sdílené ekonomiky, obvykle panuje shoda [Botsman 2018; David 2017; Sundararajan 2016; Mason 2016; a další].

Problematické nicméně je, že se na jednu stranu se sdílenou ekonomikou pojí značná idealistická očekávání, na stranu druhou se stala atraktivním trendem, pod jehož hlavičku se snaží zařadit i subjekty, které mají k těmto ideálům daleko, a tím se obsah konceptu dále roztahuje [Dobusch 2019, str. 110; Bergh et al. 2018]. Od počátku byl s projekty sdílené ekonomiky spojen příslib udržitelnějšího, heterogennějšího a rovnějšího trhu založeného na osobních transakcích a vedoucího k méně konzumnímu stylu života. A ačkoliv tyto ideály mohou být pro mnohé motivem k participaci na sdílené ekonomice, nepochybně hraje důležitou roli možnost přivýdělku. Jistě by nebylo tolik hostitelů přes Airbnb a řidičů přes Uber, pokud by měli služby poskytovat zdarma.³ Ačkoliv je běžně součástí definic, že sdílená ekonomika zahrnuje transakce zdarma i za poplatek [Botsman 2018; David 2017; Sundararajan 2016], tato **kontradikce mezi humanistickým sociálním elementem a kapitalistickou motivací generovat profit vyvolává napětí**. Nepanuje shoda na tom, nakolik má být participace na sdílené ekonomice motivovaná finančně a nakolik právě představenými ideály. Vystávají pak otázky, jestli například sdílená ekonomika v oblasti dopravy má mít takovou podobu, že řidič přibere na cestu někoho dalšího pouze v případě, že cestu podniká on sám nebo jestli je stále

² Důvěra, která mezi těmito stranami panuje, také nabývá nové formy. Jak píše Botsman [2018], již není instituční, založená na licencích, příslušném vzdělání, splňování nastavených standardů apod., ale přechází do tzv. distribuované podoby, kde zdrojem důvěryhodnosti je zejména elektronická pověst, tudíž recenze. V praxi to znamená, že k vykonávání role řidiče již není nutné být licencovaným taxikářem, ale stačí se registrovat například na Uberu a budovat si svou pověst na základě získávaných recenzí od zákazníků. Stejně tak je možné ubytovávat turisty, být chůvou, učit děti nebo třeba mít vlastní rozhlasový pořad. Pozornost konceptu elektronické pověsti (zkatkou eWOM) tak roste také v akademických kruzích. Ve sdílené ekonomice nicméně nezískává novou podobu jen důvěra, ale jak píše Mason [2016, str. 15] nebo Vătămănescu [2017, str. 15], také sdílení, vlastnictví i soukromí získává novou formu a nové významy.

³ Důkazem budiž Couchsurfing, bezplatná obdoba Airbnb, který stále zůstává spíše okrajovou záležitostí.

sdílenou ekonomikou i využít své auto k převážení lidí pouze na jejich objednávku, aniž by trasu sám jinak odjel. Stejně tak jestli prostor nabízený k přenocování by měl být pronajímán jen občasné nebo jestli je stále sdílenou ekonomikou vlastnit byt navíc a hostit v něm turisty po celý rok. Určit hranici vůbec není jednoduché. Bergh [et al. 2018] se tímto zabývá při snaze operacionalizovat koncept sdílené ekonomiky a nabízí přehled užívaných definic. Shrnuje, že jsou dva způsoby, jak nahlížet na sdílenou ekonomiku, které lze označit jako **definici úzkou** [Frenken et al. 2017; Zervas et al. 2017; Schor et al. 2016] a **širokou** [Botsman 2018; Sundararajan 2016]. Do úzké definice by spadaly jen první z výše uvedených příkladů, tedy zahrnuje pouze takové transakce, kdy participanti nabízejí výhradně dočasný přístup k nevyužitým kapacitám s motivací „(...) *sdílet a spolupracovat při produkci a spotřebě zboží a služeb*“ [Schor et al. 2014 in Bergh et al. 2018]. Motivace k účasti je tedy velmi důležitá a neměla by být jen kapitalistická, jinak se údajně jedná o klasický byznys bez přidané hodnoty. Do široké definice lze naopak zařadit nejen první, ale i druhé z příkladů, jelikož charakterizuje sdílenou ekonomiku pouze na základě typů transakcí, tedy přímým propojením mezi decentralizovanou poptávkou a nabídkou za účelem efektivnějšího využití zdrojů (majetku i času), s nárazovou povahou transakcí. Klade menší důraz na motivace k účasti a soustředí se jen na výsledný efekt, který zůstává zachován neohledně na motivaci. Bergh [et al. 2018] konstatuje, že sledovat naplnění úzké definice je v některých případech nejen až nemožné, ale především vylučuje velký objem sdílení podobných aktivit, které jsou na pomezí a nelze je tak nikam zařadit. Obdobně také argumentuje David [2017], že sdílenou ekonomiku lze chápat nejen jako alternativu k byznysu, ale také jako alternativní byznys. Dobusch [2019] pak obdobně rozlišuje mezi komunitně založenou a tržně založenou sdílenou ekonomikou. V každém případě se ale, tvrdí, jedná o nový model, který se od tradiční ekonomiky liší.

V rámci této práce chápu koncept sdílené ekonomiky v tom širším pojetí, které může zahrnovat jak morální, tak kapitalistické intence k participaci. Sdílenou ekonomiku lze tedy definovat jako úplatné i bezplatné transakce napřímo uskutečňované prostřednictvím technologické platformy s cílem efektivnějšího využití zdrojů, které potenciálně, nikoliv nutně, mohou vést k navazování osobnějších vazeb a redukování nadměrné spotřeby, znečištění a nerovností.

Tabulka 1: Přehled vybraných projektů sdílené ekonomiky dle segmentů

SEGMENT	GLOBALNÍ	DOMÁCÍ
UBYTOVÁNÍ	Airbnb HomeAway Vrbo Couchsurfing	Flatio
DOPRAVA	Uber Taxify Lift Blablacar Škoda auto DigiLab Moia Velonet Ofo	Bringr Car4way HoppyGo Liftago Zavezu Autonapůl AJO Rekola
FINANCE	Kickstarter GoFundMe Indiegogo	Hithit Zonky Fundlift Startovač
SLUŽBY, NÁJEM PRÁCE	Freelancer TaskRabbit	LidskáSíla Doginni Robeeto Hlídačky.cz
VZDĚLÁVÁNÍ	Khan Academy Coursera EdX	Nostis
NÁŘADÍ A DALŠÍ VĚCI	PeerBy Etsy	SharyGo Fler
NEBYTOVÉ PROSTORY	PivotDesk DesksNear.Me	sdílejprostor
MÉDIA	Spotify Apple Music Google Music Amazon Music Netflix Youtube	

Zdroj: Deloitte [2017, str. 14], Česko v datech [2018], vlastní zpracování

1.2 Sdílená ekonomika jako produkt post-industriální společnosti

„Věřím, že jsme na počátku kolaborativní revoluce, která bude stejně tak významná jako průmyslová revoluce.“ – Rachel Botsman [2012]⁴

V této kapitole se věnuji faktorům úspěšnosti modelů sdílené ekonomiky, které si v řadě sektorů vybudovaly významné místo na trhu vedle modelů tradičních. Výchozí tezí je, že sdílená ekonomika není jen osamoceným proudem nebo čistě technologicky hnanou změnou v oblasti produkce a spotřeby, nýbrž reflexí dlouhodobě probíhajících změn i na poli lidských hodnot a uspořádání společnosti [Botsman 2018; Hyman 2018; David 2017; Sundararajan 2016; Mason 2016; Zuboff 2010]. Jak píše Vătămănescu [2017, str.10], rozmach sektoru sdílené ekonomiky je jen vrcholem ledovce v kontextu změn způsobu myšlení a pronikající filozofie postmateriálních hodnot. Pro popsání současného stavu, ve kterém se daří sektoru sdílené ekonomiky, je proto nutné alespoň velmi stručně nastínit předcházející vývoj. Do tohoto širšího kontextu pak zasazují vznik a úspěšnost sdílené ekonomiky, včetně dalších přispívajících faktorů – vývoj technologií, environmentální problémy a ekonomická krize roku 2008.

Jako první, hlavní faktor úspěšnosti sdílené ekonomiky chápu vývoj uspořádání společnosti a lidských hodnot. Zhruba každé století dochází k významným změnám povahy produkce a spotřeby, kterým stávající uspořádání plně nevyhovuje. Jak píše Mason [2016, str. 8], kapitalismus je učitím se organismem, který se neustále adaptuje. Obecně se hovoří o třech jeho základních etapách – tradiční, industriální a post-industriální.⁵ Zuboff [2010] tento vývoj stručně popisuje jako přesun v oblasti spotřeby od elit k masám a následně od mas k individuům. Tedy stejně jako s průmyslovou revolucí došlo k posunu od diversifikovaného drobného podnikatelství k masovému způsobu produkce, od konce 60. let minulého století se přesouvá pozornost zpět od mas na individua. To je doprovázeno významnými kulturními, sociálními i ekonomickými změnami, které spolu úzce souvisí. Boltanski a Chiapello [2007] píší, že post-industriální období je vlastně vyhraněním se vůči období industriálnímu a pohlcením jeho kritiky. Industriální období je totiž charakterizováno masifikací, standardizací, centralizací a byrokratizací [Toffler a Toffler 2001, str. 1-18], což jsou zároveň rysy, které nevyhnutelně

⁴ Originální znění v angličtině je: „I believe that we are at the start of a collaborative revolution that will be as significant as the Industrial Revolution.“ – [Botsman 2012]

⁵ Označení se ale mohou lišit.

vedly k opomíjení lidských kvalit a individualit [Ditz 2014, str. 34; Boltanski a Chiapello 2007, str. 37]. Ačkoliv takové uspořádání umožnilo rychlý hospodářský růst a vysokou efektivitu, setkala se s výraznou kritikou. Nespokojenost vyústila ve vlnu protestů, nejen verbálních, ale i v podobě stávek, absencí či fluktuací, na což bylo nutné reagovat [Boltanski a Chiapello 2007, str. 169-173]. Změny navíc byly urychleny vývojem technologií, díky čemuž došlo k delegaci velké části mechanické práce na stroje a přesunu lidské síly do sektoru služeb. Od 70. let 20. století tak lze mluvit o nové fázi, někdy nazývané jako post-industriální [Bell 1976]. Tato transformace podnítila tzv. kognitivní mobilizaci [Inglehart 2008], došlo k prudkému nárůstu vzdělanosti, celkové životní úrovně a sociální komplexity. Díky poklesu pracovní doby, nárůstu mzdy i očekávané délky života vznikl nový prostor pro utváření životního stylu [Beck 2004, str. 121-125]. Inglehart [2008] tento vývoj dokumentuje ve výzkumu World Values Survey a dokládá, že se zvyšující se kvalitou života souvisí obrat lidských hodnot od materiálních k postmateriálním. Vzedmula se **nová touha po individualizaci, svobodě, seberealizaci, aktivním přístupu a možnosti mít nad věcmi kontrolu**. Jedná se o změnu postupnou, převážně mezigenerační, která se opět reflektuje i v oblasti produkce a spotřeby. V důsledku toho se orientace na produkt musela začít měnit v orientaci na konzumenta a osobní přístup [Boltanski a Chiapello 2007, str. 89-98; 170-182]. Firmy, které cílí na koncové zákazníky, již nadále nemohou být pouhými prodejci či výrobci, ale měly by mít **přesah – příběh, autenticitu, důvěryhodnost – a tvořit smysluplnou značku** [Zuboff 2010]. Nepřemýšlet jen nad tím co budou nabízet, ale také jak to budou nabízet [Boltanski a Chiapello 2007, str. 89–98, 170–182]. Značný důraz začal být kladen i na zákaznický servis. Základní konkurenční výhodou na trhu se tak staly hodnoty dříve spojované výhradně se soukromým životem. Boltanski a Chiapello [2007, str. 464–466] tento akt popisují jako komodifikaci lidských hodnot kapitalismem. Jedná se například o **empatii, emocionalitu, důvěru, sounáležitost**, které do té doby na úrovni anonymních vztahů v rámci trhu neměly místo a jsou spojovány spíše se soukromými vztahy [Hochschild 1983]⁶. Je to typickým znakem post-industrialismu, ale zároveň také částečně návrat k osobním transakcím tradičního období.

Tato **změna orientace od hmotného produktu k nehmotným kvalitám** jako jsou osobní přístup a zážitky je tak chápána jako významná okolnost pro expanzi sdílené ekonomiky vedle tradičního sektoru, jelikož by měla být schopna na tyto hodnoty lépe reagovat. Napovídá tomu i fakt, že profil typického uživatele sdílené ekonomiky odpovídá profilu nositele postmateriálních hodnot. Jedná se o lidi vzdělanější a mladší, převážně generaci mileniálů, se

⁶ Hochschild [1983] pro toto používá označení „management emocí“ či „emoční management“.

zájmem o environmentální problematiku [Mason 2016, str. 113–120; Botsman 2010]. I pro oblíbenost mladšími generacemi se očekává další značný růst, a to zejména do roku 2025, kdy mileniálové budou tvořit většinu produktivní populace.⁷ Zároveň je ve sdílené ekonomice výrazně zapojen i zmíněný emoční management díky povaze přímých transakcí bez anonymního prostředníka. Sama podstata sdílení vytváří pocit sounáležitosti [Botsman 2010].⁸ Sdílená ekonomika by tedy měla výborně cílit na dvě základní potřeby dnešního postmateriálně orientovaného konzumenta – individuální identitu a osobní přístup.

Dalším neopomenutelným faktorem a předpokladem rozmachu sdílené ekonomiky je rozvoj technologií. I když touha po změně ve způsobu spotřeby pramení ze společnosti samotné, technologie jí poskytly nový prostor.⁹ Umožňují expanzi sdílené ekonomiky do globálního měřítka – propojují jednotlivce nabízející služby s těmi, kdo je poptávají, a to velice efektivně. Technologické společnosti jako Uber nebo Airbnb tedy nechápu jako aktéry změny, nicméně jako poskytovatele platformy pro změny společenské, které již probíhají a na které velmi dobře reagují a umocňují je [Hyman 2018; Sundararajan 2016]. Sdílení mezi lidmi v nějaké míře samozřejmě probíhalo vždy, nicméně **technologie bourají jeho hranice**. Nejen že rozšiřují okruh těch, kdo se mohou propojit za účelem sdílení, ale také mezi nimi vytváří most pro překlenutí přirozené nedůvěry vůči cizím lidem [Botsman 2018].¹⁰ Díky nim se otevírá úplně nový prostor pro ekonomické aktivity, a sice výrazně flexibilnější, osobitější a heterogennější, kde vše je dostupné na jedno kliknutí. Vytváří se tak mezi lidmi síť takřka bez hranic a umožňuje mnohem osobnější přímé transakce.¹¹ Poslední významnou okolností úspěchu sdílené ekonomiky je pak **rostoucí zájem o efektivnější využití zdrojů** [Mason 2016; Botsman 2010]. Důvodem je jak stále naléhavější environmentální krize, tak ale i krize ekonomická v roce 2008 a její následky.¹² Vzniká tak potřeba využít své zdroje, ať už hmotné

⁷ Podle studie PwC [2015] by mezi lety 2013 a 2025 mohla sdílená ekonomika ve svých klíčových sektorech (finance, ubytování, přeprava, sdílení hudby a videa, práce na dálku) zvýšit příjmy z 13 miliard dolarů na 335 miliard. Polovina všech příjmu v těchto sektorech by tak patřila společnostem postaveným na principech sdílení.

⁸ Botsman [2010] dává potřebu sounáležitosti do souvislosti s technologiemi, díky kterým se člověk stává izolovanějším, a proto více touží po lidském kontaktu. Píše, že sdílená ekonomika tak propojuje dvě významné potřeby dnešního mladého člověka – oblíbené online prostředí a interakci s reálnými lidmi.

⁹ Mezi lety 2000 a 2015 došlo až k devítinásobnému růstu ve využívání internetu [Úřad vlády 2017]. Dnes jsou téměř dvě třetiny světové populace aktivními uživateli internetu, v severní Evropě je to pak dokonce 95 % populace [Statista 2020b]. Sociální sítě se stávají běžnou součástí lidských životů. V souvislosti s rozvojem sociálních médií společnost od 90. let získává nový přívlastek – síťová [Castells 1996].

¹⁰ Stává se tak běžným, že lidé dávají přednost neprofesionálnímu řidiči vybranému pouze podle recenzí nebo ověření profilu před licencovaným taxikářem. Stejně tak na svých cestách místo v hotelu splňujícím standard nabízených služeb podle počtu hvězdiček raději přespí v soukromém bytě cizího člověka. Chůvu si vyberou podle jejího online profilu, peníze si půjčí od jiných cizí lidí místo bank a tak dále.

¹¹ Možnost obcházet prostředníky, kteří mnohokrát zklamali důvěru (banky, neprůhledné korporáty), bývá také uváděna jako jeden z motivů k participaci na sdílené ekonomice [Botsman 2018].

¹² Roku 2008 a 2009 vznikly dva největší projekty sdílené ekonomiky, Airbnb a Uber.

či nehmotné, co nejlépe. Motivovaná může být jak ekologicky, tedy nepodněcovat další produkci, pokud je možné využít již existující zdroje, ale také finančně, tedy využít své zdroje k dodatečnému příjmu či naopak pořídit potřebnou službu či produkt levněji. Celkově ale, patrně v kombinaci všech zmíněných faktorů, dochází ke změně vztahu k vlastnictví, kdy se ukazuje, že mít přístup k produktům a službám je často výhodnější než je vlastnit [Botsman a Rogers 2010].¹³ Řada autorů polemizuje nad myšlenkou, zda se právě nenacházíme na pokraji další tzv. mutace kapitalismu, s čímž dávají do souvislosti i sdílenou ekonomiku, se kterou v řadě segmentů bojuje tradiční trh [Botsman 2018; David 2017; Mason 2016; Sundararajan 2016; Zuboff 2010].¹⁴

Enormní úspěch sdílené ekonomiky tedy, zdá se, tkví v propojení několika aktuálně žádaných kvalit. Těmi jsou postmateriální hodnoty (individualita, zážitkovost, sounáležitost, otevřenost, udržitelnost), dobré zvládnutí managementu emocí a chytré využití technologií i zdrojů. Zároveň je s mezigenerační obměnou očekáván další růst těchto hodnot ve společnosti, a proto i sdílené ekonomiky, která podle zmíněných autorů má potenciál je naplňovat vyšší než tradiční modely.

¹³ Ukazuje se, že zejména mileniálové nepotřebují vlastnit DVD, ale chtějí mít přístup k filmu. Nepotřebují vlastnit auto, ale mít přístup k jeho využití v době potřeby. Stejně tak se dá tato myšlenka aplikovat na vrtačku, cyklistické kolo, knihy a další produkty. Lidé tedy nepotřebují nutně kumulovat hmotný majetek, ale chtějí skrz něj uspokojovat své potřeby [Botsman a Rogers 2010].

¹⁴ Hlavní zvrát někteří autoři vidí zejména v segmentech, kde zapojení technologií znamená nulové mezní náklady, což je výzvou pro základní principy kapitalismu (nedostatkost, cenové principy), který předpokládá soutěž, jenomže informační statky se mohou šířit donekonečna bez dalších nákladů (hudba, filmy, články apod.) [David 2017; Mason 2016].

1.3 Sdílená ekonomika v segmentu ubytování – projekt Airbnb

Dosud jsem o sdílené ekonomice hovořila obecně. Jak je ale řečeno již v části o vymezení pojmu, zahrnuje mnoho projektů z různých segmentů. Pro účely této práce jsem proto zvolila jeden, a sice segment ubytování, ve kterém vedle tradičních modelů působí projekt sdílené ekonomiky Airbnb a získává si stále větší podíl na trhu.¹⁵ V této části projekt představuji a v návaznosti na části předchozí rozebírám jeho hodnoty, jak je komunikuje a slibuje naplňovat. Diskutuji také nad kontroverzností projektu a předkládám stručný přehled o aktuální situaci s ubytováním v Praze.

1.3.1 Principy a hodnoty Airbnb

Airbnb je jedním z prvních a nejúspěšnějších projektů na principech sdílené ekonomiky. Pro svůj rapidní růst si získalo značnou pozornost médií, světa byznysu, státních aparátů, ale také akademické obce. Jedná se o technologickou společnost – platformu, která napřímo mezi lidmi zprostředkovává ubytování v jejich soukromých prostorech, aniž by je sama vlastnila či hostitele zaměstnávala. Počátky Airbnb sahají teprve do roku 2008 a dnes má celosvětově zaregistrovaných více než 7 milionů ubytování ve více než 100 tisících měst a zprostředkovalo již přes 750 milionů přenocování [Airbnb, Inc.]. Jeho tržní hodnota byla v roce 2018 odhadnuta na 31 miliard, což je více než hodnota většiny hotelových řetězců, dvojnásobek například celosvětového řetězce Hilton [Statista 2020a; Cheng a Jin 2019, str. 58; Zhang 2019, str. 371]. A jak se tedy Airbnb liší svou nabídkou od tradičních modelů ubytování? Jak komunikuje svou přidanou hodnotu? Účelem této kapitoly je rozebrat hlavní principy a rétoriku Airbnb, nakolik odpovídá teoretickým úvahám kapitoly předchozí, že sdílená ekonomika lépe cílí na hodnoty postmateriálně orientovaného spotřebitele.

Specifičností Airbnb je jeho silná narativizace nabízených služeb. V roce 2014 započalo svou výraznou kampaň spojenou s logem „Belong Anywhere“ (viz Obrázek 1, str. 13). Pozměněna byla celá filozofie komunikační strategie ubytovávání přes Airbnb. Společnost se

¹⁵ Kromě projektu Airbnb působí na trhu se sdíleným ubytováním ještě HomeAway, Couchsurfing, Vrbo a další, nicméně Airbnb má zdaleka největší podíl i pozornost. Zvolila jsem jej také pro jeho diskutovanost v českém kontextu a významný podíl na trhu s ubytováním v Praze.

začala zásadně distancovat od toho, že by nabízela běžné ubytování a zaměřila pozornost ve své komunikaci na tři základní hodnoty: unikátnost, sounáležitost a autenticitu.¹⁶ Tedy hodnoty postmateriální. Za jediný rok se Airbnb podařilo více než zdvojnásobit počet ubytování na evropském trhu a zvýšit tržní hodnotu společnosti z 10 na 25,5 miliard [Brown 2014; Statista 2020a]. Lze tedy předpokládat, že uživatelům byly nabídnuty hodnoty, které jsou jimi žádané a které do té doby nebyly dostatečně k dispozici. A jak konkrétně Airbnb slibuje tyto postmateriální hodnoty uspokojovat? Zásadní odlišností vůči modelům tradičního ubytování je již samotné pojmenování nabízených prostorů k přenocování – „domovy“. Airbnb vystavělo svou filozofii na základním principu jeho fungování, a tedy že se jedná o ubytování u cizího člověka v jeho soukromém bytě, tedy vlastně v intimním prostředí [Guttentag 2015]. Z toho by měla vycházet zmíněná unikátnost, sounáležitost a autenticita. **Unikátnost** tkví v tom, že i když Airbnb nabízí několik milionů prostorů k pronájmu, žádný by neměl být stejný. Nejedná se o hromadná ubytování, řetězce, nejsou nijak standardizována. Každý prostor by měl být odrazem osobnosti jeho hostitele. Airbnb velmi dbá na diverzitu nabídky, která by díky tomu měla být jedinečná a na míru pro každého hosta.¹⁷ Ten by tak měl mít možnost uspokojit svou potřebu odlišit se. **Sounáležitost** je další výrazně komunikovanou hodnotou Airbnb. Brian Chesky [2014], jeden ze zakladatelů Airbnb, tvrdí, že misí Airbnb není nabízet ubytování, ale vytvářet celosvětovou komunitu, je o lidech a zážitcích. Jejich síla údajně tkví v tom, že vytváří most důvěry mezi cizími lidmi [Friedman 2013].¹⁸ Host by tak měl mít možnost poznat život někoho místního v jeho vlastním domově a navázat s ním osobní vztah. Botsman [2012] říká, že projekty jako Airbnb lidem umožňují opět navázat smysluplná propojení a znovuobjevit určitou lidskost, která se v neosobních transakcích vytratila. Airbnb tento příslib komunikuje ve svých kampaních hesly jako „Vítej doma“ nebo „Žij tam“ (viz Příloha 1 a 2 na str. 100). Hosté by si tak měli lépe vychutnat místo a poznat jeho autentickou tvář, nikoliv turistickou [Guttentag 2015]. **Autenticita** je tak třetí hodnotou, na kterou Airbnb klade důraz. Host by měl přebývat v autentickém domově, tak jak vypadá v daném místě, nikoliv v univerzálně vybaveném hotelu. Měl by tak zažít autenticitu místa, být v kontaktu přímo s místním a dostávat od něj tipy, jaké v průvodcích nenajde. Airbnb směřuje veškerou svou komunikaci na vytváření takové image

¹⁶ Tato komunikační strategie byla založena na rozhovorech s uživateli o tom, co oni sami chtějí [Brown 2014].

¹⁷ Na Airbnb jsou k dispozici různorodé prostory od designových bytů, domů a pokojů přes karavany, stromové domky, iglú, hrady, zámky až po soukromé ostrovy či jeskyně.

¹⁸ Airbnb se snaží pracovat na důvěryhodnosti zpracovaným systémem ověřování uživatelů, pojištěním proti škodám i non-stop linkou k dispozici, což umožňuje překonat jinak přirozenou podezřívavost vůči tomu strávit noc u někoho cizího doma, případně si pozvat někoho cizího do svého domova [Botsman 2018]. Běžně by takový krok lidé, kteří se neznají, nepodnikali.

své značky. Kromě reklamních kampaní také provozuje blog, sociální sítě, pořádá workshopy pro hostitele. Sdílí autentické příběhy, tipy mezi hostiteli, stimuluje interakci. Prostřednictvím tohoto obsahu signalizuje, že se jedná o projekt, který je převážně o komunitě lidí, místech a unikátních zážitcích.¹⁹

Obrázek 1: Význam loga Airbnb – „Belong Anywhere“



Zdroj: <https://design.studio/work/airbnb>

Airbnb tedy své služby komunikuje tak, že kromě samotného ubytování nabízí přidanou hodnotu, a těmi jsou právě zmíněné hodnoty postmateriální. Pokud by se jejich přísliby odrážely i ve zkušenostech uživatelů, pak by se jim mělo dostávat osobnějšího vztahu i přístupu, autentického zážitku z místa i ubytování a celkově vřelého pocitu jako doma, pro což na tradičním trhu obvykle není místo [Liu a Mattila 2017, str. 34]. Na příkladu Airbnb tak demonstrují, jak konkrétně mohou být komunikovány postmateriální hodnoty ve sdílené ekonomice a jak jsou užívány pro odlišení od tradičních modelů.

¹⁹ Airbnb věnuje značné úsilí do vytváření pocitu komunity a partnerství, stejně jako do šíření osvědčených postupů [Sundararajan 2014; Yannopoulou 2013]. Jsou pořádána pravidelná setkání hostitelů a je pro ně vytvořena i speciální platforma „Airbnb Community“, kde mezi sebou mohou sdílet zkušenosti a tipy.

1.3.2 Kontroverznost Airbnb a situace v České republice

Airbnb společně s Uberem²⁰ se v posledních letech staly předmětem značné kontroverze. Oba projekty se rozrostly do takových rozměrů, že začaly vyvstávat otázky nad jejich vedlejšími dopady a nutností jejich regulace. Jak jsem popsala výše (viz kapitola 1.1, str. 3), s rozrůstáním sektoru sdílené ekonomiky se rozpíná i jeho význam, dochází k odklonu od původních ideálů o čistých transakcích mezi lidmi a prosazuje se element kapitalistické ekonomiky, motiv generovat zisk. Ozývají se hlasy, že praxe ubytovávání přes Airbnb je jiná, než je jím strategicky komunikováno a že jakékoliv vyšší hodnoty se převážně vytratily a zůstal jen neregulovaný podnikatelský byznys, který nelze označit jako sdílená ekonomika [Ključnikov et al. 2018; IPR 2018]. V této části se proto zaměřuji na český kontext fungování a kritiky Airbnb.

Airbnb proniklo na český trh roku 2009 a od roku 2012 každoročně roste počet registrovaných nových ubytovacích kapacit exponenciálním tempem [Úřad vlády ČR 2017, str. 46; Deloitte 2019]. Nabídka Airbnb se rozšiřuje na téměř celé území České republiky, zejména do nejnavštěvovanějších destinací jako Český Krumlov, Karlovarsko, Valticko, zdaleka největší koncentrace ale přirozeně zůstává v Praze [Úřad vlády ČR 2017, str. 46]. Tam, podle výpočtů Deloitte [2019, str. 13], mělo Airbnb v roce 2018 již dokonce 41% podíl na celkových ubytovacích kapacitách, přičemž dlouhodobě dochází k jeho růstu, zatímco hotely kapacitně spíše stagnují [ibid., str. 16].²¹ Podobně je na tom i obsazenost [ibid., str. 18].²² Airbnb tedy neukrádá z podílu tradičních ubytovacích kapacit, jejich pozice zůstává stejná, ale více benefituje z celkového růstu turismu v Praze [Makovský 2017, str. 147].²³ Hoteliři nicméně konkurenci Airbnb často považují za nekalou. Jejich hlavním argumentem je odklon Airbnb od

²⁰ Uber je technologická společnost propojující lidi za účelem přepravy, fungující od roku 2009. Počet řidičů přes Uber se mezi lety 2013 a 2015 každých šest měsíců zdvojnásobil [Harris 2017]. Tržní hodnota Uberu neustále roste, v roce 2018 byla 76 miliard dolarů [Statista 2020a].

²¹ Mezi lety 2016 a 2018 došlo u Airbnb k 34,1% růstu, zatímco růst kapacit hotelů byl pouze 1,4 %. Stejný trend se mezi zkoumanými městy ukázal také v Barceloně, ještě výrazněji pak v Budapešti a Vídni. Opačný trend je pak v Amsterdamu, kde kapacity hotelů rostou, zatímco Airbnb klesá.

²² Obsazenost Airbnb mezi lety 2017 a 2018 vzrostla z 67,7 % na 70,5 %, zatímco hotely zažily zanedbatelný pokles z 69,5 % na 69,2 %. Tím je situace v Praze specifická oproti dalším zkoumaným městům. V Barceloně stagnují jak hotely, tak Airbnb. V Amsterdamu obsazenost hotelů vyloženě klesla, zatímco Airbnb rostlo. Ve Vídni a Budapešti je pak rostoucí trend u obou modelů ubytování.

²³ Dá se nicméně očekávat, že bez existence alternativy ubytování přes Airbnb by hotely rostly, ačkoliv Airbnb samo tvrdí, že díky jejich službě cestují lidé, kteří by jinak necestovali. Zervas [et al. 2017] odhaduje, že vstup Airbnb na trh má negativní dopad cca 8-10 % na výši tržeb hotelů. Největší vliv má zejména na hotely levnější a ty, které nezajišťují služební cesty. Václav Stárek, prezident Asociace hotelů a restaurací ČR, nicméně pro ČT24 [Ginterová 2018] uvedl, že ačkoliv v Praze hotely kapacitně stagnují, zisky jim přesto rostou díky vyšším cenám.

původních ideálů sdílené ekonomiky a **výrazná profesionalizace služeb, které se tak údajně přibližují k tradičním modelům ubytování**, aniž by toto bylo reflektováno regulacemi, kterým podléhá jejich podnikání, což považují za znevýhodnění [Úřad vlády 2017, str. 39]. Služba Airbnb se skutečně výrazně posunula od první nabídky, kterou byla nafukovací matrace v bytě zakladatelů – nyní markantně převažuje soukromé ubytování v samostatných bytech. V Praze je aktuálně k dispozici 13,5 tisíce nabídek přes Airbnb a **podíl celých bytů tvoří cca 80 %** [Insideairbnb].²⁴ Zvyšuje se zároveň podíl hostitelů, kteří mají zaregistrovaných vícero nabídek najednou, což je opět krok směrem od idealistických předpokladů. V Praze takoví hostitelé tvoří 65% podíl. Největší část z nich (20 %) má nabídky jen dvě, pak ale například hostitel s profilem Prague For You jich má aktivních dokonce 128 [ibid.]. Trh s Airbnb tak kontroluje ze 40 % pouze malá část profilů (662, tj. 7,4 %). Na druhou stranu takřka polovina uživatelů si přes Airbnb vydělá pouze přibližně dvanáct tisíc ročně [Ključnikov et al. 2018, str. 133-134].²⁵ Hlavní potíží při snaze nastavit regulace se tak opět stává najití hranice mezi sdílenou a tradiční ekonomikou a otevírá se výše nastolený problém, zda sdílenou ekonomiku odvozovat od povahy transakce či od motivů na její participaci (viz široká a úzká definice, kapitola 1.1, str. 3). Institut plánování a rozvoje hl. m. Prahy (IPR) [2018, str. 3] ve své zprávě například píše, že pokud prostor není fyzicky obýván s hostem, pak se nejedná o sdílenou ekonomiku. Podobný přístup volí i Ključnikov [et al. 2018], který píše, že 80 % nabídky v Praze lze označit za podnikatelskou aktivitu, nikoliv sdílenou ekonomiku, respektive příležitostné pronájmy.

V českém kontextu se tedy prosazuje perspektiva, že drtivá většina nabídky přes Airbnb nemůže být považována za sdílenou ekonomiku, jelikož host neobývá jeden společný prostor s hostitelem.²⁶ Znamená to tedy, že se vytratily postmateriální hodnoty spojované s transakcemi

²⁴ Z toho důvodu se také vedou debaty o potenciálním dopadu Airbnb na bytovou situaci v Praze, která je velmi špatná. Bydlení se stává stále méně dostupným pro nedostatek kapacit i vysoké ceny. Vliv Airbnb se nicméně jednoznačně nepotvrdil. Radní pro město Praha Hana Kordová Marvanová a člen výboru pro bytovou politiku Vojtěch Munzar se shodují, že na nedostatečnost bytů a jejich vysoké ceny mají vliv spíše jiné faktory jako celková změna životního stylu společnosti, kdy lidé déle žijí sami, je vyšší rozvodovost, migrace a je třeba především zvýšit počet povolených staveb, podporovat dohody města s developery, uvolnit nevyužité pozemky, podpořit družstevní bytovou výstavbu a uvolnit nevyužívané městské byty [ČT24 2020]. Airbnb dle nich hraje jen minimální roli. Závažnější dopady jsou spíše na život residentů (hluk, nepořádek, neklid) a podobu města (komodifikace každodenního života, centrum skanzenem, turistifikace) [ibid.; Sans a Domínguez 2016, str. 211-212].

²⁵ Dle dat Airbnb typický český hostitel pronajímá svůj prostor 37 dní v roce a vydělá si cca čtyřicet sedm tisíc korun [Ginterová 2018].

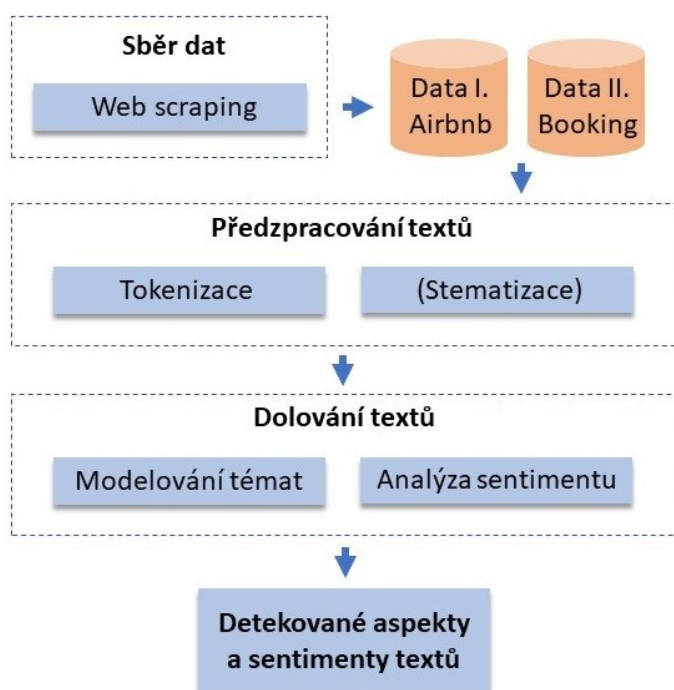
²⁶ Tato otázka bývá řešena zejména v souvislosti s nastavením regulací za účelem přesunu sdílené ekonomiky z tzv. šedé zóny. Pro mou práci je ale tato snaha klasifikovat sdílenou ekonomiku v porovnání s ekonomikou tradiční zajímavá spíše z hlediska spotřebitelské zkušenosti a naplňování potřeb, respektive hodnot než z hlediska legislativního.

v rámci sdílené ekonomiky a spotřebitelská zkušenost se příliš neodlišuje od zkušenosti s tradičním sektorem? S touto otázkou vstupuji do výzkumné části práce.

2 Metodologie a plán výzkumu

Metodologická část práce obsahuje plán výzkumu navazující na výše představená teoretická východiska. Vymezuji zde cíl výzkumu, včetně výzkumných otázek, dále popisuji zvolené analytické metody a typ dat. Jednu kapitolu také věnuji sběru dat, jejich přípravě a validizaci před analýzou. V závěru diskutuji etické hledisko zvolených postupů. Postup je shrnut graficky na Obrázku 2.

Obrázek 2: Plán výzkumu graficky



Zdroj: vlastní zpracování

2.1 Vymezení cílů výzkumné části

Hlavním cílem je prozkoumat, jak a zdali vůbec se liší dimenze, v nichž lidé popisují své zkušenosti s ubytováním se v Praze v sektoru sdílené ekonomiky v porovnání se zkušenostmi lidí ubytovaných v tradičních hromadných ubytovacích zařízeních také v Praze, a to v obou případech za poslední dva roky. K výzkumu tedy přistupuji exploračně se záměrem ze sociálněvědního hlediska přispět k porozumění tomu, proč je tento nový typ ekonomiky (nejen) v Praze tak úspěšný a jaké hodnoty v oblasti ubytování přináší navíc oproti tradičnímu sektoru z perspektivy spotřebitelské zkušenosti.

Jak bylo nastíněno v předchozích částech práce, diskurz ohledně Airbnb je momentálně velmi živý. V Praze a dalších mnoha metropolích si vydobylo významný a dále rostoucí podíl na trhu s ubytováním. Jeho úspěch bývá vysvětlován různě a z mnoha pozic. Mason [2016], Botsman [2018], David [2017], Sundararajan [2016] a další autoři ve svých teoriích argumentují úspěch sdílené ekonomiky obecně tím, že je schopna uspokojit potřeby dnešního, postmateriálně orientovaného konzumenta žijícího v západní společnosti lépe než projekty tradiční ekonomiky. Na druhou stranu se ale prosazuje perspektiva, že Airbnb nadále nemá nic společného s ideou sdílené ekonomiky a že jeho hlavní konkurenční výhodou a důvodem růstu je možnost nastavit nižší ceny díky nedostatku regulací [Ključnikov et al. 2018; IPR 2018; Úřad vlády 2017, str. 9]. Podle tohoto argumentačního tábora se jinak jedná o ubytovací kapacitu podobnou hotelu, jelikož už převládají samostatné jednotky, nikoliv fyzicky sdílené. Společným znakem těchto studií nicméně je, že pozornost věnují pouze straně nabídky na Airbnb, perspektivu spotřebitele spíše přehlíží. Dosud nejsou dostatečně empiricky prozkoumány konkrétní aspekty, v čem se uživatelská zkušenost s Airbnb liší od tradičního ubytování. Valná většina takto orientovaných prací je z amerického prostředí a jejich výsledky se ne úplně shodují [Cheng a Jin 2019, str. 59; Lawani et al. 2019; Tussyadiah a Pesonen 2016 a další]. V českém prostředí takové výzkumy takřka chybí, ačkoliv je Airbnb pro Prahu obrovským tématem. Zájem návštěvníků Prahy o Airbnb roste každým rokem, zatímco obsazenost hotelů stagnuje či mírně klesá [Deloitte 2019]. Místo na příčiny je ale pozornost směřovaná spíše na dopady a potenciální možnosti regulací. Těší se Airbnb takovému zájmu skutečně jen díky nižší ceně nebo nabízí i něco navíc oproti hromadným ubytováním? Pro přispění porozumění tomuto novodobému fenoménu formuluji následující výzkumné otázky.

VO1: *Jak se liší zkušenosti uživatelů s ubytováním se v sektoru sdílené ekonomiky a tradiční ekonomiky z hlediska slovníku, který užívají pro popis svých zkušeností?*

VO2: *Jaké jsou základní dimenze, ve kterých uživatelé Airbnb popisují svou zkušenost s ubytováním se v sektoru sdílené ekonomiky v Praze?*

VO2.1: *Jak významně jsou jednotlivé dimenze zastoupeny ve zkušenostech uživatelů Airbnb?*

VO2.2: *Nakolik se v těchto dimenzích objevuje pozitivní a nakolik negativní sentiment?*

VO3: *Jaké jsou základní dimenze, ve kterých uživatelé tradičních ubytovacích kapacit (hotelů, hostelů, penzionů) popisují svou zkušenost s ubytováním se v Praze?*

VO3.1: *Jak významně jsou jednotlivé dimenze zastoupeny ve zkušenostech uživatelů tradičních ubytovacích kapacit?*

VO3.2: *Nakolik se v těchto dimenzích objevuje pozitivní a nakolik negativní sentiment?*

VO4: *Jak se liší detekované dimenze zkušeností s ubytováním se v sektoru sdílené ekonomiky a tradiční ekonomiky?*

VO5: *Jak se liší celkové zastoupení sentimentů a emocí ve zkušenostech s ubytováním se v sektoru sdílené ekonomiky a tradiční ekonomiky?*

Pro odhalení specifických zkušeností s ubytováním v sektoru sdílené ekonomiky volím komparativní přístup. Porovnávám uživatelské zkušenosti s Airbnb oproti zkušenostem s tradičními hromadnými ubytovacími zařízeními, tedy hotely, hostely a penziony. Geograficky je výzkum zaměřen na Prahu.²⁷ Cílem je především poskytnout vhled do doposud v českém kontextu opomíjené problematiky a podkrýt souvislosti výrazného růstu Airbnb na trhu s ubytováním v Praze z pohledu spotřebitele. Není cílem výsledky zobecňovat, zvolený přístup je popisný či explorační.

²⁷ Je běžnou praxí, že jsou empirické studie o Airbnb (a obdobných projektech) zasazovány do kontextu konkrétní destinace [Lawani et al. 2019; Horn a Merante 2017 a další]. Každá totiž má svá specifika. Situace v Praze, kam je zasazena empirická část této práce, je detailně popsána v kapitole 1.3.2 na str. 14.

2.2 Data

Za účelem komparace podrobují analýze dva korpusy textů v podobě slovních výpovědí samotných uživatelů o jejich zkušenostech s ubytováním, a to jak přes Airbnb, tak v tradičních ubytovacích kapacitách. Oba korpusy jsou souborem online volně dostupných slovních recenzí jednotlivých ubytování získaných z webů metodou web scraping. První korpus obsahuje recenze ubytování ze sektoru sdílené ekonomiky z platformy Airbnb, korpus referenční pak ze sektoru tzv. tradiční ekonomiky, tedy hotelů, hostelů a penzionů z platformy Booking. Zařazeny jsou výpovědi těch uživatelů, kteří komunikují své zkušenosti v angličtině.²⁸ V obou případech se jedná o slovní recenze všech ubytovacích kapacit aktivních v době sběru dat v lokalitě Praha za poslední dva roky od sběru dat (listopad 2019).²⁹ Slovní recenze zůstávají stále relativně málo probádané, ačkoliv bylo potvrzeno, že jsou kvalitním zdrojem pro vhled do zkušeností uživatelů [Perikos et al. 2018; Liu et al. 2017; Xiang et al. 2015, str. 122; Dhuan a Whinston 2008 a další] s výrazným vlivem na další poptávku [Perikos et al. 2018; Ady a Quadry-Felitti 2015; Xiang et al. 2015]. V této kapitole popisují zdroje těchto dat, způsob jejich sběru a postup jejich validizace.

2.2.1 Zdroj dat, jejich výhody a alternativy

Vytvořené datové korpusy tvoří rozsáhlý počet slovních recenzí v počtu v řádu statisíců, které neustále přibývají v digitalizované podobě na webových stránkách daných ubytování. V této subkapitole diskutují výhody jejich užití, jejich vhodnost a případné alternativy.

Již před čtyřiceti lety charakterizoval Luhmann [1988] moderní společnost jako systém intenzivně produkující sebepopisy. A to patrně ani nemohl tušit, do jakých rozměrů se tento rys rozroste s rozmachem internetu, rozvojem technologií a s tím související digitalizací. Podle reportu IDC [2018] z roku 2018, 90 % celosvětových dat bylo vyprodukováno za poslední dva

²⁸ Problém jazykové diversity je více rozvinut v kapitole 2.3.2 o přípravě dat na str. 26.

²⁹ Vzhledem k povaze výzkumu, tedy komparaci, jsem stanovila jako konstantní proměnné lokalitu a časové období, kdy jejich zkušenost probíhala. Tyto proměnné pojí zkušenosti jednotlivých uživatelů, protože obecně nastavené standardy v ubytováních mohou být jiné napříč světovými destinacemi i v čase, a tím by se srovnávání zkušeností komplikovalo.

roky.³⁰ Internet se stal největším zdrojem informací, který kdy člověk měl a pokračuje v růstu [Weare a Lin 2000, str. 276]. Jsou to například příspěvky a profily na sociálních sítích či ve fórech, články, videa, recenze, komentáře, reakce na příspěvky a profily druhých aj. [Hájek 2014, str. 15] Značná část je navíc veřejně přístupná a snadno dohledatelná. Všechny tyto texty jsou popisem společnosti, ve které žijeme. Vzhledem k novosti tohoto fenoménu ovšem řada z nich zatím zůstává analyticky nevyužita.

Také pro zodpovězení výzkumných otázek této práce lze sebrat již existující, digitalizovaná a veřejně dostupná data. Taková, která navíc zůstávají v českém kontextu nedostatečně probádaná. A sice slovní recenze jednotlivých ubytování v Praze dostupné na webových stránkách jejich největších zprostředkovatelů. Těmi jsou za sektor sdílené ekonomiky Airbnb.com, za tradiční sektor Booking.com.³¹ Systém uživatelských recenzí je u Airbnb a Bookingu velice podobný. Obě platformy e-mailem posílají odkaz na dotazník každému uživateli, který se přes ně ubytoval. Dotazník obsahuje mj. otevřenou otázku, kde mohou volně popsat svou zkušenost s ubytováním a následně jsou všechna tato získaná data zveřejňována online u příslušných ubytování.³² Oba zprostředkovatelé ubytování garantují, že se jedná o autentická hodnocení a jsou zveřejňována všechna.³³ Dotazník uživatelé mohou vyplnit nejpozději do dvaceti osmi dnů na Bookingu a do čtrnácti na Airbnb od opuštění ubytování, jedná se tak o jejich bezprostřední zkušenost a nevznikají tam problémy s pamětí. Zároveň jsou osloveni opravdu všichni ti, kdo byli ubytováni, jelikož při registraci musejí zadat

³⁰ Podle jejich propočetů bylo v roce 2018 celosvětově celkem přibližně 33 zettabajtů dat [IDC 2018]. Do roku 2025 je IDC očekáván závratný nárůst až na 175ZB, tedy roční růst 61 %. Jeden zettabajt je trilion gigabajtů.

³¹ Ani jedna z platform nejsou vlastníky na nich nabízených ubytování, pouhými zprostředkovateli. Financovány jsou z procentuální provize za zprostředkování, která se v Česku pohybuje kolem dvaceti procent [Ginterová 2018].

³² Kromě prostoru pro slovní hodnocení mají uživatelé možnost udělit také číselné hodnocení, které se u Airbnb pohybuje na stupnici 1-5 s přesností na setiny, u Bookingu na stupnici 1-10 s přesností na desetinná místa. V případě číselného hodnocení mají přímo určeno, které oblasti mají hodnotit a následně hodnotí i celkovou zkušenost. Uživatelé Airbnb mají možnost v rámci hodnocení zanechat i neveřejnou zpětnou vazbu přímo hostiteli, to u Bookingu možné není. Na Airbnb, vzhledem k povaze ubytování, kdy hostitel pouští hosta do svého vlastního domu či bytu, je hodnocení reciproční, tedy i hostitel hodnotí hosta. Nicméně hodnocení se oběma stranám zobrazí až ve stejnou chvíli po uplynutí čtrnácti dnů od odjezdu hosta, aby bylo zamezeno vzájemnému ovlivňování se. Na platformě Airbnb i Booking mají poskytovatelé možnost komentářem reagovat na získaná hodnocení. Dotazníky jsou k nahlédnutí v Příloze 3 na str. 101 pro Airbnb a v Příloze 4 na str. 103 pro Booking.

³³ Oba zprostředkovatelé uvádí ve svých pravidlech, že jedinou výjimkou, kdy by recenze nebyla zveřejněna nebo byla smazána, je hrubé porušení pravidel jako vulgarismy, vydírání apod. Zároveň oba připouštějí možnost, že v extrémních situacích mohou provozovatelé ubytování požádat zprostředkovatele o smazání jedné konkrétní recenze, pokud k tomu mají závažné důvody a případně důkazy (uživatel se do ubytování ani nedostavil apod.). Na stránkách jednotlivých ubytování by tedy měla být veřejná všechna hodnocení, která ubytování získala, a sice na Airbnb historicky, na Bookingu za posledních 24 měsíců.

svůj e-mail. Oslovit takto různorodou cílovou skupinu by jinak bylo velice problematické v případě alternativních možností získání dat jako jsou dotazníky či rozhovory.

Užitečnost recenzí pro poskytnutí unikátního vhledu do uživatelských zkušeností již byla potvrzena mnoho výzkumnými studiemi [Zhang 2019; Perikos et al. 2018; Tussyadiah a Zach, 2017; Xiang et al. 2015; He et al. 2013; Dhuan a Whinston 2008 a další]. Obecně jsou uznávány jako jeden z nejdůležitějších obsahů pro porozumění zákaznického chování a spokojenosti [Xiang et al. 2015, str. 122]. Jak píše Xiang [et al. 2015], recenze „*reflektují způsob, jakým zákazníci popisují, znovu zažívají, rekonstruují a sdílejí své zkušenosti*“ [str. 122, přeloženo autorkou]. Zároveň jsou důležité i v rozhodovacím procesu zákazníka [Perikos et al. 2018; Ady a Quadry-Felitti 2015; Sparks a Browning 2011; Lee a Shin 2014]. Perikos [et al. 2018] také píše, že jsou bohatým zdrojem informací o základních attributech poskytované služby a názoru zákazníka na ně. Zdají se tedy být výborným zdrojem pro naplnění vymezených cílů. Jejich užití nicméně není neproblematické. Jejich obsah může být ovlivňován vnějšími faktory jako jsou již předchozí recenze, kdy se nový uživatel může inspirovat dříve použitými formulacemi nebo může mít tendenci přizpůsobovat svou rétoriku rétorice nabídky. Taková chybovost je těžko vyčíslitelná. Chybám by nicméně nebylo možné se vyhnout ani při volbě jiných postupů získání dat jako jsou dotazníky či rozhovory s uživateli o jejich zkušenostech. Tam by mohlo docházet k ovlivnění respondenta celou výzkumnou situací, chybám při vytváření výběrového souboru a dalším.³⁴ K datům a interpretaci výsledků každopádně přistupuji kriticky a s uvědoměním si jejich limitů, včetně nemožnosti generalizace výsledků.

2.2.2 Sběr dat

Technikou sběru dat byl v obou případech tzv. web scraping, což je automatizovaný proces extrakce a uložení velkého množství informací z webových stránek na základě algoritmu [Parvez a kol. 2018]. Zdrojem mi byly internetové platformy, které zprostředkovávají tyto dva typy ubytování. Pro ubytování ze sektoru sdílené ekonomiky je to Airbnb.com, celosvětově zdaleka největší zprostředkovatel sdíleného ubytování a také jeho průkopník (viz kapitola 1.3,

³⁴ Použít jiný typ sběru dat by nicméně bylo vhodným postupem pro validizaci výsledků této práce. Vzhledem k náročnosti a omezenému prostoru diplomové práce je ovšem realizace ponechána případným dalším výzkumníkům.

str. 11). Pro hotely, hostely a penziony pak Booking.com, největší platforma sdružující nabídku ubytování v tradičních ubytovacích kapacitách [Trajkovska 2018]. Pro Airbnb jsem využila data, která již byla shromážděna a jsou k dispozici na webu Insideairbnb, zatímco pro Booking jsem je shromažďovala samostatně. Data jsem následně kontrolovala ručně a v případě Bookingu také oproti databázi hromadných ubytovacích zařízení v Praze od Českého statistického úřadu [2019]. Jelikož se postupy získání i následné validizace dat z platform Airbnb a Booking liší, popisuji je také odděleně.

A. Airbnb

Datový korpus za Airbnb je tvořen daty sebranými Murray Coxem, počítačovým expertem a specialistou na technologie. Cox pravidelně každý měsíc stahuje snad všechna veřejně přístupná data z webu Airbnb za vybrané metropole, včetně Prahy. Následně je pouze v očištěné podobě poskytuje volně veřejnosti na svém webu Insideairbnb. Tato data jsou běžně využívána v akademických výzkumech [Lawani et al. 2019; Cheng a Jin 2019, Xu et al. 2019 a další]. Po pečlivém ověření kvality tato data užívám i pro účely této práce.³⁵ Využívám nejaktuálnější data, která byla momentálně k dispozici, a sice z listopadu 2019. V datovém souboru jsou tedy zahrnuta ta ubytování, která byla v tomto období aktivní a dostupná na platformě Airbnb pro lokalitu Praha. Celkově se jedná o více než čtrnáct tisíc nabízených ubytování, pro které byly staženy všechny recenze. Součástí datového souboru z Insideairbnb je pro každé ubytování i jeho URL, je tedy možné se prokliknout na jeho webovou stránku a zkontrolovat, zda informace v datovém souboru odpovídají údajům o ubytování online. Ručně jsem kontrolovala vzorek sta náhodných ubytování a nikde jsem neshledala chybu nebo vynechání údajů. Datový soubor tedy považuji za kvalitní a používám k analýze.

B. Booking

Projekt Insideairbnb, ze kterého jsem čerpala data za Airbnb, je spíše výjimkou. Pro data za Booking se mi nepodařilo žádný spolehlivý zdroj nalézt. Data z platformy Booking jsem tedy sbírala samostatně s využitím zmíněné techniky web scraping. Užila jsem software Octoparse,

³⁵ Stejná data bych mohla potenciálně sbírat i sama při použití obdobného postupu jako u Bookingu, nicméně vzhledem ke kvalitě a rozsahu dat z Insideairbnb jsem to shledala zbytným.

který pracuje s jazykem X-path. S jeho využitím jsem softwaru určila cestu k potřebným datům (elementům HTML), a tak označila části webu určené ke stažení. Následně jím byly tyto kroky vykonávány automatizovaně pro každé jednotlivé ubytování. Data byla sbírána také v listopadu 2019.³⁶ Zahrnuta jsou ubytování, která byla v tomto období dostupná v Praze na platformě Booking a spadá do kategorie hotel, hostel nebo penzion, tedy hromadné ubytovací kapacity. Celkem se jedná o téměř sedm set ubytovacích zařízení. Data jsem validizovala oproti databázi hromadných ubytovacích zařízení v Praze k 18.11.2019 [ČSÚ 2019].³⁷ Mohla jsem tak zjistit, kolik z celkového počtu hromadných ubytování v Praze je registrováno na Bookingu, a tedy o jakém podílu z nich jsou dostupná data pro analýzu.³⁸ Z téměř sedmi set ubytování jich na Bookingu nebylo registrovaných jen padesát šest. Dalších jedenáct registrovaných bylo a byly k nim k dispozici recenze z minulosti, nicméně v době stahování dat nebyly aktivní, a proto je nezahrnuji, aby byly dodrženy podmínky pro sbírání dat u obou datových souborů stejné. Celkem tedy není v datovém souboru za Booking zahrnuto přibližně jedno procento na ČSÚ registrovaných hromadných ubytování v Praze, zbytek ano.

Ve výsledku jsem tedy vytvořila dva datové soubory obsahující slovní recenze ubytování za poslední dva roky. Data jsem sbírala z webových stránek zprostředkovatelů Airbnb a Booking v průběhu listopadu 2019. Zahrnuta jsou v obou případech ta ubytování, která byla v době sběru dat aktivní v lokalitě Praha.

³⁶ Konkrétně v období od 3. do 22.11.2019. Naprosto ideální by samozřejmě bylo provést sběr dat v jednom časovém okamžiku, protože webové stránky jsou neustále aktualizovány, nicméně vzhledem k využití techniky pouze průměrného výkonu a neustálého průběžného kontrolování úplnosti stahovaných dat trval sběr téměř tři týdny takřka nepřetržitě.

³⁷ Databáze zahrnuje také údaje o studentských kolejkách a firemních ubytovnách, ty nejsou zahrnuty do analýzy.

³⁸ V některých případech jsou v databázi ČSÚ ubytovací zařízení registrována pod jiným názvem než na Bookingu, taková byla kontrolována podle adresy.

2.3 Metody a postupy

Data jsem analyzovala s využitím metod pro analýzu textů. Pracovala jsem tedy se dvěma rozsáhlými korpusy nestrukturovaných textových dat – prvním tvořeným slovními recenzemi ubytování přes Airbnb a v druhém případě recenzemi ubytování v hromadných ubytovacích kapacitách. Po pročištění dat a převedení do modelu bag-of-words³⁹ jsem na oba korpusy aplikovala stejné postupy. Kromě prozkoumání souborů z hlediska poměrného zastoupení jednotlivých výrazů jsem užila také metodu Structural topic modeling pro detekování opakujících se témat, výrazů je tvořících a jejich zastoupení v celkovém textu. Korpusy textů jsem zkoumala také s využitím analýzy sentimentu jak na úrovni dokumentů, tak témat. Výsledky jsem komparovala. Účelem této kapitoly je odůvodnit tyto volby a představit zvolené metody.

2.3.1 Text mining

Systematická analýza obsahu textů byla ustanovena coby metoda již v padesátých letech minulého století Berelsonem [1952].⁴⁰ Stala se běžně užívanou technikou pro kvantifikaci nestrukturovaných dat jako jsou rozhovory, terénní poznámky, produkce médií a mnohé další [Kronick 1997, str. 57]. S exponenciálním růstem digitalizovaných dat, kde až osmdesát procent z nich má podobu textu [IDC 2018], dále nabývá na významu [Roberts et al. 2019, Weare a Lin 2000, str. 272]. Nicméně jedná se o texty v takovém množství, že jejich ruční čtení a kódování již není v silách výzkumníka. Zapojují se proto významně více počítače a rozvíjí se tak tzv. text mining, tedy těžení z nestrukturovaných textů, zahrnující různé techniky, jak získat užitečné informace z objemného datového souboru s pomocí algoritmů [Hájek 2014, str. 196-7]. Množství i pokročilost těchto technik roste každým dnem a jejich užívání se stává novým trendem v datové analýze [Downer et al., 2019]. Hlavními výhodami těžení textů jsou vyšší

³⁹ Model, kde dokument je reprezentován jednotlivými slovy, u kterých není důležité jejich pořadí, nýbrž jejich četnost a spoluvýskyty.

⁴⁰ Berelson definuje obsahovou analýzu jako výzkumnou techniku určenou „(...) k objektivnímu, systematickému a kvantitativnímu popisu zjevného obsahu komunikace“ [Berelson 1952, str. 18]. Vyvolává ovšem diskuze a řada autorů se k ní již nehlásí pro náročný požadavek objektivitu i problém vypořádat se s tím, co je vlastně obsah, zejména pokud se jedná o analýzu komunikace [Hájek 2014, str. 57-61]. Dnes je proto často užíván termín „textová analýza“.

efektivita, odhalování skrytých informací ve velkém rozsahu, odkrývání nové perspektivy, eliminování lidských chyb a preciznost, umožnění srovnání výsledků při aplikaci stejných postupů [McDonald a Kelly 2012]. Nevýhody přirozeně vyplývají ze samotného principu automatizace, na kterém tyto techniky stojí. Stroj není schopen zachytit ironii či číst mezi řádky, přistupuje k textu jako k surovému předmětu [Hájek 2014, str. 17]. Může se tak stát, že počítači unikne informace, kterou by lidský čtenář nepřehlédl. Je proto důležité věnovat zvýšenou pozornost validitě výsledků a vhodnosti dat pro jednotlivé techniky.

Slovní recenze, které jsou předmětem analýzy v této práci, se počtem pohybují v řádech statisíců a mají podobu nestrukturovaných, textových dat. Pro dosažení cíle výzkumu je třeba získat z nich informaci o důležitých aspektech v nich popisované služby a tyto následně porovnávat. Z těchto důvodů jsem pro dosažení cílů zvolila přístup dolování z textů, a to konkrétně techniku pro strukturní modelování témat, s jejímž využitím je možné najít nejkoherentněji zastoupená témata v textu, kterou doplňuji o analýzu sentimentů. Aby bylo možné text tímto způsobem analyzovat, je ale nutné jej nejprve převést do vhodné podoby, tzv. modelu bag-of-words.

2.3.2 Příprava dat

Předzpracování textu umožňuje produkovat kvalitnější klasifikace [Perikos et al. 2018]. Nejprve jsem texty očistila o jiné jazyky než angličtinu a následně převedla do modelu bag-of-words pro jeho jednoduchost a vhodnost pro klasifikační procedury. To především znamená, že text je tokenizován, v tomto případě do slov. Vhodnost dalších úprav se liší podle jednotlivých procedur, ale obecně platí, že je vhodné snížit dimenzionalitu textu pro efektivnější a přínosnější práci s ním, tedy bývají odstraňována tzv. stop slova, slova s velmi nízkou frekvencí výskytu a pro některé postupy je vhodná také stematizace.

Jak píše Hájek [2014, str. 196-8], jednou ze skutečností, se kterými je nutné se vypořádat při dolování rozsáhlých textů v podobě tzv. velkých dat, je jazyková diverzita. Algoritmus pracuje se stejným slovem v jiném jazyce jako se slovem odlišným. Vzhledem k tomu, že cílovou populací této práce jsou všichni z celého světa, kdo se ubytovali v Praze přes Airbnb či v hotelu, hostelu a penzionu přes Booking, recenze, ve kterých popisují své zkušenosti, jsou

psány v různých jazycích. Stálo tedy přede mnou rozhodnutí, jak s tímto naložit. Ruční překlad všech recenzí do jednoho stejného jazyka není možný pro velikost datového souboru, ale i mé nedostatečné jazykové znalosti. Automatizovaný překlad zatím není možné považovat za spolehlivý, má výrazné nedostatky, a tím by značně utrpěla validita výsledků. Navíc, jak píše Hájek, „(...) z analytického hlediska neexistuje překlad, jehož výsledkem by byl identický text ve dvou jazykových formách“ [Hájek 2014, str. 198]. Rozhodla jsem se tedy následovat postup již existujících empirických výzkumů užívajících techniky dolování z textů [Zhou et al. 2018; Cheng a Jin 2019; Lawani et al. 2019] a pro analýzu jsem vybrala nejzastoupenější jazyk v obou datových souborech, kterým je angličtina.⁴¹ V případě Airbnb tvoří angličtina dokonce 80 % veškerých recenzí, u referenčního korpusu z Bookingu je to cca 36 %.⁴² U Airbnb je druhým nejzastoupenějším jazykem francouzština (5 %), dále vietnamština (4 %), španělština (3 %), němčina (3 %) a italština (2 %). U Bookingu je to němčina (16 %), čeština (15 %), italština (8 %), španělština (7 %) a francouzština (5 %). Všechny zbylé jazyky tvoří jen malou část, a sice u Airbnb 3 %, u Bookingu 13 %. Recenze v dalších jazycích jsou ponechány ke zpracování dalším výzkumníkům. Nepochybně by bylo zajímavé pak porovnat výsledky, zdali se mezi angličtinou a dalšími jazyky liší. Jazyk jednotlivých recenzí byl automatizovaně detekován s pomocí Google Spreadsheets a následně jsem vyfiltrovala pouze recenze v angličtině.⁴³ Celkem tedy dále pracuji s více než 300 000 recenzemi z Airbnb a více než 120 000 recenzí z Bookingu.

V dalším kroku jsem soubory převedla do modelu bag-of-words, tedy jeden token na řádek, který je vhodný pro zvolené analytické postupy [Silge a Robinson 2017]. Token je jedna smysluplná jednotka textu. Jednotkou v tomto případě je slovo. Proces převedení volného textu do jednotlivých slov se tak nazývá tokenizace.⁴⁴ Dále byla odstraněna čísla a tzv. stop slova,

⁴¹ Rozhodnutí by patrně mohlo být problematictější u menších datových souborů, nicméně v tomto případě i po filtraci všech ostatních jazyků kromě angličtiny, jsou datové soubory tvořeny množstvím recenzí v řádu statisíců. Je tedy důležité spíše klást důraz na to, že je pracováno se zkušenostmi pouze jasně vymezené skupiny těch, kdo je píšou v angličtině, a není možné výsledky generalizovat na všechny.

⁴² Rozdíl může být v tom, že u hotelů jsou většinou ubytovaní zvyklí, že na recepci mluví jejich jazykem, zatímco u Airbnb je ubytovává hostitel komunikující v mezinárodním jazyce, nikoliv profesionální jazykově vybavený recepční. Zároveň typický uživatel Airbnb je zpravidla mladší než uživatel hotelů a mladší lidé častěji ovládají angličtinu jako druhý jazyk než lidé starší. Tím, že z obou souborů jsou vybrány pouze recenze v angličtině, se dá očekávat, že skupina recenzentů bude koherentnější, tedy potenciálně pro komparaci vhodnější, ovšem to jsou pouze odhady a bylo by do budoucna velmi přínosné zanalyzovat i další jazyky a výsledky porovnat.

⁴³ V případě Airbnb jsem odstranila také automatické zprávy o zrušení ubytování hostitelem, které se vkládají mezi recenze a pak také automatické zprávy o tom, že Airbnb něco skrylo v rámci jednotlivé recenze, zpravidla telefonní číslo, webový odkaz nebo e-mail. Tento text není součástí zkušeností hostů a není pro analýzu zajímavý.

⁴⁴ V tidy formátu, který byl zvolen pro tuto práci s daty, jsou při tokenizaci automaticky také převedena všechna písmena na malá a jsou smazány všechny symboly v textech.

tedy taková slova, která nenesou žádný význam, ale jsou velmi četná a ztěžovala by identifikaci důležitých klíčových slov [Perikos 2018]. V angličtině to jsou převážně členy, předložky, spojky, tvary slovesa být apod. Užít k tomu byl předpřipravený, běžně užívaný a validizovaný lexikon *snowball* v softwaru R, který jsem doplnila navíc vlastním seznamem stop slov vytvořeným na míru konkrétně užívaným datovým souborům.⁴⁵ Graficky jsou tokenizace i odstranění stop slov znázorněné na Obrázku 3, str. 29. Překlepy, jazykové chyby v kmenu slova či zkratky jsem pak řešila ručním převedením slov na jeden tvar. Příkladem je slovo *apartment*, které se ale velmi často objevovalo také ve tvaru *apartment*, *apt*, *apartement* apod. Pokud by tyto tvary nebyly ručně převedeny na jeden, bylo by s nimi počítáno jako s odlišnými slovy.⁴⁶ Dále jsem do analýzy nezahrnula slova, která mají velmi nízkou frekvenci výskytu.⁴⁷ Pro některé procedury jsem provedla také stematizaci, kdy jsou slova převedena na společný kmen. To řeší problém například s třetí osobou v angličtině (koncovkou -s), množným číslem (koncovkou -s) nebo různými tvary (například slova *recommend*, *recommended* a *recommendable* se tak sloučí). Stematizace nicméně není doporučována pro modelování témat v textech, proto jsem stematizovanou verzi souborů užila pouze pro účely frekvenčních analýz, kde je přínosná.

⁴⁵ Všechny přípravy textu jsou prováděny v tzv. tidy formátu v softwaru R. V tomto formátu jsou běžně užívány tři lexikony stop slov [Silge a Robinson 2017]. Jsou jimi *snowball*, *SMART* a *onix*. Každý má jiný počet slov a je doporučováno individuálně hodnotit, který se nejvíce hodí na konkrétní datový soubor. Pro datové soubory užívané pro účely této práce byl shledán nejvhodnějším lexikon *snowball*, který obsahuje nejméně stop slov (necelých 200) a jsou nejvíce neutrální, neobsahují takřka žádná slova nesoucí význam (k nahlédnutí je obsah tohoto lexikonu v Příloze 5 na str. 105). Lexikony *SMART* a *onix* totiž obsahují slova, která nepovažuji za vhodné z recenzí odstranit, protože jsou pro analýzu v tomto kontextu relevantní (například slovo *help*, *appreciate*, *near*, *new*, *thanks*). Běžným postupem také bývá doplnit tyto základní lexikony ještě vlastním seznamem stop slov, který analytik vytvoří ručně a na míru, kdy individuálně hodnotí, která slova nejsou zajímavá pro analýzu v daném kontextu a zbytečně komplikují odhalení těch podstatných. Takový seznam jsem také vytvořila a aplikovala, k nahlédnutí je v Příloze 6 na str. 106. V souladu se Zipfovým zákonem tak byl objem datového souboru významně redukován.

⁴⁶ Takový krok jsem provedla pouze pro slova, která mají frekvenci v řádu alespoň stovek. Pro slova s menší frekvencí to v první řadě není časově proveditelné, ale ani efektivní, protože ta už nemají potenciál výsledky zvolených procedur výrazně ovlivnit.

⁴⁷ Pro frekvenční analýzy jsem nevyužila slova s frekvencí menší než padesát, pro modelování témat pak slova s frekvencí menší než tři sta. Pro analýzu frekvencí jsou totiž zajímavá i slova s menšími frekvencemi, pro modelování témat už nikoliv. Navíc pro frekvenční analýzy byla dimenzionalita souboru značně snížena i stematizací, ta není u modelování témat doporučována, proto byla snížena odstraněním většího počtu slov s nízkou frekvencí.

Obrázek 3: Proces tokenizace

Původní text recenze: Andrea was so helpful! Thank you. 😊
Tokenizovaný text: "andrea", "was", "so", "helpful", "thank", "you"
Po odstranění stop slov: "andrea", "helpful", "thank"

Zdroj: vlastní zpracování

Díky těmto postupům byly významně redukovány dimenze textů a s takto připravenými soubory jsem vstupovala do analýz.

2.3.3 Metoda Structural topic modeling (STM)

Modelování témat ve velkém objemu textu je v posledních letech stále užívanější technikou strojového učení pro organizaci, sumarizaci a porozumění obsahu rozsáhlého množství textových informací [Roberts et al. 2019, str. 1]. Metod a možností jejich využití je mnoho. Pro účely této analýzy jsem zvolila metodu STM, která funguje na principu nejužívanější metody Latentní Dirichletovy alokace (LDA), nicméně je novější a LDA v několika vlastnostech překonává – vytváří korelovaná témata, což lépe odpovídá přirozenému užívání jazyka, umožňuje zapojení metadat do analýzy a v rámci balíčku *stm* v softwaru R je implementována také řada užitečných diagnostik nápomocných pro výběr nejvhodnějšího počtu témat. Silge [2018b], specialistka na textovou analýzu v R, považuje v tuto chvíli STM za nejlepší implementaci pro modelování témat v R.

Autory balíčku *stm* v R jsou Roberts, Steward a Tingley [2014]. STM, stejně jako další obdobné techniky, odhaluje strukturu ve zdánlivě nestrukturovaném souboru textů a vytváří shluky sémanticky obdobně užívaných slov, kdy každému slovu vypočítává pravděpodobnost náležitosti k danému shluku, tedy tématu. Jedná se o generativní model, který v několika opakováních vyhledává to nejpravděpodobnější řešení [Roberts et al. 2019, str. 2]. Předpokladem všech technik pro modelování témat je, že každý dokument je směsí témat a každé téma je směsí slov, což je zásadně odlišuje od jednodušších postupů shlukování jako

například k-means shlukování, kdy každý dokument může náležet pouze k jednomu shluku.⁴⁸ Výstup STM tak nejen umožňuje identifikovat jednotlivá témata objevující se napříč dokumenty, ale také je důkladně prozkoumat přes slova k nim náležící, a tedy vidět, jak jsou v textech konstruována. Témata spolu mohou korelovat a může se stát, že jedno slovo je užíváno ve vícero tématech, tedy se objeví ve vícero shlucích, což odpovídá vlastnostem přirozeného jazyka. STM také umožňuje vypočítat očekávané proporcionální zastoupení jednotlivých témat v celkovém textu, což je užitečné pro naplnění cílů této práce, jelikož se dá očekávat, že některé aspekty se budou objevovat u obou typů ubytování, nicméně mohou mít odlišnou důležitost. STM spadá do kategorie metod strojového učení bez učitele. Není tedy potřeba ručně kódovat trénovací množinu, je ovšem nezbytné předem zadat počet témat tvořeného modelu. Standardně, a tak je tomu i v této práci, je vytvářeno několik modelů s různými počty témat, které jsou následně porovnávány se snahou vybrat ten nejlepší pro konkrétní účel. STM má k dispozici několik diagnostik, které napomohou k určení co nejlepšího počtu témat. Ideální řešení nicméně neexistuje, není žádný definitivní výsledek, jedná se o explorační techniku a důležitou roli musí stále hrát úsudek analytika a smysluplná věcná interpretace [Silge 2018b; Roberts et al. 2014, str. 4-7]. STM je doporučováno pro aplikaci na velké soubory dat, kdy jednotlivé dokumenty mají konzistentní strukturu, pak produkuje nejlepší výsledky. Data užívaná v této práci tyto požadavky splňují.

STM jsem aplikovala na oba korpusy slovních recenzí zvlášť a následně výsledky porovnála. Tento postup umožnil odhalit témata, ve kterých uživatelé ubytování ze sektoru sdílené ekonomiky a zvlášť uživatelé tradičních ubytovacích kapacit popisují své zkušenosti, která jsou celkově nejzastoupenější a jakými výrazy je konstruují. Díky shodným postupům při sběru i zpracování dat je možné toto porovnat a zjistit, liší-li se, případně jak.

2.3.4 Analýza sentimentu

STM, představená v předchozí kapitole, sice umožňuje odhalit nejčastěji zmiňovaná témata v nestrukturovaném textu, v tomto případě slovních recenzích, nicméně neumožňuje již určit, jsou-li zmiňované v pozitivním či negativním smyslu a do jaké míry. To je ovšem podstatné

⁴⁸ Kromě toho užívají sofistikovanější techniky pro určení pravděpodobnosti náležitosti dokumentů k jednotlivým tématům. Opět jsou tyto pravděpodobnosti zpřesňovány během mnoha iterací.

pro hlubší porozumění důležitosti odhalených aspektů pro uživatele. Proto ji v této práci doplňuji analýzou sentimentu, technikou určenou pro identifikaci subjektivity. Provádím jak aspektově orientovanou analýzu sentimentu, tak také na úrovni celých dokumentů.

Analýza sentimentu nebo také technika Dolování názorů (anglicky Opinion mining) narůstá v posledních letech na oblibě. A to zejména díky enormnímu růstu uživatelsky generovaného obsahu jako jsou právě recenze, ale také různé diskuze, komentáře, příspěvky na sociálních sítích apod. [Valdivia et al. 2017, str. 72]. Analýza sentimentu je technikou, jak s využitím počítače zpracovat názory, sentimenty a subjektivitu v nestrukturovaných textech [Pang a Lee 2006]. Její počátky sahají do roku 2001 [Volcani a Fogel 2001]. Užívána je pro porozumění názorům lidí na široké spektrum témat, nejčastěji pro zjištění spokojenosti zákazníků s produktem či službou [Karlgrén et al. 2012]. Analýza sentimentu má různé postupy i podoby. Může být prováděna na úrovni věty, dokumentu či aspektu [Vinodhini a Chandrasekaran 2012, str. 283]. V této práci využívám poslední dvě možnosti za účelem zjištění sentimentů v celých textech, ale také v jednotlivých aspektech/tématech odhalených s pomocí STM. Důležitým rozhodnutím při užívání analýzy sentimentů je, zdali sentimenty identifikovat s pomocí strojového učení, tedy ručního kódování na trénovací množině, nebo, a to je užíváno častěji, s využitím předpřipravených lexikonů [Saberí a Saad 2017]. Vzhledem k dostupnosti kvalitních a ověřených lexikonů pro angličtinu jsem zvolila tuto možnost. Využívám validizovaný volně dostupný lexikon *nrc* v softwaru R.⁴⁹ Na základě lexikonu je tedy jednotlivým slovům přiřazován sentiment [Díaz et al. 2018, str.1]. Základem je vypočítání polarity textu na škále negativní až pozitivní, možné je ale dopočítat také specifitější emoce jako očekávání, vztek, důvěra, překvapení, smutek, radost, strach a znechucení.⁵⁰ I přes mnoho studiemi ověřený přínos této techniky je ovšem nezbytné mít stále na paměti, že se jedná o automatizované zpracování přirozeného jazyka a lidských emocí. Výsledek už z principu nemůže být stoprocentní a nebyl by ani při ručním zpracování několika kodéry/lidmi, jelikož jsou jemné nuance mezi tím, jak kdo text čte a chápe. Karlgrén [et al. 2012] uvádí, že přesnost klasifikace textu jako pozitivního či negativního se pohybuje mezi 70 až 90 % v závislosti na tématu. Zároveň čím je text objemnější, tím je výsledek spolehlivější.

⁴⁹ Autory lexikonu jsou Mohammad a Turney [2010]. Lexikon byl vytvářen a validizován s využitím crowdsourcing na recenzích restaurací, filmů, Twitterových datech a podobně, tedy datech podobných těm, se kterými pracuji v této práci.

⁵⁰ Jedná se o základní emoce z Plutchikova kola emocí [Plutchik 1994 in Luo a Tang 2019, str. 148].

Díky analýze sentimentu tak je možné lépe porozumět, jakou zkušenost uživatelé v ubytování zažívají a o kterých aspektech se vyjadřují spíše pozitivně a o kterých negativně. Stejně tak je možné na této škále pozitivního až negativního sentimentu zhodnotit i soubory recenzí celkově, a komparovat tak korpusy recenzí z Airbnb a Bookingu i z tohoto hlediska.

2.4 Etické hledisko

Dodržení etiky výzkumu kontroluji oproti Mezinárodnímu kodexu ICC a Esomar pro výzkum trhu, výzkum veřejného mínění, sociologický výzkum a datovou analytiku [2016]. Díky povaze dat, která jsem využila pro účely předloženého výzkumu, se eliminuje celá řada potenciálních etických problémů, které mohou vznikat při kontaktu se subjektem. Kozinets [2010, str. 151] píše, že při stahování již existujících dat se nejedná výslovně o výzkum lidských subjektů, není tedy nutné usilovat například o informovaný souhlas. I přesto není web scraping, tedy automatizované stahování velkého množství dat z webů, zcela neproblematický. Jedná se o tzv. „šedou zónu“, která není legislativně jasně upravena. Dle Krotova a Silvy [2018] je zásadami nesbírat obsah, který není veřejný, nepoškozovat webovou stránku (například přetížením) a nezveřejňovat žádné osobní údaje ani citlivé údaje o organizaci vlastníci webovou stránku. Tato pravidla jsem dodržela při sbírání dat pro tuto práci i jejich následném zpracování.

V případě dat o Airbnb bylo jejich shromáždění provedeno Murray Coxem, který data archivuje na webové stránce Insideairbnb, a dovoluje je užívat každému pod podmínkami licence Creative Commons CCO 1.0 Universal „Potvrzení o statusu volného díla“ (Creative Commons). To znamená, že data poskytuje ve prospěch veřejnosti i bez nutnosti jeho souhlasu a vzdává se jakýchkoliv nároků na ně. Data z webové stránky Booking, která jsem shromáždila samostatně, užívám výhradně k akademickým účelům a dále je nešířím. V obou případech pracuji výhradně s online veřejně dostupnými daty, která si může volně prohlížet každý, kdo vstoupí na webovou stránku společností Airbnb nebo Booking. Nebyly sbírány žádné soukromé a citlivé osobní údaje a data jsou anonymizována. Analyzuji je pak v souladu s vědeckými metodami, technikami a principy výzkumu.

3 Analytická část

Analytická část práce sestává ze čtyř celků. Nejprve stručně popisují datové korpusy a jejich základní charakteristiky. Ve druhém oddílu prezentují výsledky distribučních frekvencí užívaných výrazů a tyto porovnávám mezi korpusy. Třetí oddíl věnuji výsledkům modelování témat a jejich následné komparaci mezi korpusy. Na to navazuje obsah oddílu čtvrtého, kde představuji výsledky analýzy sentimentu na úrovni jednotlivých témat i celých dokumentů.

3.1 Popis dat

V této části jsou představeny vlastnosti textových korpusů, se kterými jsem vstupovala do analýz a stručně zrekapitulovány jednotlivé kroky přípravy textů pro analýzu.

Analyzovala jsem dva korpusy textů. První korpus je tvořen uživatelskými recenzemi ubytování na platformě Airbnb, druhý pak recenzemi hromadných ubytovacích zařízení (hotelů, hostelů, penzionů) z platformy Booking. Při přípravě dat pro jednotlivé procedury jsem následovala obecně doporučované postupy [např. Silge a Robinson 2017]. Detailně je celý proces sběru dat a jejich přípravy popsán a vysvětlen v metodologické části práce. Pro stručný přehled o vlastnostech datových souborů i jejich velikosti po jednotlivých úpravách v rámci přípravy před vstupem do samotných procedur, je přiložena Tabulka 2 na str. 34. Výchozích 315 tisíc (Airbnb) a 123 tisíc recenzí (Booking) v angličtině jsem tokenizovala do slov. Následně s cílem snížit dimenzionalitu a pracovat skutečně jen s výrazy nesoucími význam jsem odstranila tzv. stop slova a čísla. V případě analýzy distribučních frekvencí následovala stematizace, tedy sloučení slov se stejným kořenem, a odstranění slov s velmi nízkou frekvencí (pod 50). Z původních 13 milionů (Airbnb) a 5 milionů (Booking) slov tak v souborech zůstalo 4,5 milionu (Airbnb) a 2 miliony (Booking) pouze plnovýznamových slov. Unikátních slov jsou to 3 tisíce (Airbnb) a 2 tisíce (Booking).⁵¹ Pro techniku modelování témat pak již není stematizace doporučována, a proto prováděna nebyla. Odstranila jsem nicméně vícero slov s velmi nízkou frekvencí (pod 300), která nejsou při redukci obsahu textu na nejdůležitější témata podstatná, a tím se snížila dimenzionalita pro svižnější propočítávání a evaluaci modelů.

⁵¹ Jak se ale během analýzy ukáže, soubor Airbnb je navyšován pouze výrazně větším zastoupením unikátních křestních jmen, po jejich sloučení jsou počty unikátních slov shodné.

Délka recenzí je v průměru stejná, a sice 40 slov, ať už uživatelé hodnotí ubytování přes Airbnb nebo v hromadných ubytovacích kapacitách. Na jedno ubytování přes Airbnb připadá v průměru 22 recenzí, na jedno ubytování přes Booking pak 186 recenzí.

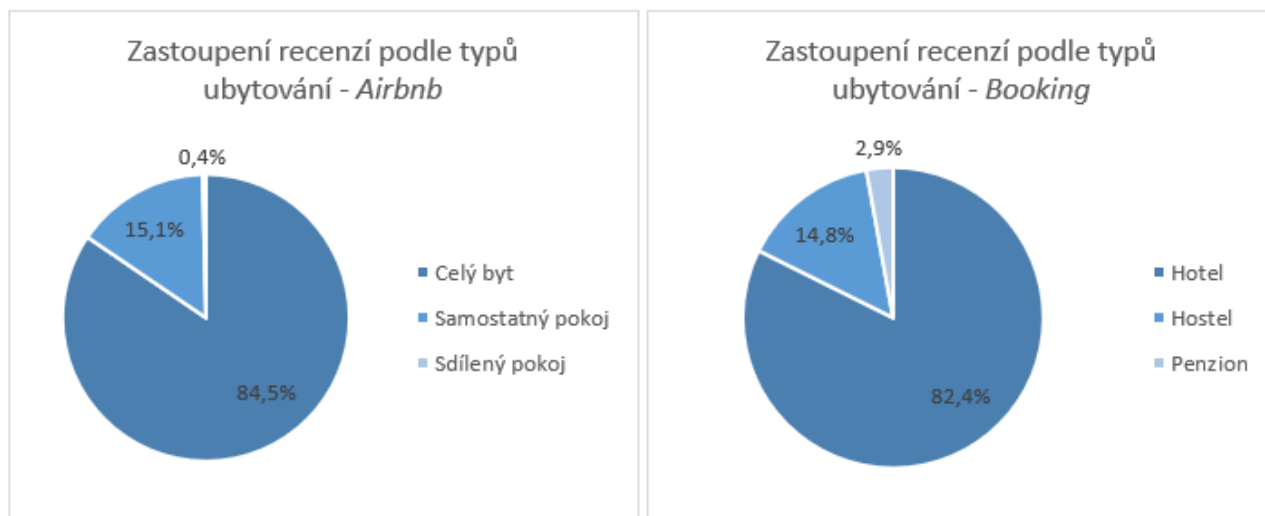
Tabulka 2: Základní vlastnosti textových korpusů recenzí z Airbnb a Booking

	Airbnb	Booking
Počet ubytování	14 185	659
Počet recenzí	315 238	122 582
Průměrný počet recenzí na ubytování	22	186
Celkový počet slov recenzí	12 799 653	4 731 565
Průměrný počet slov na recenzi	41	39
Počet slov po odstranění stop slov a čísel	5 430 350	2 037 436
Počet slov po odstranění slov s frekvencí výskytu pod 50	4 557 631	1 930 391
Počet unikátních slov (po všech úpravách + stematizaci)	2 930	2 005
Počet slov po odstranění slov s frekvencí výskytu pod 300	4 055 740	1 572 580

Zdroj: vlastní zpracování

Co se týče metadat, jak je k nahlédnutí v Grafu 1 na str. 35, naprostá většina recenzí z Airbnb za poslední dva roky jsou recenze samostatných bytů (85 %), tedy hostitel nepobývá ve stejném bytě. Následují recenze samostatných pokojů (15 %), kde zbytek bytu bývá sdílený, a naprosto zanedbatelnou část tvoří recenze sdílených pokojů (0,4 %). Recenze hromadných ubytovacích zařízení za poslední dva roky jsou pak tvořeny převážně recenzemi hotelů (82 %), následují hostely (15 %) a penziony (3 %). To zhruba odpovídá zastoupení jednotlivých typů ubytování v nabídce, kdy z celkového počtu přibližně 14 tisíc ubytování na Airbnb bylo v době sběru dat aktivních 11 tisíc samostatných bytů (78 %), 3 tisíce samostatných pokojů (21 %) a 200 sdílených pokojů (1 %). Na Bookingu bylo v době sběru dat aktivních téměř 700 ubytování, z toho 510 hotelových pokojů (77 %), 80 ubytování v hostelu (12 %) a 69 v penzionu (10 %).

Graf 1: Procentuální zastoupení recenzí podle typů ubytování



Zdroj: vlastní zpracování

Veškeré úpravy v rámci přípravy dat byly prováděny na obou korpusech naprosto totožné pro zajištění porovnatelnosti výsledků. Zároveň je možné vidět, že po provádění jednotlivých kroků jsou oba soubory ovlivňovány velmi podobně co do počtu slov. S takto očištěnými daty, která obsahují již pouze slova nesoucí význam, tedy interpretačně zajímavá v kontextu této práce, jsem vstupovala do analýzy. Zároveň s vědomím toho, že, ačkoliv jsou oba korpusy tvořeny recenzemi třech typů ubytování, v obou případech výrazně převládá jen jeden, a sice u Airbnb celý byt a u hromadných ubytovacích zařízení přes Booking hotelový pokoj.

3.2 Analýzy frekvencí

Nejprve jsem data prozkoumala z hlediska frekvenčních distribucí za účelem získání představy o obsahu korpusů a komparace na základě odlišností a podobností užívaných výrazů. Zkoumala jsem tedy slovník, jaký uživatelé při popisu svých zkušeností užívají a zda se mezi uživateli sdílených modelů a tradičních modelů ubytování liší. Při analýze distribucí slov v textech bývá rozlišováno mezi slovy s vysokou, střední a nízkou frekvencí, nejsou ovšem stanoveny přesné hranice [Baayen 2008, str. 44; Bathea et al. 2017, str. 348]. V souladu s tím i v této práci nejprve věnuji pozornost slovům s nejvyšší frekvencí a následně prozkoumávám slova s frekvencí střední a nízkou.

3.2.1 Komparace z hlediska nejčastěji užívaných výrazů

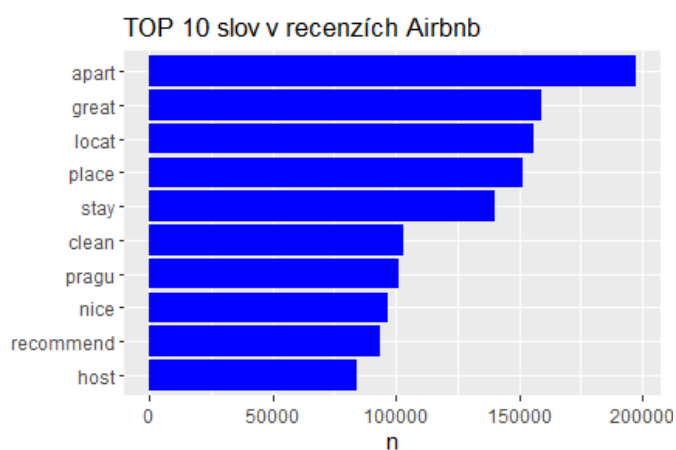
V této části komparuji soubory na základě nejčastěji užívaných výrazů při popisu zkušenosti s ubytováním, sdíleným přes Airbnb a tradičním přes Booking. Nejprve pozornost věnuji deseti úplně nejčastěji užívaným výrazům a následně sedmdesáti nejčastějším slovům. Těchto sedmdesát slov tvoří v obou případech polovinu celých souborů o jinak celkem dvou (Airbnb) a třech tisících (Booking) slov.

Deset nejčastěji užívaných výrazů se v jednotlivých textových korpusech příliš neliší. V obou případech uživatelé nejčastěji zmiňují přímo prostor, který si pronajímají, tedy *byt* u Airbnb a *pokoj* u Bookingu.⁵² V obou případech jsou na prvních místech také *lokalita* a člověk či lidé, se kterými se v ubytování setkávají či s nimi ohledně něj komunikují, tedy u Airbnb *hostitel* a u Bookingu *personál*. Při pohledu na deset nejfrekventovaněji užívaných slov (viz Graf 2 a 3 na str. 37) se jeví jako velmi překvapivé, že by snad uživatelé Airbnb dávali ve svých zkušenostech menší prostor lidskému faktoru než uživatelé tradičních ubytovacích kapacit. Slovo *personál* je u Bookingu již na třetím místě, zatímco u Airbnb slovo *hostitel* až na desátém, což neodpovídá výše představeným teoriím předpokládajícím, že jednou z hlavních přidaných hodnot pro konzumenty sdílené ekonomiky je lidský kontakt. Při dalších detailnějších

⁵² Jak bylo přestaveno při popisu dat v prvním oddílu analytické části, 85 % recenzí ubytování přes Airbnb jsou recenze celých bytů.

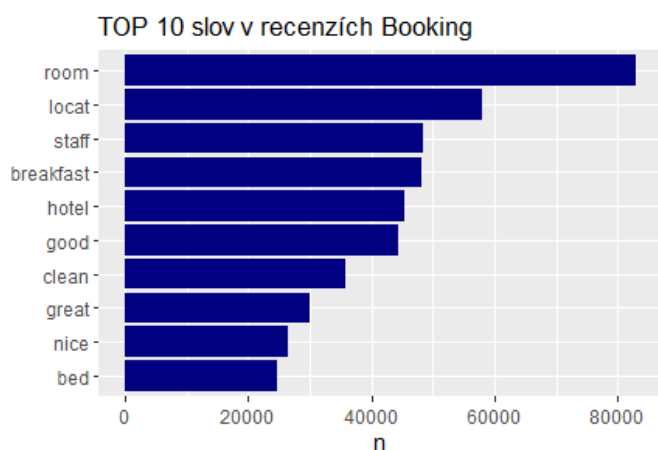
analýzách se nicméně ukázalo, proč tyto první statistiky nejsou úplně směrodatné. Společným znakem je dále velký prostor věnovaný *čistotě*. Odlišují se naopak tím, že mezi deseti nejfrekventovanějšími slovy u Airbnb je *Praha* a slovo *doporučit*, zatímco u Bookingu je to *snídaně* a *postel*. To by potenciálně mohl být poměrně zajímavý rozdíl, nicméně zatím není možné posoudit, jestli by tato slova nebyla hned mezi pár dalšími po těchto deseti nejužívanějších, proto je dále pozornost věnována detailnější analýze frekvencí doplněné také o poměry šancí pro vyčíslení míry rozdílů, že se určitý výraz objeví ve zkušenostech uživatele Airbnb více či méně krát oproti uživateli Bookingu a obráceně.

Graf 2: 10 nejfrekventovanějších slov v recenzích ubytování přes Airbnb



Zdroj: vlastní zpracování

Graf 3: 10 nejfrekventovanějších slov v recenzích ubytování přes Booking



Zdroj: vlastní zpracování

Z hlediska přehlednosti dále není vhodné porovnávat všechny nejčteněji užívané výrazy ve frekvenčních tabulkách, proto si čtenář může pro rychlý přehled prohlédnout mraky slov (viz Graf 4 na str. 39), ve kterých je vizualizováno sedmdesát nejfrekventovanějších slov, respektive jejich kořenů vzhledem k úpravě stematizací, ve zkušenostech zvláště s Airbnb a tradičními ubytovacími kapacitami. Čím je slovo v mraku větší, tím vyšší má frekvenční zastoupení v daném souboru. Ačkoliv celkem je v souboru Airbnb téměř tři tisíce a u Bookingu dva tisíce unikátních slov⁵³ i po všech úpravách⁵⁴, těchto sedmdesát nejfrekventovanějších slov tvoří u Airbnb i Bookingu zhruba polovinu celého souboru všech užívaných plnovýznamových slov.⁵⁵ Zatímco deset nejčastějších slov se příliš nelišilo, tady se již objevují zajímavé rozdílnosti. Ve zkušenostech s Airbnb se mezi těmito sedmdesáti nejfrekventovanějšími slovy objevují výrazy jako *komunikace*, *reakce/odpověď*, *děkuji*, *domov*, *místo*, *útulný*, *užívat*, *cítit*, *poskytnout*, *kuchyň* a *tip*. Jsou to tedy především slova související buď s osobou hostitele a komunikací s ním, anebo s pocity či dojmy z pobytu. Tato slova se neobjevují mezi sedmdesáti nejfrekventovanějšími slovy u Bookingu. Zkušenosti s Bookingem se naopak od Airbnb liší slovy *servis*, *bar*, *recepce*, *jídlo*, *wifi*, *káva*, *klimatizace*, *hostel*, *hotel*, *parkování*, *bar*, *lidé*, *hlučný*, *peníze* a *cena*. Uživatelé klasických ubytovacích kapacit tedy v tomto souboru recenzí častěji reflektují ve svých zkušenostech komfort a služby, které se jim v rámci ubytování dostaly a za jakou cenu. Dále se sice objevují v obou souborech sedmdesáti nejčastějších slov, ale je mezi nimi výrazný rozdíl v poměrném zastoupení v textu, kmeny slov *doporučit*, *Praha*, *jednoduché*, *navštívit*, *divit/uvažovat/obdivovat*, která se častěji objevují u Airbnb, u Bookingu je to pak slovo *snídaně*, *postel*, *excelentní*, *přátelský/přítel*, *sprcha*, *koupelna*, *vybavení*, *horko*. I tyto výrazy naznačují, že zkušenosti s Airbnb jsou specifické spíše reflektováním dojmů, zatímco u Bookingu spíše služeb a komfortu.⁵⁶ Společný základ užívaných výrazů je ale také značný, většina souvisí již s výše prezentovanými kategoriemi jako je lokalita, byt či pokoj, hostitel či zaměstnanci, ale také významnou část tvoří slova ohledně dopravní dostupnosti (*metro*, *tramvaj*, *doprava*). Pro lepší představu o souvislosti souborů jsem dopočítala korelaci,

⁵³ Hlavním důvodem pro poměrně výrazně rozdílný počet unikátních slov se při dalších užitých procedurách ukáže být velké množství křestních jmen v souboru recenzí z Airbnb.

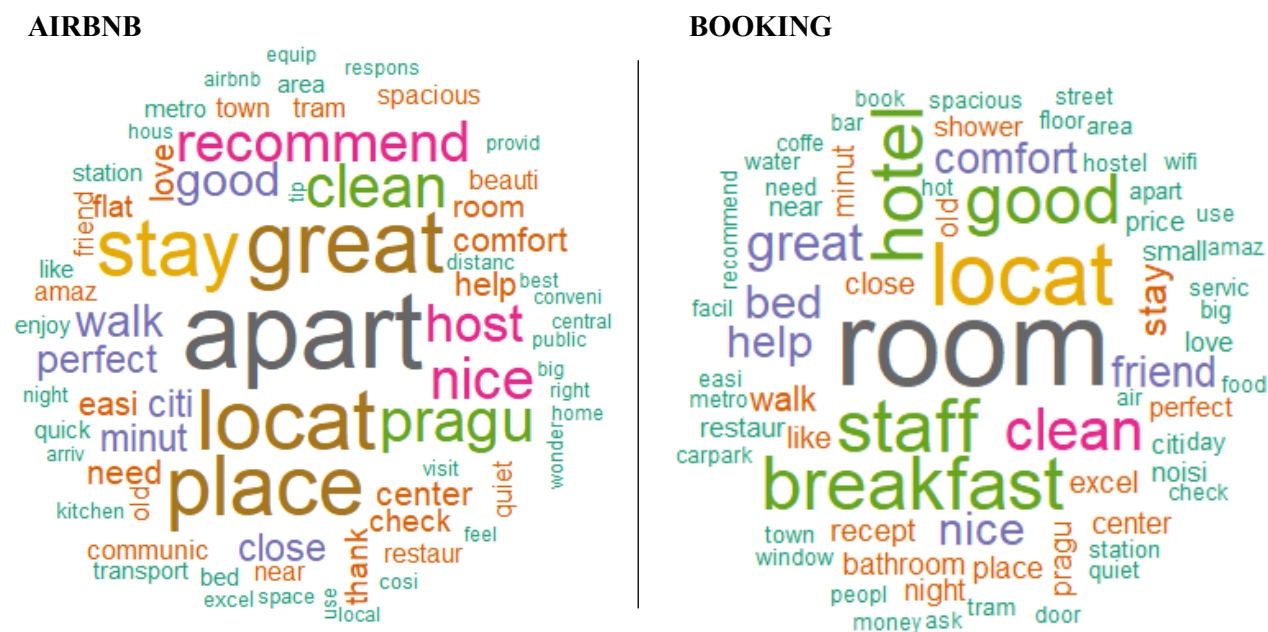
⁵⁴ Úpravami je míněna standardní příprava dat jako je tokenizace, odstranění stop slov, korekce překlepů a jazykových chyb a u těchto frekvenčních analýz stematizace a odstranění slov s frekvencí menší než padesát. Veškeré úpravy, jimiž data prošla, jsou detailně popsány v metodologické části práce v podkapitole 2.3.2 Příprava dat na str. 26.

⁵⁵ U Airbnb je to 56,3 % souboru a u Bookingu 51,6 %.

⁵⁶ Objevují se i další výrazy, jejich četnost užívání se mezi těmito soubory liší, nicméně nejsou v tuto chvíli interpretačně zajímavé, například slova *hotel* a *hostel* u Bookingu, kde je naprosto očekávatelné, že ve zkušenostech s Airbnb se objevovat nebudou.

jejíž hodnota je 0,6.⁵⁷ Jedná se tedy o podstatný až velmi silný vztah [de Vaus 2002], tedy značná část obsahu obou souborů je shodná a zkušenosti s těmito dvěma typy ubytování mají významný společný základ.

Graf 4: Mraky slov – sedmdesát nejfrekventovanějších slov – Airbnb a Booking



Zdroj: vlastní zpracování

Pro první jednoduché vyčíslení, nakolik tyto dva soubory diferencují zmíněné výrazy, kterými se liší, jsem dopočítala poměry šancí.⁵⁸ V tabulkách 3 na str. 40 pro Airbnb a 4 na str. 41 pro Booking jsou k nahlédnutí hodnoty poměru šancí a údaje, z nichž byly dopočítány. Zahrnuty jsou pouze výrazy, kde je šance pro užití v jednom ze souborů alespoň dvakrát větší než u druhého. Mezi výrazy nejvíce odlišující zkušenosti s Airbnb od Bookingu z hlediska poměrů užití v těchto sedmdesáti nejčastěji užívaných výrazech jsou *komunikace*, *reakce/odpověď* a *tip*, u kterých je dokonce třináct až sedmnáctkrát větší šance, že se objeví ve zkušenostech uživatelů Airbnb než uživatelů tradičních ubytovacích kapacit na Bookingu. Dvakrát až pětkrát větší šanci být zmíněny od uživatelů Airbnb oproti uživatelům Bookingu

⁵⁷ Pro spočítání korelace mezi frekvencemi slov byl použit Pearsonův korelační koeficient.

⁵⁸ Šance jsou využity kvůli odlišené velikosti souborů, porovnávat absolutní četnosti není možné.

pak mají výrazy *doporučit, domov, děkuji, útulný, užívat, Praha, snadný, navštívit a divit se/uvažovat/obdivovat*. Soubor zkušeností s Bookingem naopak od Airbnb nejvíce odlišují výrazy *recepce, snídaně a klimatizace*. Tyto mají třináct až dvacet sedmkrát větší šanci být zmíněny uživatelem Bookingu. Dvakrát až pětkrát je tato šance pak větší u výrazů *zařízení, služby, cena, hlučný a sprcha*. U zbylých výrazů, které se objevily mezi sedmdesáti nejfrekventovanějšími u jednoho ze souborů a u druhého nikoliv, je již poměr šancí menší než dva, tedy z tohoto hlediska soubory natolik výrazně nediferencují.

Tabulka 3: Poměry šancí užití výrazů diferencujících soubor Airbnb oproti Bookingu

TOP 70 – Slova nejvíce diferencující soubor se zkušenostmi s Airbnb oproti Booking		Airbnb	Booking	Poměry šancí
communic	<i>poměrná část souboru</i>	0,00621%	0,00037%	17,0
respons	<i>poměrná část souboru</i>	0,00469%	0,00034%	13,9
tip	<i>poměrná část souboru</i>	0,00282%	0,00022%	13,1
recommend	<i>poměrná část souboru</i>	0,0177%	0,00355%	5,0
home	<i>poměrná část souboru</i>	0,00271%	0,00056%	4,8
thank	<i>poměrná část souboru</i>	0,00669%	0,00144%	4,7
cosi	<i>poměrná část souboru</i>	0,00371%	0,00111%	3,3
enjoy	<i>poměrná část souboru</i>	0,00435%	0,00138%	3,2
pragu	<i>poměrná část souboru</i>	0,01919%	0,00605%	3,2
easi	<i>poměrná část souboru</i>	0,00764%	0,00268%	2,9
visit	<i>poměrná část souboru</i>	0,00361%	0,00131%	2,8
wonder	<i>poměrná část souboru</i>	0,00343%	0,00129%	2,7

Zdroj: vlastní zpracování

Tabulka 4: Poměry šancí užití výrazů diferencujících soubor z Bookingu oproti Airbnb

TOP 70 – Slova nejvíce diferencující soubor se zkušenostmi s Bookingem oproti Airbnb		Booking	Airbnb	Poměry šancí
recept	<i>poměrná část souboru</i>	0,00703%	0,00026%	27,4
breakfast	<i>poměrná část souboru</i>	0,02502%	0,00106%	23,6
aircon	<i>poměrná část souboru</i>	0,00362%	0,00028%	12,7
facil	<i>poměrná část souboru</i>	0,00293%	0,00059%	5,0
servic	<i>poměrná část souboru</i>	0,00383%	0,0008%	4,8
price	<i>poměrná část souboru</i>	0,00548%	0,0021%	2,7
noisi	<i>poměrná část souboru</i>	0,00534%	0,00194%	2,7
shower	<i>poměrná část souboru</i>	0,00464%	0,00175%	2,6

Zdroj: vlastní zpracování

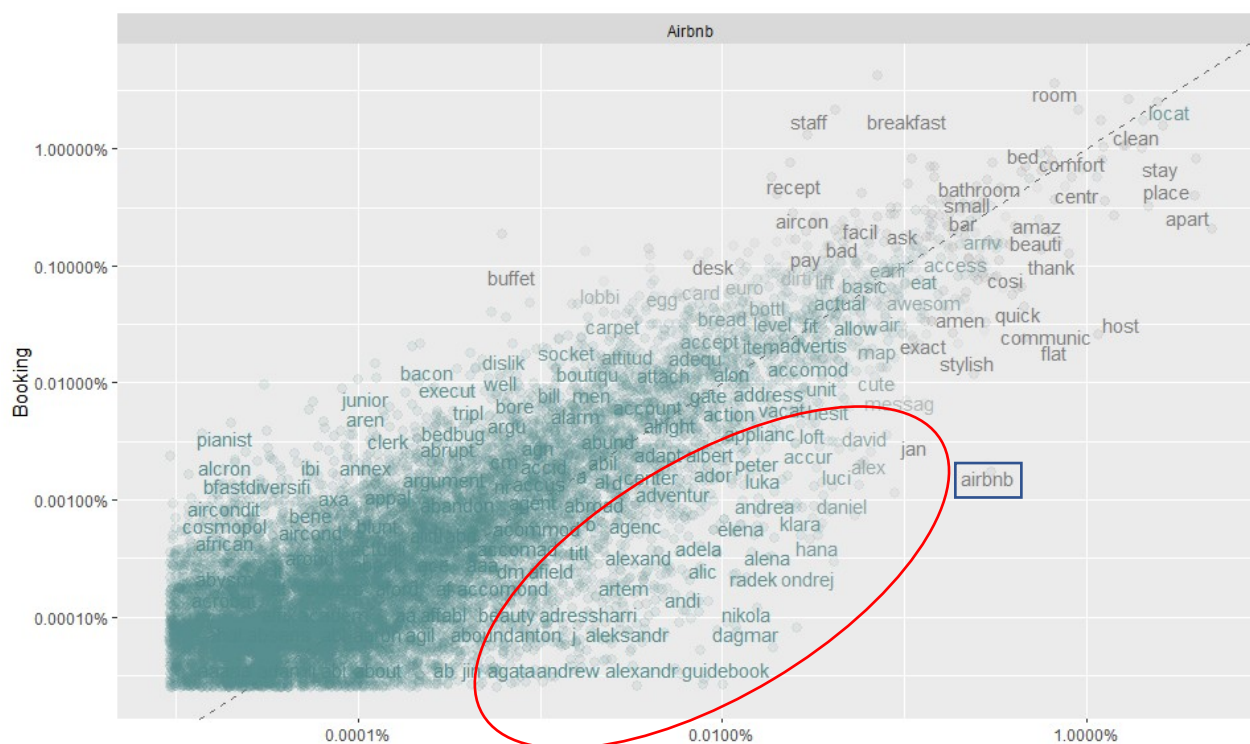
Shrnutí

Z hlediska frekvenční distribuce a korelace bylo prozkoumáno sedmdesát čtenostně nejzastoupenějších slov, které tvoří přes polovinu obou upravených korpusů z celkového počtu tří a dvou tisíc slov unikátních. Bylo zjištěno, že zkušenosti uživatelů s Airbnb a zkušenosti uživatelů s hromadnými ubytovacími zařízeními mají významný společný základ. V obou korpusech je pozornost věnována lokalitě, samotnému pokoji či bytu, ve kterém uživatel pobývá, jeho vybavení, čistotě a dopravní dostupnosti. Zároveň se ale v obou korpusech mezi těmito nejfrekventovanějšími slovy objevují taková, která zkušenosti značně diferencují. Pro lepší představu byly tyto rozdíly vyčísleny v šancích užití nějakého výrazu. Ukázalo se, že uživatelé Airbnb několikanásobně častěji hodnotí své pocity z ubytování, zmiňují lokalitu (Praha) a především kontakt s tím, kdo je ubytovává. Oproti tomu uživatelé Bookingu několikanásobně častěji zmiňují služby, komfort ubytování, ale také jeho cenu, což je překvapivé, jelikož, jak bylo popsáno v teoretické části, objevuje se řada zastánců názoru, že spíše uživatelé sdíleného ubytování kladou důraz na cenu a vybírají si ho pro finanční výhodnost.

3.2.2 Explorace kompletních souborů z hlediska doporučených frekvencí

Již prozkoumáním horní poloviny souborů se ukázaly nejen zajímavé odlišnosti mezi nimi, ale také jejich společný základ. Zároveň se jeví jako zvláštní, že ačkoliv sedmdesát čtenostně nejzastoupenějších slov tvoří u obou souborů jejich více než polovinu, zkušenosti s Airbnb ale jinak obsahují o přibližně třetinu více unikátních slov než Booking. Pro získání přehledu o obsahu celých souborů, včetně středně a málo frekvenčně zastoupených slov, vizualizují do grafu doporučené frekvence (viz Graf 5).

Graf 5: Porovnání doporučených frekvencí slov mezi recenzemi na Airbnb a Bookingu



Zdroj: vlastní zpracování

V grafu může čtenář vidět data za Airbnb i za Booking. Nad osou jsou výrazy, které se podle doporučených frekvencí objevují častěji v souboru se zkušenostmi s ubytováním přes Booking, pod osou pak ve zkušenostech s Airbnb. Slova, která mají přibližně stejné poměrné frekvenční zastoupení v těchto souborech, leží přímo na ose. Čím dále od osy je slovo umístěno,

tím je pro daný soubor specifitější v porovnání s referenčním. Zároveň čím výše na ose je slovo umístěno, tím vyšší má frekvenci. Pravý horní roh grafu byl již pečlivě prozkoumán v předchozí podkapitole a čtenář může vidět, že skutečně *lokalita*, *čistota* a u Airbnb *byt*, u Bookingu *pokoj* jsou společnými znaky. Dále se opět vydělují u Bookingu výrazy spojené především se službami či vybavením, a to i na nižší frekvenční úrovni (*bufet*, *lobby*, *přepážka*, *slanina* a další), u Airbnb to jsou pak opět i na té nižší frekvenci spíše pocity či dojmy a slova spojená s komunikací (*stylové*, *zpráva*, *rychlé*, *roztomilé*, *mapa*, *průvodce* a další). Co ovšem strhává pozornost, je hlouček slov vydělující se u Airbnb v části se slovy s nižší frekvencí. Pro přehlednost je v grafu označen červeně. Ukazuje se, že se ve zkušenostech s ubytováním přes Airbnb objevuje velké množství křestních jmen. Zároveň je vidět, že u Bookingu se žádná jména nevydělují. To je velmi zajímavé dílčí zjištění, věnuji mu proto větší pozornost.

Vzhledem k tomu, že vůbec byla tato křestní jména do grafu vizualizována, nasvědčuje to tomu, že mají v souboru významné proporční frekvenční zastoupení. Dá se zároveň předpokládat, že jmen bude v souborech ještě mnohem více. Vzhledem k tomu, že jedno konkrétní křestní jméno jako takové není interpretačně zajímavé, alespoň v kontextu této práce, rozhodla jsem se detekovat v textech co nejvíce jmen bude možné a sloučit je do jedné kategorie pro přehled, jaké zastoupení v textech mají.⁵⁹ Odhalil se tak velmi pozoruhodný fakt, a sice že lidské jméno je celkově dokonce už druhým nejčastěji používaným slovem ve zkušenostech s ubytováním přes Airbnb (viz Graf 6 na str. 45). Předčilo tak *lokaci* i *čistotu*, jediné slovo, které zůstává před ním, je *byt*. Ještě zajímavějším je pak srovnání se souborem zkušeností s ubytováním v tradičních ubytovacích kapacitách přes Booking. Tam se z hlediska četností užívaných slov dostala kategorie s lidskými jmény na padesáté sedmé místo a je zmiňováno zhruba stejně často jako *parkování*, *úžasný*, *jídlo* frekvenčně těsně nad ním nebo *horko*, *okno* a *káva* frekvenčně těsně pod ním. Absolutní četnosti nemá význam srovnávat⁶⁰, jelikož soubor se

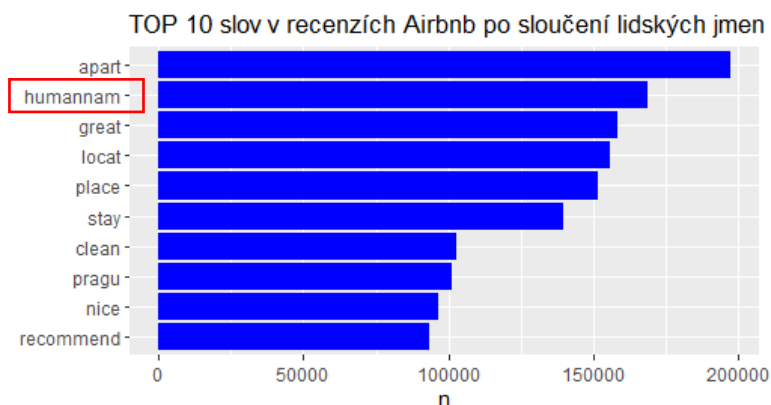
⁵⁹ Z hlediska praktické proveditelnosti byl toto poměrně náročný krok. V textech se objevují jména v různých tvarech a také s různými chybami, jelikož se často jedná o česká jména, která ve svých zkušenostech zmiňují cizinci, tedy v nich chybují. Zároveň se ale rozhodně nejedná výhradně o česká jména. Využít tedy jen určité seznamy jmen by nebylo příliš efektivní. Ty tedy vytvořily pouze základ doplněný o jména dohledaná v souborech s využitím procedury TF-IDF pro hledání klíčových slov, kde jména měla vysokou hodnotu, protože se jedná o specifická slova, a byla tak rozpoznatelná. Základy fungování této procedury popsali velmi přehledně Silge a Robinson [2017, kapitola 3.3]. Tento postup jsem aplikovala na oba soubory. V dalším kroku jsem doplnila seznam jmen o jména hostitelů, která byla společně s recenzemi sbírána z platformy Airbnb. Posledním krokem pro detekování zbylých nezahrnutých jmen pak bylo ještě ruční dohledávání jmen, a to v obou souborech. Takto jsem vytvořila seznam jmen, detekovala je v textech a sloučila do jedné kategorie. Nedá se vyloučit, že nějaké jméno bylo opomenuto, ale rozhodně pokud ano, pak se jedná o naprosté minimum oproti množství, které se podařilo zachytit.

⁶⁰ Jen pro přehled čtenáře, lidské jméno se v souboru z Airbnb objevilo téměř sto sedmdesát tisíckrát z celkového počtu více než pět milionů slov. V souboru z Bookingu o téměř dvou milionů slovech se objevilo přes pět tisíckrát.

zkušenostmi z Airbnb je celkově větší, ale při porovnání doporučených frekvencí jsem zjistila, že lidské jméno se ve zkušenostech s Airbnb objevuje více než desetkrát častěji než ve zkušenostech s tradičními ubytovacími kapacitami přes Booking. To zaprvé vysvětluje, proč bylo v souboru o Airbnb o třetinu více unikátních slov, nyní se počet téměř dorovnal, ale především to vyvolává otázky, proč tomu tak je a co to znamená celkově pro zkušenost, kterou uživatel Airbnb v porovnání s uživatelem Bookingu má. Je nutné dodat, že pokud chce uživatel ve své recenzi jmenovat člověka, se kterým během pobytu přišel do styku, ať už je to právě hostitel u Airbnb nebo zaměstnanec u Bookingu, pak uživatel Airbnb to má jednodušší vzhledem k tomu, že na online profilu hostitele přímo vidí jeho jméno. Uživatel Bookingu by si jméno musel pamatovat ze jmenovky, kterou zaměstnanci nosí, případně e-mailové, telefonické či jiné komunikace, pokud nějaká s ubytováním proběhla, aby ho mohl použít v recenzi. Do jisté míry tam tedy roli hraje paměť a vyvstává otázka, jestli by ta jména použil, pokud by je měl viditelně k dispozici při psaní recenze stejně jako je má uživatel Airbnb. Nicméně tak to není, k dispozici je nemá, a to je podstata tohoto zajímavého rozdílu. Zatímco u Airbnb uživatel/host už od okamžiku prohlížení si nabídky ubytování ví, kdo je hostitelem/hostiteli, tedy s kým bude komunikovat a kdo o něj bude v ubytování pečovat, u hromadných ubytovacích zařízení tomu tak není.⁶¹ Tam se host zpravidla setká s celou řadou zaměstnanců, které pozná až na místě, každý se zpravidla stará jen o jednu dílčí službu a střídají se. Evidentně se toto promítá i do popisu zkušeností s ubytováním, kdy u Airbnb je to zkušenost na úrovni křestních jmen, to znamená osobní a neformálnější, zatímco u Bookingu, pokud se mluví o lidském faktoru v ubytování, tak je referováno k recepci, zaměstnanci, číšníkovi a podobně, tedy k roli a jak je vykonávána, nikoliv ke konkrétnímu člověku tuto roli vykonávající. Zkušenost s hromadnými ubytovacími zařízeními lze tedy považovat za anonymnější i dle toho, jak je popisována samotnými uživateli. Zároveň se tím vysvětluje, proč před objevením této skupiny křestních jmen bylo u Airbnb zmiňováno slovo *hostitel* méně často než u Bookingu slovo *zaměstnanec*, uživatelé Airbnb své hostitele mnohem častěji nazývají přímo křestním jménem.

⁶¹ Pokud hostitel má pro provozování ubytování výpomoc nějakého dalšího člověka, se kterým může host přijít do kontaktu, či je hostitelů více pro jedno ubytování (například partnerský pár), stává se běžné, že i jeho jméno a často také fotka jsou u ubytování zobrazeny.

Graf 6: 10 nejčastějších slov v recenzích Airbnb po sloučení jmen do jedné kategorie



Zdroj: vlastní zpracování

Kromě hloučku křestních jmen na straně Airbnb se ve zmíněném grafu 5 na str. 42 proporčních frekvencí slov ukázala ještě jedna odlišnost mezi soubory. V souboru se zkušenostmi s ubytováním přes Airbnb má výrazně vyšší proporční frekvenci slovo *airbnb* (v grafu označeno modře), což je logické, nicméně graf neukazuje, že by snad v souboru se zkušenostmi s hromadným ubytováním přes Booking bylo oproti souboru Airbnb významně více zastoupené slovo *booking*, ačkoliv oba jsou pouze v roli zprostředkovatele, tedy v podstatě ve stejné roli. Tady se nicméně mohl projevit jeden z nedostatků automatizovaného zpracování textů, kdy slovo *booking* je v angličtině plnovýznamovým slovem a algoritmus nepozná, v jakém smyslu je zrovna používáno. Navíc po stematizaci se sloučilo se slovem *book*. Proto jsem toto prozkoumala ručně. Nezdá se ovšem, že by uživatelé Bookingu odkazovali ve svých zkušenostech na Booking, ve smyslu zprostředkovatele, významně často, jelikož slovo *booking* (ještě před stematizací) má v souboru se zkušenostmi s Bookingem pouze 2,5krát větší šanci na užití než v souboru se zkušenostmi s Airbnb. Z toho bylo dopočítáno, kolikrát zhruba mohl tedy být výraz *booking* užit ve smyslu Booking – zprostředkovatel.⁶² Je ale nutné vzít v potaz, že v souboru se zkušenostmi s Bookingem se objevuje s relativně vysokou četností také výraz *booking.com*.⁶³ Zajímavé je, že výraz *airbnb.com* se objevuje v úplně celém souboru se zkušenostmi s Airbnb obsahující několik milionů slov pouze čtyřikrát (viz Tabulka 5 na str.

⁶² V recenzích z Airbnb bylo slovo *booking* užito 2124krát, tedy tvoří celkový soubor z 0,000402 %. V recenzích z Bookingu bylo slovo *booking* použito 1978krát, tedy tvoří celkový soubor z 0,001029 %. U Bookingu je tedy slovo *booking* užíváno 2,5krát častěji a předpokládám, že toto častější užití by se dalo připsat tomu, že je užíváno ve smyslu Booking – zprostředkovatel, a to tedy zhruba 1187krát.

⁶³ Časté užívání tohoto výrazu jsem opět detekovala také díky využití procedury TF-IDF, kterou jsem užila i pro detekování velkého množství křestních jmen.

46). Celkově je pak v souboru Airbnb sedmkrát větší šance, že je přímo zmíněn zprostředkovatel oproti souboru Booking. Pro lepší pochopení, jak jsou tato slova užívána, a validnější interpretaci, proč k takovým rozdílům dochází, jsem nahlédla do nejčtetnějších trigramů, tedy třech nejčastěji po sobě jdoucích slovech, která obsahují slova *airbnb* u souboru z Airbnb a *booking* či *booking.com* u souboru z Bookingu.⁶⁴ Jak je vidět v Tabulce 6 na str. 47, výraz *airbnb* uživatelé Airbnb neuvádějí vyloženě ve smyslu zprostředkovávající online platformy, ale spíše ve významu konceptu ubytování, namísto slova *pokoj*, *byt*, *ubytování* nebo *pobyt*. Evidentně ubytovat se *přes* Airbnb pro ně znamená ubytovat se *v* Airbnb. Patrně ubytování přes Airbnb nese význam ve smyslu konceptu, zdá se, že je to pojem s určitými charakteristikami. To by mohlo být zajímavým východiskem pro další spíše kvalitativní studie. U Bookingu nic takového neplatí, ten je převážně užíván ve smyslu online platformy.

Tabulka 5: Četnosti užití slov *airbnb* a *booking* v příslušných korpusech

Airbnb		Booking	
výraz	n	výraz	n
<i>airbnb</i>	15189	<i>booking</i> (ve významu zprostředkovatel)	1187*
<i>airbnb.com</i>	4	<i>booking.com</i>	321

*jedná se o přibližnou dopočítanou četnost

Zdroj: vlastní zpracování

⁶⁴ Pro tyto účely bylo slovo *booking.com* rozděleno na *booking a com*, aby slovo *booking* mělo stejnou šanci dostat se mezi nejčtetnější trigramy v souboru z Bookingu jak s příponou *com*, tak bez ní.

Tabulka 6: Porovnání 10 nejčastějších trigramů obsahující slovo *airbnb* a *booking*

Airbnb (celkový počet slov: 5 280 230)		Booking (celkový počet slov: 1 923 138)	
trigram	n	trigram	n
<i>the best airbnb</i>	1116	<i>on booking com</i>	280
<i>recommend this airbnb</i>	483	<i>through booking com</i>	94
<i>best airbnb experience</i>	341	<i>booking com and</i>	91
<i>best airbnb experiences</i>	213	<i>the booking com</i>	89
<i>this airbnb is</i>	210	<i>in booking com</i>	84
<i>airbnb i have</i>	209	<i>on the booking</i>	73
<i>best airbnb i</i>	199	<i>in the booking</i>	72
<i>the airbnb is</i>	184	<i>booking com the</i>	57
<i>my best airbnb</i>	156	<i>from booking com</i>	54
<i>the airbnb was</i>	156	<i>by booking com</i>	52

Zdroj: vlastní zpracování

Shrnutí

Při prozkoumání zbytku souboru z hlediska proporčních frekvencí se opět ukázalo, že i na nižších frekvencích mají uživatelé Airbnb tendenci zmiňovat spíše slova spojená s pocity z ubytování, komunikací s hostitelem apod., zatímco uživatelé Bookingu častěji zmiňují komfort a služby, jaké se jim v ubytování dostaly. Hlavním zjištěním nicméně je, že se v souboru se zkušenostmi s ubytováním přes Airbnb objevují velmi často křestní jména a po jejich sloučení do jedné kategorie se ukázalo, že jsou dokonce druhým nejužívanějším slovem vůbec. V referenčním souboru z Bookingu se jména objevují výrazně méně často (více než desetkrát méně). Uživatelská zkušenost s Airbnb se tak zdá být na mnohem osobnější úrovni. Kromě toho bylo také odhaleno, že uživatelé Airbnb mluví o Airbnb jako o konceptu ubytování, nikoliv jako o pouhém online zprostředkovateli, u Bookingu je tomu obráceně.

3.3 Modelování témat

Analýzy frekvenčních distribucí byly dobrým začátkem pro prozkoumání textů a vytvoření si představy o jejich obsahu. Byly tak objeveny náznaky, že mezi texty jsou společné i diferencující kategorie. Pro systematičtější a koherentnější popsání těchto vlastností a jejich rozdílností mezi korpusy jsem s využitím nástrojů pro těžení z textů modelovala témata, která se v korpusech nejsystematičtěji objevují. Pro detekování latentních témat jsem zvolila metodu *Structural topic modeling* (STM) obsaženou v balíčku *stm* v softwaru R [Roberts et al. 2014]. Principy této metody jsem představila v metodologické části práce v kapitole 2.3.3 na str. 29. Zjednodušeně řečeno tato metoda slouží k redukci obsahu textu odhalením nejvíce se opakujících témat.⁶⁵ S jejím využitím jsem zvolila nejvhodnější počet koherentních témat, tato prozkoumala a podle výrazů je tvořících pojmenovala. Následně jsem tato témata porovнала mezi korpusy textů z Airbnb a Bookingu, včetně jejich podílného zastoupení v textech.

3.3.1 Trénování, evaluace a interpretace modelů – postup

Při modelování témat jsem na oba korpusy aplikovala stejný postup, nejprve ho tedy popíši obecně a v dalších částech prezentuji výsledky za každý korpus zvlášť, které posléze v posledním kroku komparuji. Při modelování témat je postupováno v několika fázích. Prvním krokem je vůbec nalezení vhodného počtu témat, následuje jejich prozkoumání a pojmenování a poslední fází je pak dopočítání jejich proporce v textech.

Vzhledem k explorační povaze technik pro modelování témat v textech není předem znám přesný počet témat, který by korpus měl obsahovat. Standardním postupem pro nalezení vhodného počtu témat při využití technik tohoto typu je vytvoření vícero modelů o různých počtech témat a jejich následná evaluace. Hlavním kritériem by vždy měla být věcná smysluplnost výsledků, velkou roli tedy hraje úsudek analytika a rozhodně není vhodné nechávat výběr počtu témat pouze na algoritmech. Přesto má balíček *stm* implementovaných několik diagnostik, které mohou při rozhodování být nápomocné. Nejprve jsem tedy vytvořila

⁶⁵ Metoda STM pracuje s unigramy, tedy jednotlivými slovy. Téma lze proto chápat jako shluky slov, která jsou v textech užívána společně a v sémanticky podobných způsobech.

modely o dvou až jedenácti tématech pro oba korpusy zvlášť a sama zhodnotila, nakolik jsou všechna témata interpretovatelná a smysluplná. Vytvořila jsem si tak odhady, jaký počet témat by mohl být nejvhodnějším řešením a následně jsem dopočítala nápomocné diagnostiky pro konečné rozhodnutí, ačkoliv je nutné zdůraznit, že jedno jednoznačně ideální řešení většinou neexistuje [Silge 2018b]. První diagnostikou, kterou jsem využila, je *held-out likelihood*. Jedná se o kritérium na principu křížové validace, kdy množina dat je rozdělena na podmnožiny a je ověřována vhodnost řešení pro každou takovou podmnožinu [Wallach et al. 2009]. Dalším jsou *rezidua* [Taddy 2012], u kterých by disperze měla být co nejmenší, aby model co nejlépe odpovídal struktuře textu. Dalším je pak *Lower bound* [Roberts et al. 2014, str. 10], což je kritérium konvergence, které pracuje s tím, jak velké změny se v modelu dějí při iteracích a posledním kritériem užívaným v této práci je *sémantická koherence* [Mimno et al. 2011], která hodnotí frekvenci, nakolik se nejpravděpodobnější slova tvořící každé modelované téma objevují společně. *Sémantická koherence* nicméně již z principu, na kterém funguje, vždy naznačuje, že je lepší volit menší počet témat, proto jsem ji porovnávala také s *exkluzivitou*, která pomáhá najít vhodný kompromis – cílem, aby témata byla sémanticky koherentní, ale také jasně vydělená a nepřekrývala se příliš [Roberts 2014, str. 6].

Jakmile je vybrán model o nejvhodnějším počtu témat (což může i odviset od účelu celého výzkumu), je tento model interpretován na základě výrazů, které tvoří jednotlivá témata a ta jsou vzhledem k interpretaci pojmenována. Tyto výrazy jsou přiřazeny k tématům na základě pravděpodobnosti (*beta*), se kterou s ním souvisejí. Je nicméně nutné brát v potaz, že slova u každého tématu s nejvyšší pravděpodobností jsou zároveň přirozeně ta s nejvyšší četností. Dále je dopočítáno proporční zastoupení jednotlivých témat v celkovém korpusu (*gamma* pravděpodobnost), na jehož základě je možné usuzovat na důležitost tématu pro autory analyzovaných textů.

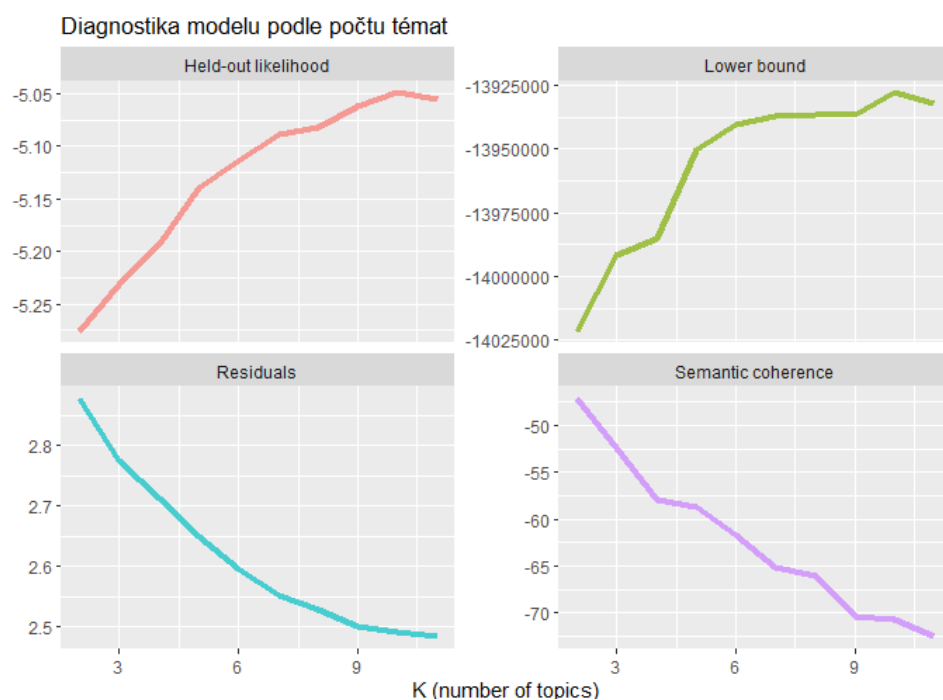
Vytváření modelů i jejich evaluace jsem prováděla zvlášť na korpusu se zkušenostmi s ubytováním přes Airbnb a zvlášť na korpusu se zkušenostmi s ubytováním v hromadných ubytovacích zařízeních přes Booking, proto i prezentace výsledků je rozdělená do dvou oddílů a následně třetí oddíl je věnován jejich komparaci.

3.3.2 Modelování témat v korpusu recenzí z Airbnb

Pro korpus slovních recenzí ubytování přes Airbnb jsem namodelovala deset různých modelů o dvou až jedenácti tématech. Výsledky diagnostik (vysvětlených v předchozím oddílu) pro tyto modely jsou vizualizovány do Grafu 7 na str. 51. Tyto diagnostiky naznačují, že nejvhodnější počet témat by měl být zhruba mezi sedmi a desíti. Pro takové modely jsou *held-out likelihood* a *lower bound* nejvyšší, *rezidua* naopak nejnižší. *Sémantická koherence* je pak porovnána navíc s *exkluzivitou* v Grafu 8 na str. 51 pro modely o různém počtu témat⁶⁶, kdy jako vhodný kompromis se jeví osm až deset témat. Všechny tyto varianty jsem prozkoumala podle slov, která by je s největší pravděpodobností tvořila a vybrala model s nejsmysluplnější interpretací. Postupně jsem tvořila modely od nejnižšího počtu témat po nejvyšší. Až do počtu osmi témat každé nově vytvořené téma dávalo interpretačně smysl a bylo jedinečné, nicméně od devíti témat výše se již tvořila témata interpretačně nejasná, nekoherentní a začala se výrazně překrývat. Zvolila jsem tedy model o osmi tématech.

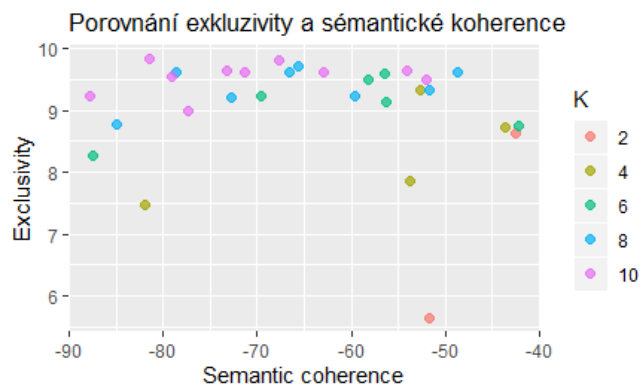
⁶⁶ Pro přehlednost jsou vizualizovány jen modely o dvou, čtyřech, šesti, osmi a desíti tématech. Diagnostiky jsou pouze orientační, není tedy nutné lpět na drobných nuancích a při vizualizaci všech modelů se graf stal nepřehledným.

Graf 7: Diagnostika modelu podle počtu témat – Airbnb



Zdroj: vlastní zpracování

Graf 8: Porovnání exkluzivity a sémantické koherence témat – Airbnb



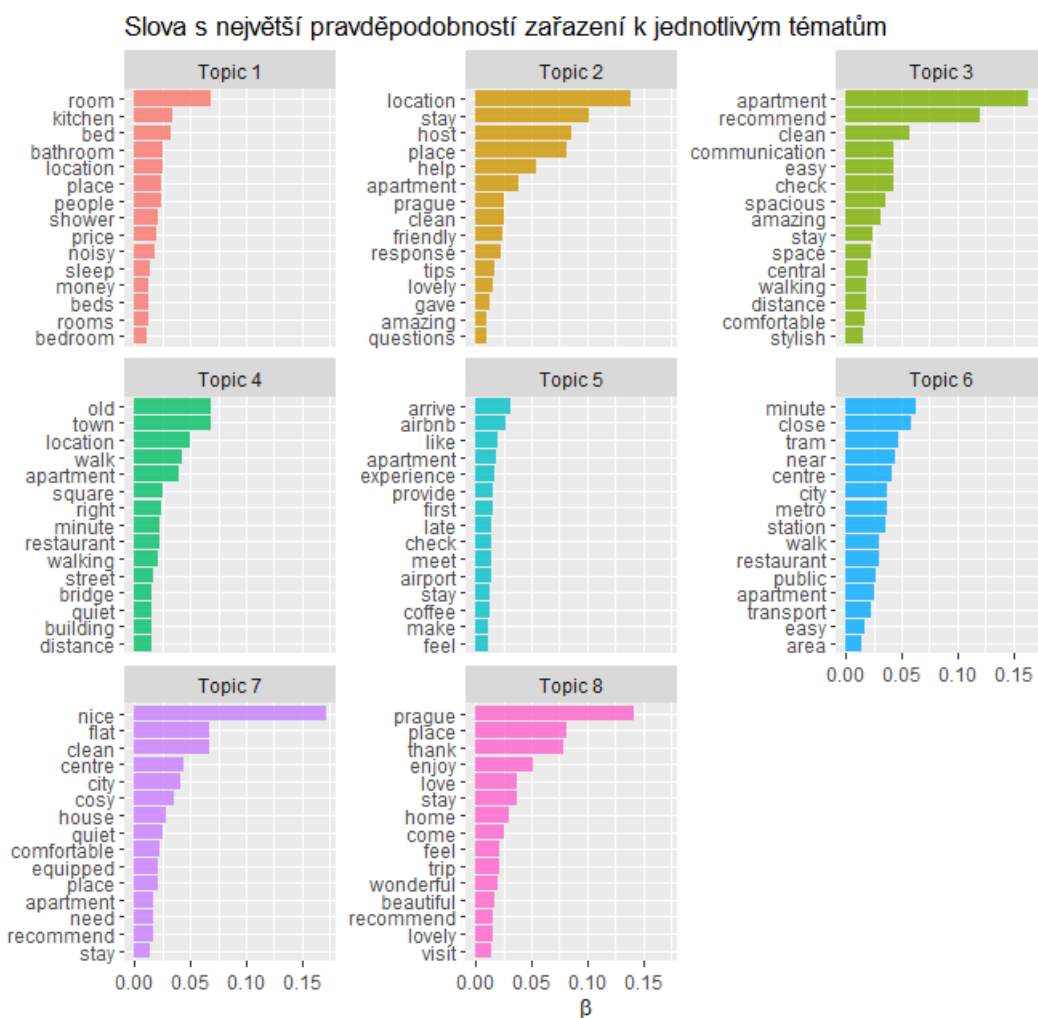
Zdroj: vlastní zpracování

Výsledný model redukující obsah recenzí uživatelů Airbnb je k nahlédnutí v grafu 9 na str. 53. Model je snadno interpretovatelný a každé téma zachycuje něco nového, jak bylo cílem. Pro přehlednost je v tomto grafu zaneseno pro každé z témat patnáct slov s nejvyšší pravděpodobností náležitosti k tématům. STM pracuje na principu pravděpodobnosti,

dopočítává tedy pravděpodobnost, se kterou slovo patří k danému tématu (*beta* na ose *x*) a je možné, že jedno slovo je používáno v souvislosti s vícero tématy, nicméně s jinou pravděpodobností. Slova s nejvyšší pravděpodobností ale zpravidla bývají i slova s nejvyšší četností, pro detailnější vhled do obsahu detekovaných témat jsem proto vygenerovala celkem padesát slov pro každé téma a tato jsou k nahlédnutí níže v Tabulce 7 na str. 55. Při interpretaci témat samozřejmě přihlížím i k nim. U prvního tématu se nejvíce objevují slova ohledně částí ubytovacího prostoru jako *pokoj*, *kuchyň*, *koupelna*, *sprcha* či *ložnice* a dále slova ohledně vybavení (*postel*, *wifi*, *ručníky*), toto téma jsem tedy pojmenovala **pokoj a jeho vybavení** (Topic 1). Druhé téma nejvíce obsahuje slova týkající se hostitele, ať už přímo slovo *hostitel* nebo křestní jména (*Jan*, *Jana*, *David*, *Tomáš*, *Alex*), ale také slova spojená s jeho vlastnostmi (*přátelský*, *milý*, *vstřícný*) nebo péči, kterou hostovi poskytl (*pomoc*, *odpověď*, *tipy*, *uvítat*, *pohostinnost*, *informace*) a komunikaci s ním (*odpověď*, *rychlý*, *komunikace*, *zprávy*). Pojmenovala jsem ho tedy jako **hostitel a jeho péče o hosta** (Topic 2). Třetí téma je méně jednoznačné, obsahuje slova spojená s jinými tématy v souvislosti se slovy jako *doporučit*, *jednoduché*, *pobyt*, *prostor*, *pohodlný*, dále *hodnota*, *místo*, *průběh* aj. Dá se tedy říct, že se jedná o celkové zhodnocení, toto téma jsem pojmenovala jako **celkový dojem z ubytování** (Topic 3). Čtvrté téma je už opět naprosto jednoznačné, objevují se slova týkající se konkrétní lokace (*Staré*, *město*, *náměstí*, *most*, *Karlův*, *řeka*, *hrad*) spolu se slovy ohledně vzdálenosti (*pěšky*, *přímo*, *vzdálenost*, *pohodlný*, *blízko*). Pojmenovala jsem jej tedy **lokalita ubytování** (Topic 4). Páté téma je tvořeno slovy týkajícími se příjezdu nebo odjezdu (*přijet*, *přihlásit*, *pozdě*, *setkat*, *letišť*, *brzy*, *klíče*, *zavazadla*, *let*, *taxi*, *vyzvednout*), pojmenovala jsem ho **příjezd a odjezd** (Topic 5). Téma šest obsahuje opět slova ohledně vzdálenosti (*minuta*, *blízko*, *poblíž*, *pěšky*), ale tentokrát v kombinaci se slovy ohledně dopravy (*tramvaj*, *metro*, *veřejná*, *doprava*). Pojmenovala jsem ho **dopravní dostupnost** (Topic 6). Sedmé téma je opět o něco méně jednoznačné a dává smysl až při zhodnocení všech padesáti slov, nejen prvních patnácti, nicméně obsahuje podstatná jména jako *byt*, *město*, *dům*, *místo*, *pobyt*, *oblast*, *sousedství*, *parkování* ve spojení s přídavnými jmény *milý/hezký*, *útulný*, *tichý*, *vybavený*, *vřelý*, *klidný*, *dostupný*, *světlý*. Uživatelé tedy popisují celý prostor, ve kterém se daná ubytovací jednotka nachází, proto jsem toto téma pojmenovala jako **celý dům a jeho sousedství** (Topic 7). Poslední a v kontextu této práce nejzajímavější koherentní téma jsem pojmenovala jako **zážitky a emoce z pobytu** (Topic 8), jelikož je tvořeno slovy jako *děkuji*, *užít*, *milovat*, *domov*, *cítit se*, *úžasný*, *krásný*, *šťastný*, *speciální*, *potěšení*, *uvolňující*. A co je zajímavé, téma obsahuje také několik křestních jmen (*Martin*, *Zuzana*, *Petra*, *Kateřina*), tedy tyto pocity jsou často zmiňovány

společně s odkazem na hostitele. Toto jsem ještě ověřila při kontrole korelací mezi tématy. STM vytváří korelovaná témata a druhá nejsilnější korelace (0,35) se ukázala právě mezi druhým a osmým tématem, tedy mezi hostitelem a zážitky a emocemi z pobytu. Dá se tedy říci, že z určité části jsou tyto emoce a zážitky spojovány uživateli právě s osobou hostitele. Osoba hostitele se pak ukázala jako důležitá i pro celkové zhodnocení pobytu, kdy mezi těmito dvěma tématy byla hodnota korelace nejvyšší (0,5).⁶⁷ Převážná část zkušeností uživatelů Airbnb je tedy popisována v těchto osmi kategoriích, to jsou ta pro ně nejdůležitější témata. Zatím nicméně nejsou seřazena podle důležitosti, ale pouze detekována, prozkoumána a pojmenována.

Graf 9: Zvolený výsledný model o osmi tématěch – Airbnb



Zdroj: vlastní zpracování

⁶⁷ Třetí nejvyšší korelace je pak mezi pokojem a jeho vybavením (téma 1) a příjezdem a odjezdem (téma 5). Zbylá témata již korelují jen velmi slabě (mezi 0,03 a 0,05). Korelace mezi tématy ukazuje, nakolik se témata v textech objevují společně.

Dalším krokem bylo dopočítání pravděpodobnosti *gamma*, ve které je vyjádřena proporce každého tématu v celkovém korpusu textu.⁶⁸ Je tak možné seřadit témata podle důležitosti, respektive podle toho, jak velký prostor jim uživatelé věnují v popisu svých zkušeností. Hodnoty *gamma* jsou k nahlédnutí také v Tabulce 7 ve sloupci „Proporce tématu“. Zároveň jsou tam již témata seřazena sestupně podle důležitosti. Ukazuje se tedy, že v korpusu recenzí zpracovávaných v této práci věnují uživatelé Airbnb největší prostor, a to téměř čtvrtinu z celku, hodnocení hostitele. Osoba hostitele je pro ně tedy v jejich zkušenosti s Airbnb zdaleka nejdůležitější, a to i přesto že převážnou část korpusu tvoří recenze celých bytů, tedy fyzicky nesdílených s hostitelem. Evidentně ale komunikace s ním, jeho rady a tipy, nakolik se cítí v jeho bytě vítání, je pro ně zásadní. Druhým nejdůležitějším tématem je celkový dojem z ubytování (15 %) a s takřka stejnou proporcí (13 %) následuje celý dům a jeho sousedství a dopravní dostupnost. Pátým tématem z hlediska jemu věnovaného prostoru jsou pak zážitky a emoce z pobytu (10 %), což je pozoruhodné zjištění vzhledem k tomu, že teprve poté následuje pokoj a jeho vybavení (9 %). Ačkoliv rozdíl v proporcích není tak zásadní, pro uživatele Airbnb v tomto korpusu recenzí je evidentně důležitější, jak se v daném ubytování cítí, než jak je vybavené. Posledními tématy z hlediska jim věnovaného prostoru je pak lokalita (9 %) a příjezd a odjezd (9 %).

⁶⁸ Respektive jak velká část dokumentu je asociována s konkrétním tématem [Roberts et al. 2014, str. 2].

Tabulka 7: Identifikovaná témata v recenzích Airbnb řazená podle proporce zastoupení

Téma	Proporce tématu	50 slov s nejvyšší pravděpodobností zařazení k tématu
Topic 2 <i>Hostitel a jeho péče o hosta</i>	0.212	location, stay, host, place, help, apartment, prague, clean, friendly, response, tips, lovely, gave, amazing, questions, places, wonderful, local, needed, welcome, quickly, hospitality, beautiful, quick, visit, provide, exactly, need, jan, kind, information, comfortable, accommodating, described, experience, visiting, jana, useful, expected, eat, communication, perfect, city, accommodation, messages, david, josef, tomas, suggestions, alex
Topic 3 <i>Celkový dojem z ubytování</i>	0.156	apartment, recommend, clean, communication, easy, check, spacious, amazing, stay, space, central, walking, distance, comfortable, stylish, value, fantastic, place, beautiful, quick, amenities, view, host, modern, attractions, awesome, prague, convenient, response, ideal, spot, lovely, smooth, group, centrally, brilliant, process, terrace, incredible, top, daniel, facilities, close, perfect, worth, balcony, sparkling, fabulous, enough, service
Topic 7 <i>Celý dům a jeho sousedství</i>	0.134	nice, flat, clean, centre, city, cosy, house, quiet, comfortable, equipped, place, apartment, need, recommend, stay, kind, area, friendly, fast, close, far, like, new, host, neighborhood, warm, location, carparking, person, cool, tidy, situated, liked, calm, reach, pretty, balcony, cute, prague, near, owner, people, foot, neighbourhood, looks, pictures, bright, beautiful, spend, communication
Topic 6 <i>Dopravní dostupnost</i>	0.129	minute, close, tram, near, centre, city, metro, station, walk, restaurant, public, apartment, transport, easy, area, stop, bus, train, transportation, bar, shops, place, need, supermarket, stay, easily, access, convenient, neighborhood, park, quiet, stops, cafes, prague, clean, food, trams, shopping, far, subway, local, corner, outside, right, grocery, carparking, accessible, couple, ride, front
Topic 8 <i>Zážitky a emoce z pobytu</i>	0.102	prague, place, thank, enjoy, love, stay, home, come, feel, trip, wonderful, beautiful, recommend, lovely, visit, apartment, comfortable, pleasant, needed, martin, happy, return, weekend, explore, family, looking, like, hope, welcome, long, need, decorated, exploring, make, city, location, making, special, pleasure, future, book, visiting, zuzana, clean, petra, base, relax, beautifully, katerina, relaxing
Topic 1 <i>Pokoj a jeho vybavení</i>	0.090	room, kitchen, bed, bathroom, location, place, people, shower, price, noisy, sleep, money, beds, rooms, bedroom, clean, comfortable, large, water, floor, stay, toilet, problem, fine, windows, use, door, wifi, space, light, hot, enough, apartment, open, private, basic, pretty, towels, issue, like, guests, window, shared, issues, loud, street, bedrooms, expected, group, size
Topic 4 <i>Lokalita ubytování</i>	0.088	old, town, location, walk, apartment, square, right, minute, restaurant, walking, street, bridge, quiet, building, distance, charles, castle, bar, prague, comfortable, plenty, river, heart, close, tourist, large, sights, stairs, area, spacious, sites, steps, elevator, beautiful, high, stay, machine, floor, attractions, coffee, space, wenceslas, safe, family, walked, cafe, busy, including, views, lift
Topic 5 <i>Přijezd a odjezd</i>	0.088	arrive, airbnb, like, apartment, experience, provide, first, late, check, meet, airport, stay, coffee, make, feel, left, air, early, keys, luggage, find, use, extra, leave, hot, breakfast, morning, flight, found, key, asked, appreciated, star, tea, hour, host, give, machine, touch, food, taxi, pick, available, checked, booked, door, booking, service, showed, evening

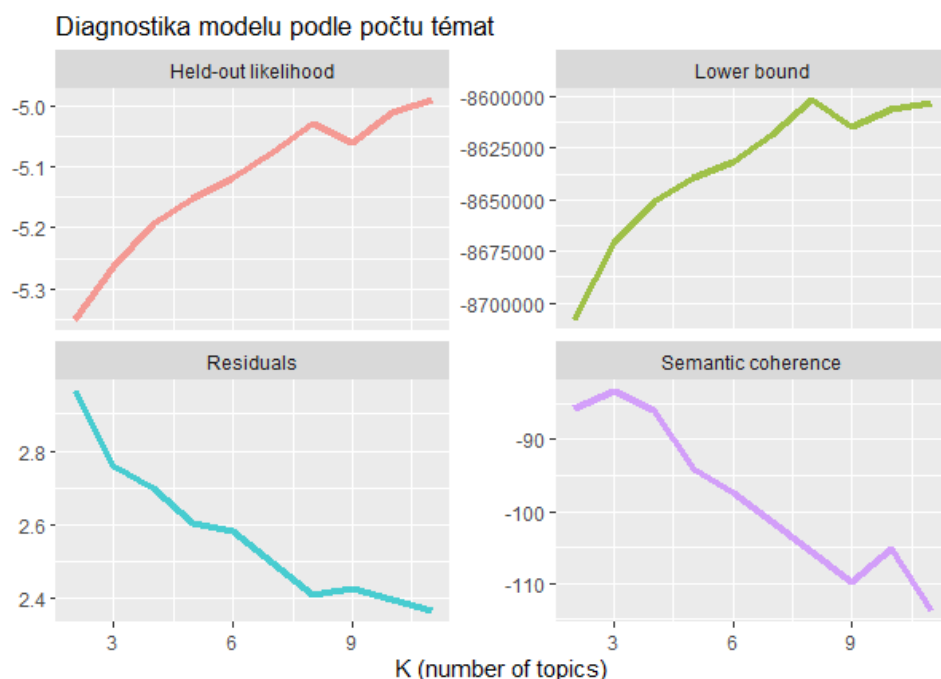
Shrnutí

V korpusu recenzí uživatelů Airbnb bylo detekováno osm nejdůležitějších témat, která se napříč recenzemi opakují. Největší prostor při popisu svých zkušeností s ubytováním přes Airbnb věnují uživatelé osobě hostitele. Často hodnotí komunikaci s ním, jeho vlastnosti, pohostinnost a nápomocnost, co týče poskytnutých tipů a informací o místě. Při odkazování na hostitele často užívají jeho křestní jméno. Druhou největší část jejich zkušeností tvoří zhodnocení celkového dojmu z ubytování, kde uživatelé případně dávají doporučení, přičemž toto téma středně silně koreluje s tématem hostitele, tedy jeho osoba má významný vliv i na tento celkový dojem, jinak ale v menší míře obsahuje výrazy spojené i se všemi ostatními tématy. Třetím nejzastoupenějším tématem je dům a jeho sousedství ve smyslu čistoty, útulnosti, vřelosti a podobně. Dopravní dostupnost místa pak také tvoří samostatné téma. Hned po něm následuje téma tvořené slovy označujícími emoce a zážitky. Evidentně i ty jsou významnou součástí zkušenosti uživatele Airbnb a opět silně korelují s tématem hostitele, tedy patrně má na těchto zážitcích a emocích svůj podíl. Teprve poté je z hlediska důležitosti pokoj a jeho vybavení, tedy hostitel, zážitky či místo pobytu a jeho dostupnost má pro uživatele Airbnb při popisu jeho zkušeností větší význam než samotný pokoj. Následuje už pak jen lokalita ve smyslu vzdálenosti od turisticky významných míst a proces příjezdu a odjezdu.

3.3.3 Modelování témat v korpusu recenzí tradičních ubytování

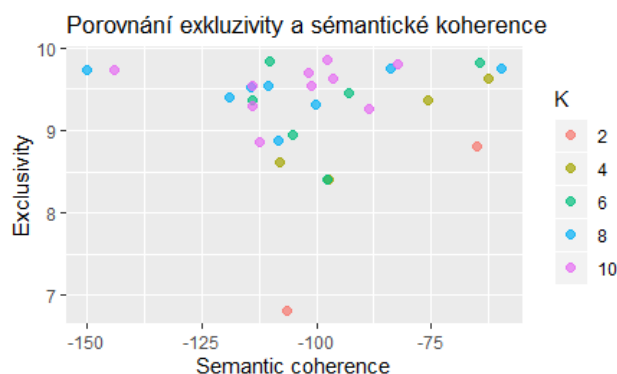
Stejný postup, tak jak byl popsán u korpusu z Airbnb, jsem aplikovala také na korpus se zkušenostmi uživatelů s hromadnými ubytovacími kapacitami přes Booking. Vytvořila jsem tedy deset modelů se dvěma až jedenácti tématy. Vizualizované metriky pro diagnostiku může čtenář vidět v Grafu 10. Diagnostiky naznačují, že by nejvhodnější počet témat mohl být osm, u devíti se diagnostiky lehce horší, pak deset, případně víc. Cílem je maximalizovat *held-out likelihood* a *lower bound*, naopak minimalizovat *rezidua* a také najít vhodný kompromis mezi *sémantickou koherencí* a *exkluzivitou* (viz Graf 11 na str. 58). Stejně jako u Airbnb jsem všechny tyto možnosti namodelovala a posoudila i pouhým okem z hlediska interpretovatelnosti slov témata tvořící a věcného smyslu, přičemž osm témat bylo koherentních a přinášela ještě novou informaci, při větším počtu se pak už začínala tvořit témata, které se opakovala nebo byla nejasná. Rozhodla jsem se tedy pro počet osmi témat.

Graf 10: Diagnostika modelu podle počtu témat – Booking



Zdroj: vlastní zpracování

Graf 11: Porovnání exkluzivity a sémantické koherence témat – Booking



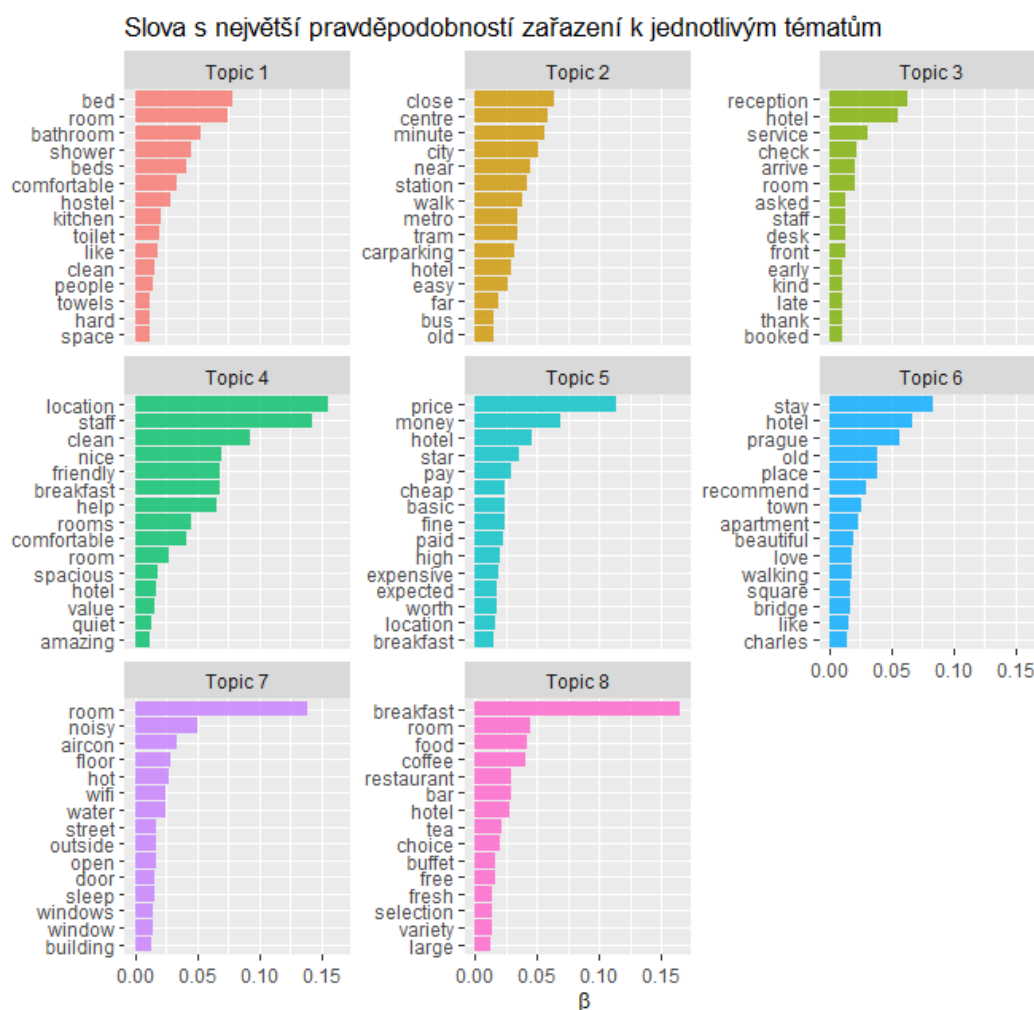
Zdroj: vlastní zpracování

Výsledný model, tentokrát redukující obsah zkušeností uživatelů Bookingu, jsem opět vizualizovala do grafu (viz Graf 12 na str. 60) pouze s patnácti slovy s nejvyšší pravděpodobností (*beta*) náležitosti k tématu. Detailnější přehled slov tvořících jednotlivá témata je pak níže v Tabulce 8 na str. 61. Dalším krokem bylo pojmenování těchto témat. První téma jsem pojmenovala jako **pokoje a jeho vybavení** (Topic 1), je velice podobné prvnímu tématu ve zkušenostech s Airbnb, také obsahuje slova týkající se prostoru ubytování (*pokoje, koupelna, sprcha, kuchyně, toaleta* aj.) a základního vybavení ubytování (*postel, ručníky, polštáře, matrace, závěsy, prostěradlo*), doplněné o přídavná jména s tím spojená (*pohodlný, čistý, špinavý, nepohodlný, hezký*). Druhé téma je obdobou šestého tématu u Airbnb, obsahuje slova týkající se dopravy (*metro, tramvaj, zastávka, autobus, vlak, doprava, auto*) a vzdálenosti (*blízko, centrum, minuta, město, poblíž* apod.). Pojmenovala jsem ho také **dopravní dostupnost** (Topic 2). Třetí téma je oproti tématům objevujícím se u Airbnb specifické, zahrnuje slova jako *recepce, hotel, servis, zaměstnanci, personál, laskavý, hrubý, angličtina/anglicky, taxi, poskytovat, karta, navíc, zdarma/volný*. Zvolila jsem pojmenování **personál a jeho služby** (Topic 3). Čtvrté téma je podobné třetímu tématu u Airbnb, také oproti jiným není tak jednoznačné, obsahuje slova ze všech ostatních témat, nicméně v kombinaci se slovy jako *hodnota, atmosféra, doporučit, stěžovat* nebo *bezchybný*. Pojmenovala jsem ho také **celkový dojem z ubytování** (Topic 4). Páté téma tvořené slovy ohledně ceny ubytování (*cena, peníze, platit, levný, drahý* a další) je poměrně překvapivé. Nejen proto, že se objevuje ve zkušenostech s Bookingem natolik často, že tvoří samostatné téma, ale především že to není tématem ve zkušenostech s Airbnb. Přitom častým argumentem oponentů Airbnb je, že se Airbnb daří

zejména proto, že si díky nedostatečné regulaci mohou hostitelé dovolit nasadit nižší ceny než tradiční ubytovací kapacity. Vzniká otázka, proč tedy uživatelé Airbnb při popisu své zkušenosti s ubytováním nezmiňují cenu natolik pravidelně, aby byla vyhodnocena jako důležité téma? A proč naopak uživatelé tradičních ubytovacích kapacit ano? Neznamená to samozřejmě automaticky, že cena není pro uživatele Airbnb důležitá, nejspíše hraje při výběru ubytování roli, ale proč ji pak ve svých zkušenostech nereflektují a věnují pozornost spíše jiným kategoriím? To mohou být otázky pro další výzkumné studie. Páté téma jsem tedy pojmenovala jako **cena vs hodnota** (Topic 5). Šesté téma je obdobou čtvrtého tématu u Airbnb, obsahuje velice podobná slova týkající se buď nějakého konkrétního místa (*staré, město, náměstí, most, karlův, hrad, řeka*) nebo vzdálenosti (*pěší, vzdálenost, přímo, pěšky, minuta*), proto jsem ho také pojmenovala **lokalita** (Topic 6). Sedmé téma se opět týká pokoje, nicméně tentokrát již nikoliv základního vybavení, ale spíše vybavení navíc či pohodlí na pokoji (např. *hlučný, klimatizace, podlaha/poschodí, horko, wifi, okna, televize, výtah, chladno, internet, topení, teplota*). Pojmenovala jsem ho **(dis)komfort na pokoji** (Topic 7). Poslední osmé téma jsem pojmenovala jako **stravování** (Topic 8), jelikož je tvořeno slovy *snídaně, jídlo, káva, restaurace, bar, čaj, bufet* a dalšími podobnými. Kolem těchto osmi témat především se točí zkušenosti uživatelů tradičních ubytovacích kapacit. Opět jsem kontrolovala i korelovanost mezi jednotlivými tématy, přičemž nejvyšší, i když stále slabá, korelace (0,12) je mezi prvním a sedmým tématem, tedy mezi vybavením pokoje a (dis)komfortem na pokoji, což je předpokládatelné, že se tato témata budou objevovat v textech spolu.⁶⁹ Dalším krokem je seřazení témat podle důležitosti, zatím je pořadí náhodné.

⁶⁹ Další témata již nekorelují vůbec, maximální hodnota Pearsonova korelačního koeficientu je 0,03. Neobjevují se tedy v textech nijak systematicky společně.

Graf 12: Zvolený výsledný model o osmi tématech – Booking



Zdroj: vlastní zpracování

V Tabulce 8 jsou témata seřazena sestupně podle jejich důležitosti, respektive podle *gamma* pravděpodobnosti, tedy proporce, ze které tvoří celkový korpus zkušeností s hromadnými ubytovacími kapacitami přes Booking. Nejčastěji tedy uživatelé věnují prostor v recenzích zhodnocení celkového dojmu z ubytování, což tvoří celou čtvrtinu textů. Další čtvrtinu pak věnují hodnocení pokoje, ať už jeho vybavení (13 %) nebo (dis)komfortu (13 %). Čtvrtým nejdůležitějším tématem pro ně je lokalita (12 %), následuje personál a jeho služby (12 %), dostupnost (12 %), stravování (9 %) a posledním tématem z hlediska jeho proporce je cena ubytování (5 %).

Tabulka 8: Identifikovaná témata v recenzích Booking řazená podle proporce zastoupení

Téma	Proporce tématu	50 slov s nejvyšší pravděpodobností zařazení k tématu
Topic 4 <i>Celkový dojem z ubytování</i>	0.254	location, staff, clean, nice, friendly, breakfast, help, rooms, comfortable, room, spacious, hotel, value, quiet, amazing, central, facilities, lovely, modern, fantastic, polite, kind, superb, design, atmosphere, brilliant, tidy, convenient, awesome, decor, recommend, old, new, interior, complain, pleasant, fabulous, warm, equipped, liked, efficient, personnel, stylish, outdated, view, centrally, maintained, attractions, spotless, decorated
Topic 1 <i>Pokoj a jeho vybavení</i>	0.126	bed, room, bathroom, shower, beds, comfortable, hostel, kitchen, toilet, like, clean, people, towels, hard, space, dirty, double, pillows, uncomfortable, enough, bath, single, light, bathrooms, bedroom, size, use, mattress, large, broken, curtains, showers, sleep, need, bathtub, sheets, cleaning, changed, new, feel, nice, towel, shared, hair, soft, toilets, separate, wall, used, head
Topic 7 <i>(Dis)komfo rt na pokoji</i>	0.126	room, noisy, aircon, floor, hot, wifi, water, street, outside, open, door, sleep, windows, window, building, tv, elevator, poor, cold, people, morning, warm, loud, hear, lift, rooms, fan, stairs, fridge, problem, smell, inside, location, internet, kettle, difficult, heating, top, connection, view, first, summer, slow, temperature, entrance, closed, need, air, worked, walls
Topic 6 <i>Lokalita</i>	0.124	stay, hotel, prague, old, place, recommend, town, apartment, beautiful, love, walking, square, bridge, like, charles, distance, feel, right, view, wonderful, lovely, amazing, restaurant, location, castle, walk, visit, places, property, family, welcome, attractions, area, come, building, quiet, river, home, hostel, suite, tourist, expectations, minute, trip, terrace, bar, heart, meet, sights, apartments
Topic 3 <i>Personál a jeho služby</i>	0.119	reception, hotel, service, check, arrive, room, asked, staff, desk, front, early, kind, late, thank, booked, first, airport, english, like, morning, help, checkin, gave, rude, ask, experience, make, left, taxi, booking, guests, stay, lady, leave, speak, welcome, give, hour, personnel, wanted, provide, wait, card, find, found, cleaning, free, checked, come, extra
Topic 2 <i>Dopravní dostupnost</i>	0.116	close, centre, minute, city, near, station, walk, metro, tram, carparking, hotel, easy, far, bus, old, public, restaurant, town, train, transport, stop, area, car, access, walking, location, airport, distance, pool, park, quiet, convenient, shopping, right, find, need, reach, shops, transportation, free, trams, stops, street, easily, subway, foot, front, nice, use, sauna
Topic 8 <i>Stravování</i>	0.088	breakfast, room, food, coffee, restaurant, bar, hotel, tea, choice, buffet, free, fresh, selection, variety, large, options, delicious, eat, plenty, location, quality, available, nice, tasty, eggs, included, area, drinks, machine, dinner, limited, choices, position, lobby, choose, fruit, drink, comfortable, bread, served, liked, spa, wide, including, cold, making, complimentary, poor, juice, enough
Topic 5 <i>Cena vs hodnota</i>	0.047	price, money, hotel, star, pay, cheap, basic, fine, paid, high, expensive, expected, worth, location, breakfast, decent, quality, simple, expect, room, standard, accommodation, reasonable, euro, cost, rate, euros, extra, low, perfect, charged, included, exchange, place, hotels, option, facilities, amount, ok, complain, deal, fair, needs, cash, need, pretty, proximity, booking, budget, higher

Zdroj: vlastní zpracování

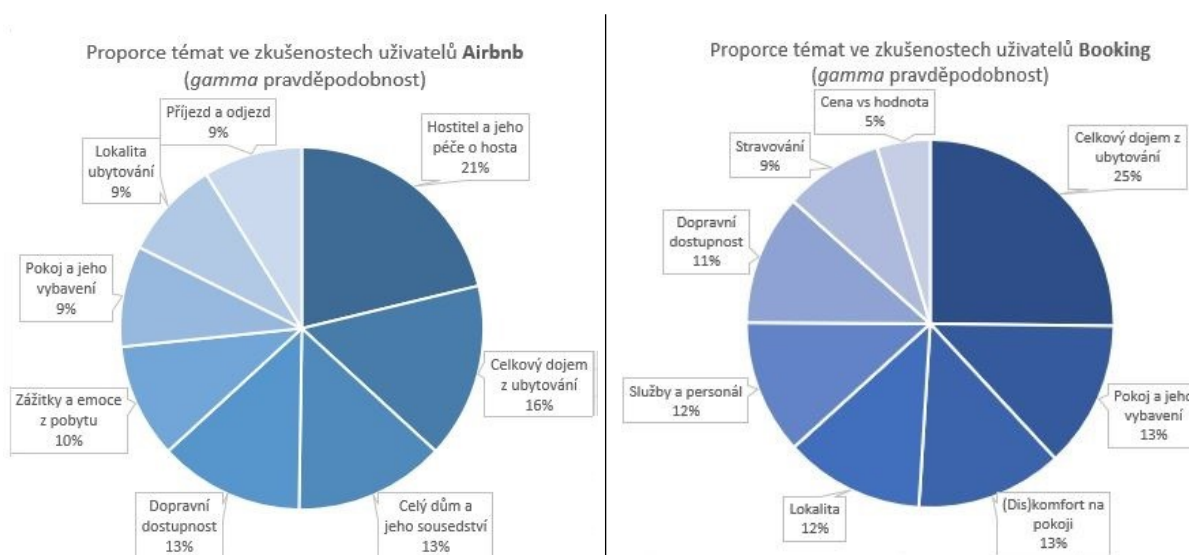
Shrnutí

V korpusu recenzí hromadných ubytovacích zařízení přes Booking se také ukázalo osm témat jako nejvhodnější počet. Tato témata ale nejsou stejná jako u korpusu z Airbnb. Nejčastěji zmiňovaným tématem uživateli hromadných ubytovacích zařízení je celkový dojem z ubytování, tedy jakési zhodnocení celku a případné doporučení. Druhým tématem z hlediska zastoupení je pak pokoj a jeho základní vybavení jako jsou postel, sprcha, toaleta, ručníky a další, přičemž hned následuje téma ohledně spíše nadstandardního vybavení v podobě klimatizace či wifi spojeného s komfortem či diskomfortem na pokoji. Tato dvě témata spolu mírně korelují. Obdobně důležitá je také lokalita, tedy vzdálenost ubytování od turisticky významných míst, stejně tak personál, jaké služby poskytuje a jaké kvality jsou. Následuje dopravní dostupnost zmiňovaná v souvislosti s centrem města, stravování, u kterého je kladen velký důraz na výběr, a posledním tématem je cena v porovnání s hodnotou.

3.3.4 Komparace obsahu zkušeností se sdíleným a tradičním ubytováním

Posledním, podstatným krokem je srovnání vytvořených modelů redukujících obsah zkušeností uživatelů s Airbnb a s hromadnými ubytovacími kapacitami. V obou korpusech bylo nalezeno osm nejdůležitějších a nejkoherentnějších témat, jak jsem detailně rozebrala výše, nicméně tato témata a jejich zastoupení v textech se liší. Pro přehlednou komparaci jsem pravděpodobnostní zastoupení témat (*gamma* pravděpodobnost) načrtla do Grafu 13.

Graf 13: Komparace modelovaných témat a jejich proporcí – Airbnb vs Booking



Zdroj: vlastní zpracování

Zvlášť v obou korpusech recenzí jsem s využitím STM detekovala osm témat, přičemž komparace ukázala, že čtyři z nich jsou obdobné, čtyři další pak korpuse odlišují.⁷⁰ Z hlediska proporčního zastoupení těchto témat, ta společná tvoří u Airbnb téměř polovinu korpuse, u Bookingu pak téměř dvě třetiny, zbytek obsahu korpuse se liší. Mezi společná témata patří popis celkového dojmu z ubytování, které má v obou korpusech značné zastoupení a kde, jak bylo představeno výše, uživatelé shrnují jednotlivé přednosti a nedostatky ubytování napříč oblastmi a případně dávají doporučení. U Bookingu tvoří dokonce čtvrtinu celého textu, a je

⁷⁰ Rozlišování mezi společným a unikátním obsahem je samozřejmě určitým zjednodušením, stále se výzkum pohybuje pouze v rovině pravděpodobnosti, tedy ve skutečnosti se jedná o nějaké oblasti textu, které jsou si více či méně podobné, není jednoduché (pomyslně) nakreslit jasnou dělící čáru.

tak nejzastoupenějším tématem, u Airbnb pak zhruba 16 % a je druhým nejzastoupenějším. Dalším společným znakem je vyjadřování se k dopravní dostupnosti ubytování, které má zhruba stejné zastoupení v popisu zkušeností obou skupin uživatelů, tvoří přibližně 12–13 % celých textů a má i obdobnou relativní důležitost vzhledem k dalším tématům – u Airbnb je na čtvrtém místě, u Bookingu na šestém. Stejně tak lokalita, která u obou sad recenzí tvoří zhruba desetinu textu (9 a 12 %), u Airbnb je na sedmém, u Bookingu na čtvrtém místě. V těchto ohledech jsou si zkušenosti uživatelů, jak je sami reflektují, vcelku podobné. Posledním společným tématem je pak pokoj a jeho vybavení. Ačkoliv zastoupení toto téma nemá výrazně odlišné, a sice u Airbnb je to 9 % a u Bookingu 13 %, z hlediska relativní důležitosti se ale liší významně. Zatímco u Airbnb je celá řada pro uživatele důležitějších témat a pokoj a jeho vybavení je mezi třemi posledními, u Bookingu je naopak hned druhým nejdůležitějším. Kromě toho je to u Airbnb jediné téma týkající se vybavení pronajímaného prostoru, zatímco u Bookingu je se stejným zastoupením (13 %) navíc téma obsahující slova týkající se komfortu či diskomfortu na pokoji. Dohromady tedy je v popisu zkušeností uživatelů hromadných ubytovacích zařízení přes Booking věnována čtvrtina textu jen pro popis pokoje a jeho vybavení, ať už standardního (postele, polštáře atd – viz obsah tématu Pokoj a jeho vybavení) nebo nadstandardního umožňujícího vyšší komfort (klimatizace, wifi, situovanost oken atd – viz obsah tématu (Dis)komfort na pokoji), u Airbnb je to jen desetina. Uživatelé Airbnb naopak věnují v popisech svých zkušeností prostor celému sousedství či domu, ve kterém se ubytování nachází, nicméně nikoliv vyloženě z hlediska vybavení, ale spíše toho, jak prostor působí (téma Celý dům a jeho sousedství obsahuje slova jako vřelý, útulný, přátelský). Obecně se dá říci, že uživatelé Airbnb jsou v popisu svých zkušeností orientovaní výrazně více na kvality ubytování nemateriální povahy, zatímco uživatelé Bookingu se soustředí na kvality materiální. Tato tendence je evidentní i v dalších tématech, kterými se korpusy naopak liší. Pátým nejvíce zastoupeným tématem, které předběhlo například právě i vybavení pokoje, jsou u Airbnb zážitky a emoce z pobytu. Toto téma obsahuje výrazy, které jsou ukázkovým příkladem nemateriálních hodnot (například užívat, cítit se, domov, speciální, příjemný a řada dalších viz výše v Tabulce 7) a tvoří desetinu celého korpusu. U Bookingu žádné takové téma detekováno nebylo. Dalším tématem, které odlišuje korpus recenzí Airbnb od hromadných ubytovacích kapacit, je Hostitel a jeho péče o hosta. Ačkoliv korpus z Bookingu také obsahuje téma, kde jsou výrazy spojené s lidským faktorem v ubytování (nazvané Personál a služby), obsah i zastoupení těchto témat jsou jiné. U Airbnb je téma hostitele na prvním místě a tvoří téměř čtvrtinu korpusu, u Bookingu tvoří téma Personál a služby necelou desetinu obsahu textů a je

na místě pátém. Co se pak týče obsahu těchto dvou témat, hostitele uživatelé Airbnb zmiňují ve spojitosti s komunikací s ním, jeho vstřícností a pohostinností, tedy vlastnostmi, personál zmiňují uživatelé Bookingu převážně ve spojitosti s poskytovanými službami (jazyková vybavenost, průběh příjezdu, taxi a další). Mnohem osobnější vztah k hostiteli je pak demonstrován i hojným užíváním jeho křestního jména, která jsou také obsahem tohoto tématu, což se u hromadných ubytovacích zařízení ve vztahu k personálu neděje. Dále uživatelé Bookingu reflektují jako důležité stravování v hotelu, kterému věnují desetinu prostoru ve svých recenzích, u Airbnb se takové téma neobjevilo, nicméně to je očekávatelné vzhledem k tomu, že ubytování přes Airbnb většinou nezahrnuje stravování. Stejně tak naopak uživatelé Airbnb věnují desetinu svých textů popisu procesu příjezdu a odjezdu z ubytování, zatímco u Bookingu toto samostatným tématem není. Nabízí se logické vysvětlení, že hromadná ubytovací zařízení většinou mají recepce, tedy proces příjezdu a odjezdu je standardizovaný, zatímco u Airbnb jsou klíče předávány různým způsobem, záleží na hostiteli, jak předání zorganizuje a hladký průběh evidentně hraje roli. Posledním tématem, které ještě nebylo zmíněno a které tvoří 5 % u recenzí z Bookingu a u Airbnb se neobjevuje, je cena v porovnání s hodnotou ubytování. Je překvapivé, že cena se podle obsahu recenzí zdá být pro uživatele hromadných ubytovacích zařízení důležitější než pro uživatele sdíleného ubytování. Častým argumentem přitom je, že hlavní konkurenční výhodou Airbnb je možnost nastavit nižší ceny. Podle obsahu recenzí to nicméně vypadá, že hlavní výhodou je spíše uspokojování nehmotných potřeb, jako je osobní přístup, komunikace a zážitek.

Shrnutí

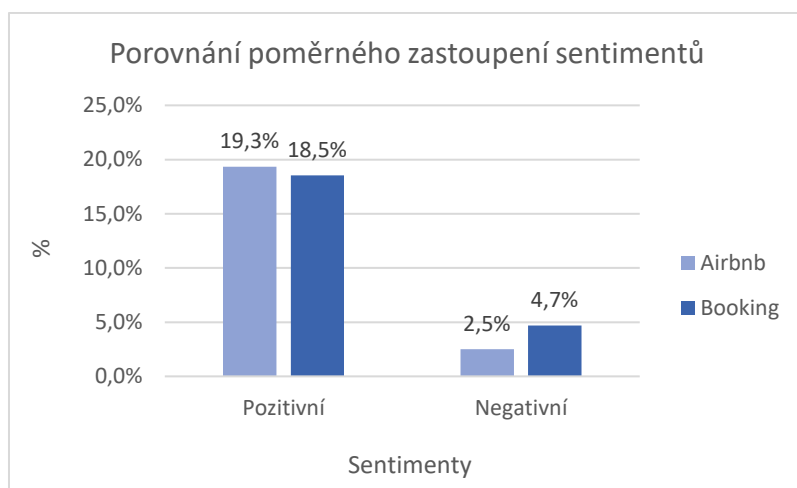
Uživatelé obou typů ubytování, sdíleného přes Airbnb i tradičního hromadného přes Booking, věnují v popisech svých zkušeností pozornost celkovému zhodnocení dojmu z ubytování, dopravní dostupnosti, lokalitě a základnímu vybavení pokoje. První tři jmenovaná témata se výrazně neliší ani prostorem, který je jim uživateli věnován, nicméně poslední z nich – vybavení pokoje – je relativně důležitější pro uživatele Bookingu než uživatele Airbnb. Kromě toho uživatelé Bookingu zmiňují kromě základního vybavení pokoje ještě vybavení nadstandardní, související s komfortem na pokoji, a tak ve výsledku věnují vybavení pokoje dvakrát větší prostor než uživatelé Airbnb. S komfortem v ubytování souvisí i další témata objevující se ve zkušenostech uživatelů Bookingu, přes desetinu svých textů věnují službám a personálu tyto služby vykonávající a necelou desetinu textů pak stravování v ubytování.

Uživatelé Airbnb věnují méně pozornosti službám a vybavení a mnohem více se při popisu zkušeností zaměřují na své pocity a zážitky z ubytování. Stejně tak při popisu hostitele se méně zaměřují na konkrétní služby, které jim poskytl, ale více na jeho vlastnosti a komunikaci s ním, přičemž velmi často ho označují křestním jménem. Uživatel Bookingu mluví o personálu mnohem méně osobně. Posledními tématy odlišující korpusy jsou pak u Bookingu cena a hodnota ubytování, u Airbnb proces příjezdu a odjezdu a také popis celého domu, ve kterém se ubytování nachází, a jeho sousedství. Společná témata tvoří celkově v korpusu textů z Airbnb menší podílnou část než v korpusu z Bookingu. Obecně se dá říci, že uživatelé Bookingu ve svých zkušenostech častěji popisují komfort, služby a jejich výhodnost, zatímco uživatelé Airbnb věnují pozornost spíše osobnímu kontaktu, prožitku či zážitku a příjemnosti domu a sousedství.

3.4 Porovnání obsahu recenzí z hlediska sentimentu

V předchozích kapitolách byly popsány četnosti jednotlivých užívaných výrazů a detekována témata, do kterých se shlukují. Byla tak vytvořena již poměrně detailní představa o obsahu obou korpusů a odlišnostech i podobnostech mezi nimi. Pro komplexnější popis zkušeností uživatelů obou typů ubytování nicméně zbývá ještě zhodnotit sentimenty, které se v jejich textech objevují. Zatím jsou detekována témata, která se opakují, ale zůstává otázka, která z nich zmiňují spíše s pozitivní a která spíše s negativní konotací? S užitím analýzy sentimentu zhodnocuji, která témata nejvíce přispívají k pozitivnímu či negativnímu sentimentu, stejně tak který korpus je celkově pozitivnější či negativnější. To doplňuji také o podrobnější analýzu emocí jako jsou očekávání, vztek, důvěra, překvapení, smutek, radost, strach a znechucení.

Graf 14: Porovnání poměrného zastoupení negativního a pozitivního sentimentu



Zdroj: vlastní zpracování

S využitím předpřipraveného lexikonu *nrc* [Mohammad a Turney 2010] jsem v softwaru R detekovala zvlášť slova s pozitivním a negativním sentimentem, v každém korpusu odděleně.⁷¹ Jednoduchým podílem počtu slov s daným sentimentem oproti celkovému počtu slov v souboru je tak možné dopočítat přibližné procentuální zastoupení sentimentu (viz

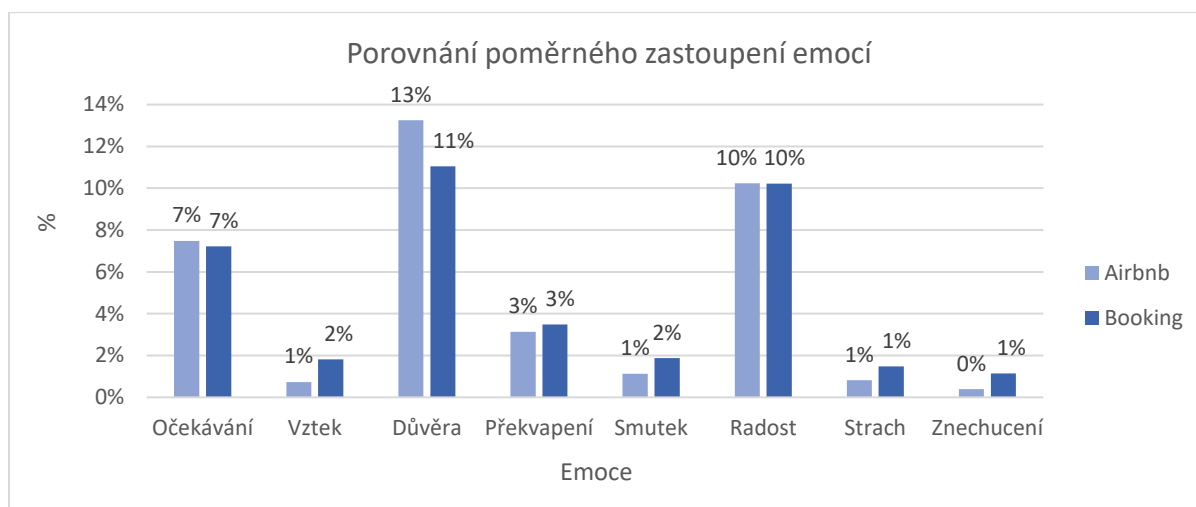
⁷¹ Nejčtenější slova s pozitivním a negativním sentimentem, která byla detekována v jednotlivých souborech, jsou k nahlédnutí v Příloze 9 na straně 112.

Graf 14, str. 67). Zbylá slova jsou z toho hlediska neutrální, což je v přirozeném jazyce většina [Silge a Robinson 2017]. V obou souborech recenzí výrazně převažují výrazy s pozitivním sentimentem nad těmi negativními. Uživatelé obou typů ubytování v tomto souboru recenzí mají tedy tendenci věnovat větší prostor pozitivním aspektům své zkušenosti než těm negativní. S pozitivním sentimentem je v obou případech zhruba pětina plnovýznamového textu.⁷² Zatímco pozitivní jsou uživatelé obou typů ubytování takřka stejně, u negativního sentimentu je výrazný rozdíl. Uživatelé tradičních ubytovacích zařízení užívají při popisu svých zkušeností slova s negativním sentimentem dvakrát častěji. Nemusí to nicméně nutně znamenat, že by zkušenost s tradičními modely ubytování byla negativnější. Roli může hrát to, že u Airbnb je vztah s hostitelem, který recenzi obdrží, mnohem méně anonymní, proto také potenciálně může být těžší jeho ubytování kritizovat. Pro ještě detailnější vhled do uživatelů popisovaných emocí jsem dopočítala také poměr zastoupení slov spojených s emocí očekávání, vzteku, důvěry, překvapení, smutku, radosti, strachu a znechucení (viz Graf 15 na str. 69). V kontextu této práce je zajímavá zejména emoce důvěry, která, jak bylo řečeno v teoretické části, je často skloňovaným tématem v souvislosti se sdílenou ekonomikou. Skutečně se ve zkušenostech uživatelů Airbnb objevují slova spojená s emocí důvěry častěji než u uživatelů tradičních ubytovacích zařízení, nicméně rozdíl zdaleka není tak výrazný, jak by se dalo očekávat. Emoce pozitivního sentimentu (očekávání, překvapení, radost) jsou zastoupené v obou souborech takřka stejně, přičemž nejčastěji z nich uživatelé vyjadřují v obou případech právě důvěru a radost. Emoce negativní se pak objevují výrazně méně častěji a u všech z nich (vztek, smutek, strach, znechucení) je zastoupení ve zkušenostech uživatelů tradičního ubytování dvakrát vyšší než u uživatelů ubytování sdíleného.⁷³

⁷² I do této analýzy jsem vstupovala s upraveným souborem, očištěným od tzv. stop slov.

⁷³ Nejčastější slova, která byla na základě lexikonu přiřazena ke konkrétním emocím jsou k nahlédnutí v mraku slov v Příloze 7 na str. 111 pro korpus recenzí Airbnb a v Příloze 8 pro korpus recenzí tradičních ubytovacích zařízení.

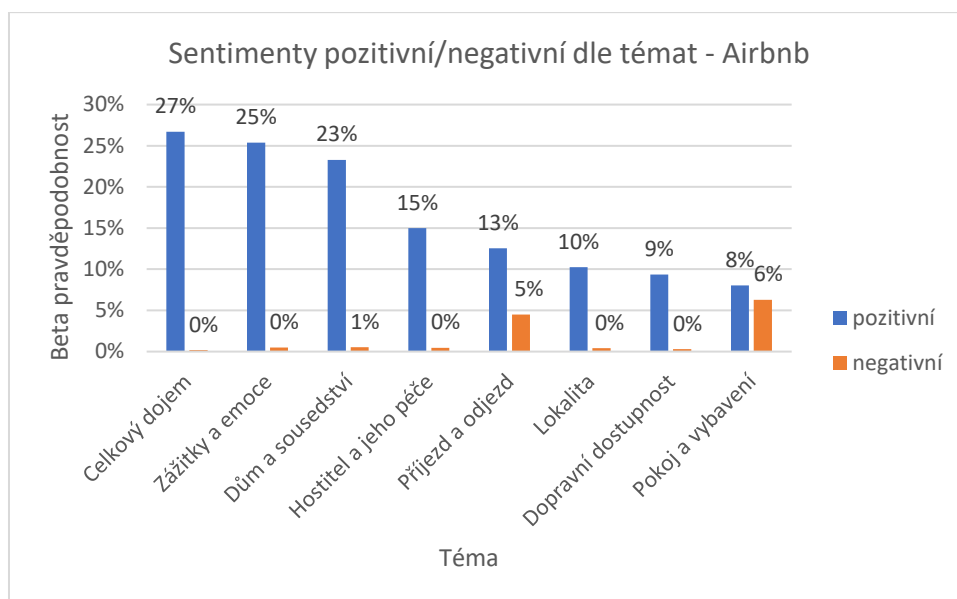
Graf 15: Porovnání poměrného zastoupení emocí – Airbnb vs Booking



Zdroj: vlastní zpracování

Dalším krokem bylo zjistit, nakolik k jednotlivým sentimentům (pozitivní/negativní) přispívají témata, která se v textech systematicky objevují a která byla výše detekována. Slova pojící se k jednotlivým tématům a jejich *beta* pravděpodobnosti (tj. pravděpodobnost náležitosti k tématu) jsem vyfiltrovala a opakovala postup pro přiřazení sentimentů, tentokrát zvlášť na jednotlivých tématech, která tvoří celkový text. Součet *beta* pravděpodobností slov s pozitivním a zvlášť negativním sentimentem jsem zanesla do Grafu 16 na str. 70 pro případ Airbnb a do Grafu 17 na str. 71 pro případ tradičních ubytovacích zařízení přes Booking. I v této analýze se ukazuje, že přidanou hodnotou Airbnb je především celkový dojem z ubytování, zážitky a emoce, dům a jeho sousedství, tedy celé místo, ve kterém host pobývá, a také osoba hostitele. Tato témata mají nejvýraznější podíl na pozitivním sentimentu, uživatelé v souvislosti s nimi nejčastěji užívají slova s pozitivní konotací a takřka žádná s konotací negativní. Naopak nejčastěji vyjadřují nespokojenost v souvislosti s pokojem a jeho vybavením a s procesem příjezdu a odjezdu. Z toho se dá vyvozovat, že, alespoň v tomto souboru recenzí, ubytování přes Airbnb uživatelům skutečně nabídlo spíše postmateriální hodnoty v podobě lidského kontaktu a prožitků, se kterými jsou spokojeni, přičemž ty materiální jako vybavení pokoje možná pokulhávají. Ještě zajímavější je tento výsledek v kontrastu se zkušenostmi s tradičními modely ubytování, které jsou z hlediska sentimentu téměř přesně opačné.

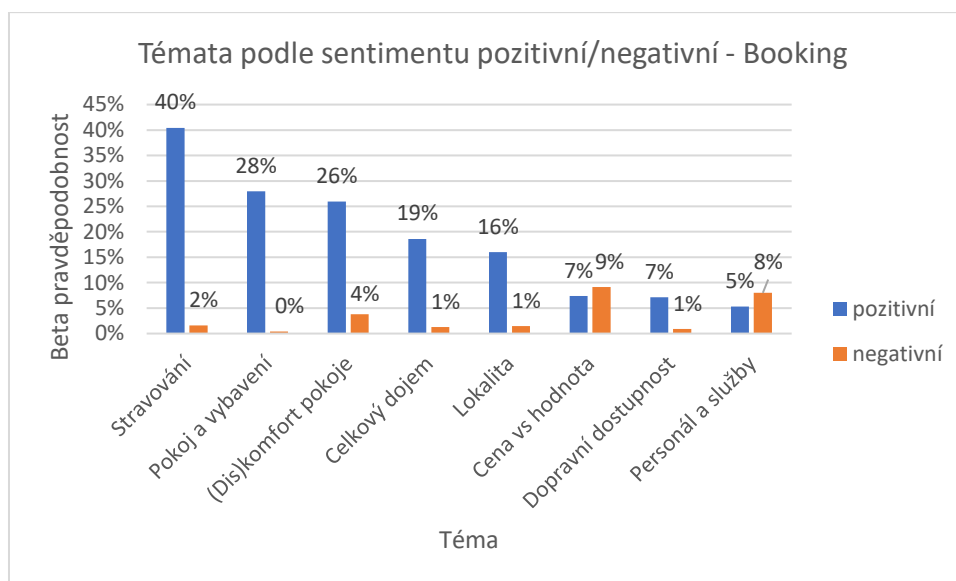
Graf 16: Sentimenty pozitivní/negativní dle témat – Airbnb



Zdroj: vlastní zpracování

Pokoj, jeho vybavení a komfort si u uživatelů tradičních modelů ubytování stojí velmi vysoko z hlediska pozitivní konotace, nicméně zdaleka nejpozitivnější výrazy užívají v souvislosti se stravováním. Na opačné straně škály je naopak personál a jeho služby, o kterém dokonce častěji hovoří negativně než pozitivně. Stejně tak i cenu zmiňují spíše v negativním než pozitivním smyslu, což ale bohužel není možné srovnat s vyjadřováním uživatelů Airbnb o ceně, jelikož ti ji nezmiňují tak systematicky často, aby tvořila celé téma. Jak bylo ale řečeno, obecně služby sektoru sdílené ekonomiky bývají levnější, Airbnb není výjimkou. Soudím tedy, že modely tradičního ubytování přináší uživatelům spíše hodnoty materiální jako právě stravování a komfortně vybavený pokoj, zatímco ty postmateriální se v první řadě ve zkušenostech uživatelů nevyskytují tak často jako u uživatelů Airbnb, ale v podobě alespoň zmiňovaného personálu a jeho služeb i tak evidentně uživatele neuspokojují a spíše o nich mluví negativně. Celkový dojem ale pak hodnotí spíše pozitivně. Dopravní dostupnost je na tom v obou souborech z hlediska sentimentu velmi podobně, lokalitu pak zmiňují uživatelé tradičních ubytování v pozitivním smyslu přibližně o třetinu častěji.

Graf 17: Sentimenty pozitivní/negativní dle témat – Booking



Zdroj: vlastní zpracování

Shrnutí

Ukázalo se, že nejen že uživatelé modelu sdíleného ubytování přes Airbnb věnují při popisu svých zkušeností větší prostor postmateriálním hodnotám, ale také tyto hodnoty zmiňují v nejvíce pozitivním smyslu. Nejnegativněji naopak hovoří o samotném pokoji a jeho vybavení. Zkušenosti uživatelů tradičních modelů ubytování jsou přesně opačné. K popisu pokoje, jeho vybavení a komfortu užívají největší počet slov s pozitivní konotací, výrazným zdrojem pozitivní zkušenosti pro ně bývá také stravování. Nejnegativněji naopak hovoří o lidském faktoru, tedy personálu a jeho službách. Celkově ale ve zkušenostech uživatelů obou typů ubytování významně převažuje pozitivní sentiment nad negativním. Negativní sentiment je v obou souborech zastoupen minimálně. Z hlediska konkrétních emocí je pak zajímavý rozdíl v zastoupení emoce důvěry, která se ve zkušenostech uživatelů Airbnb objevuje mírně častěji.

3.5 Diskuse nad výsledky

Výzkum naplnil stanovený cíl práce, a sice odhalit odlišnosti ve spotřebitelské zkušenosti s modely sdíleného ubytování od těch tradičních. V této závěrečné diskusi věnuji pozornost teoretickým i praktickým implikacím svého výzkumu a porovnávám své výsledky s výsledky již existujících podobně orientovaných studií. Polemizuji také nad nedostatky a limity zvolených postupů i nad možnostmi dalšího výzkumu. V závěru diskuse pak předkládám své úvahy o hlavních přínosech této práce.

Dle stanovených výzkumných otázek jsem zkoumala odlišnosti ve zkušenostech uživatelů sdíleného a tradičního modelu ubytování dle slovníku, který k popisu své zkušenosti používají; témat, která systematicky ve svých zkušenostech zmiňují; a sentimentů, které se v jejich zkušenostech objevují jak celkově, tak v souvislosti s jednotlivými tématy. Ve všech těchto oblastech se ukázalo, že Airbnb výrazně překonává tradiční ubytování v naplňování emočních a sociálních hodnot, zatímco uživatelé tradičních ubytování kladou důraz spíše na fyzické atributy jako vybavení. Lidský faktor (personál) nejen že zmiňují výrazně méně často, ale také spíše s negativní konotací, zatímco pro uživatele Airbnb je hostitel kategorií s druhým nejsilnějším pozitivním sentimentem. I přesto, že v Praze téměř všechna sdílená ubytování přes Airbnb nejsou skutečně fyzicky sdílena, ale jedná se o celé byty, nezdá se, že by tím vymizely benefity spojované se sdílenou ekonomikou obecně. Ačkoliv hosté fyzicky nepobývají s hostitelem v jednom společném prostoru, kontakt s ním je stále klíčovou dimenzí jejich zkušenosti a často ji dávají i do souvislosti se zážitky a emocemi z pobytu, což je téma, které se ve zkušenostech uživatelů tradičního ubytování ani neobjevuje. I přesto, že tedy většina ubytování přes Airbnb v Praze je pronajímána častěji než jen příležitostně, a roli tedy bude velmi pravděpodobně hrát i kapitalistická motivace generovat zisk, z hlediska spotřebitelské zkušenosti se ale tento trh stále významně liší od trhu tradičního. Tento výsledek je v souladu s tezí Botsman [2012], Sundararajana [2016], Masona [2016], Davida [2017] a dalších zastánců široké definice sdílené ekonomiky, že projekty jako Airbnb vytvářejí trh, kde mohou vznikat osobnější vztahy namísto neosobních transakcí trhu tradičního. A nemusí přitom mít podobu jen alternativy k byznysu, ale také alternativního byznysu [David 2017]. Rozchází se naopak s předpokladem, že pokud sdílené ubytování není skutečně fyzicky sdílené, pak se jedná o službu podobnou hotelu, pouze levnější [IPR 2018; Zervas et al. 2017; Úřad vlády 2017; Festila

a Müller 2017]. Minimálně v Praze se ukázalo, že z hlediska spotřebitelské zkušenosti tomu tak není.

Belarmino, Whalen, Koh a Bowen [2019], jako dosud jedni z mála, provedli přímou komparaci modelů ubytování s využitím slovních hodnocení z webových stránek poskytovatelů. Užívali nicméně pouze relativně malý vzorek osmi set recenzí ze čtyř amerických měst a analyzovali je pouze z hlediska distribučních frekvencí užívaných výrazů. Jejich výsledky jsou ale velmi podobné dílčím výsledkům výzkumu této práce. Jako zásadní rozdíl identifikovali také, že uživatelé Airbnb věnují největší pozornost vztahu s hostitelem, kdežto uživatelé tradičních ubytování popisují spíše fyzické vlastnosti pokoje. Zaznamenali také, že hostitel je často uživateli označován jeho křestním jménem, zatímco v souboru hodnocení tradičních ubytování se křestní jména objevují výrazně méně často a odkazují spíše ke konkrétním pozicím (číšník, recepční atd.). I jejich další poznatky se shodují s těmi této práce. Jako další odlišující kategorii totiž rovněž shledali u tradičních ubytování stravování a u sdíleného ubytování důraz na sousedství domu. Další studie s komparativní povahou je pak od autorů Mody, Sues a Lehto [2017], kteří provedli dotazníkové šetření se 630 respondenty a ve svém výzkumu také potvrzují, že Airbnb překonává hotely ve všech dimenzích spojených se zážitky. Nemateriální hodnoty ve zkušenostech uživatelů Airbnb identifikovali jako důležité i další autoři využívající tradiční metody sběru dat [Festila a Müller 2017; Guttentag a Smith 2017; Tussyadiah a Pesonen 2016; Tussyadiah 2016; Yannopoulou 2013]. Analyzování recenzí v rozsahu tzv. velkých dat, jako je tomu v této práci, je pak záležitostí převážně až přibližně posledního roku a amerického prostředí. Recenze Airbnb za různá města s využitím technik pro modelování témat analyzovali například Tang a McNicholas [2019], Cheng a Jin [2019], Zhang [2019] nebo Luo a Tang [2019]. Ačkoliv každý identifikoval jiný počet témat s jinou pravděpodobností výskytu⁷⁴, na důležitosti hostitele a komunikace s ním se shodují. Opakovaně se objevují témata lokality, doporučení, zážitků a vybavení, což je v souladu i s mými výsledky. Mezi témata detekovanými v mém výzkumu jsou navíc příjezd a odjezd, dopravní dostupnost a sousedství, které ale Zhang [2019] detekoval také, i když se liší pojmenováním a některá má rozdělená podrobněji do vícero témat. Luo a Tang [2019], Tang a McNicholas [2019] a Cheng a Jin [2019] modelování témat pak také stejně jako v této práci doplnili navíc o analýzu sentimentu. I v jejich souboru recenzí se objevuje negativní sentiment naprosto minimálně a výrazně převažuje ten pozitivní. Zároveň se i v jejich souborech ukazuje nejsilnější pozitivní

⁷⁴ Důvodem bude patrně explorační povaha přístupu, kde řada rozhodnutí leží na výzkumníkovi. Zároveň softwaru pro zpracování těchto dat nabízejí celou řadu možností a technik, které mohou přinášet lehce se odlišující výsledky.

sentiment u témat doporučení a hostitele. Žádná z těchto studií nicméně nemá komparativní povahu a recenze tradičních ubytovacích zařízení již neanalyzovali. Výsledky za tradiční ubytování tedy mohu porovnat jedině opět zvlášť se studiemi zaměřující se konkrétně jen na zkušenosti s tradičním ubytováním. Výsledky mého výzkumu se od nich také nijak výrazně neliší. Jako klíčové dimenze identifikují služby, lokalitu, pokoj, vybavení, cenu a stravu, případně personál [Akhtar et al. 2017; Liu et al. 2013; Crnojevac et al. 2010; Dolnicar a Otter 2003]. Hodnoty spojené se zážitky, emocemi či lidským kontaktem se mezi nimi také neobjevují. Akhtar [et al. 2017, str. 570] navíc provedl analýzu sentimentu na souboru recenzí tradičních ubytování. I v jeho souboru převládal sentiment pozitivní, i když v menší převaze než v tom mém, nicméně také u něj se ukázalo téma personál jako nejvíce negativní a celkově málo zastoupené oproti tématům jiným. Dle výsledků těchto studií se dá tedy říci, že nehledě na geografický kontext, uživatelé Airbnb se jeví jako výrazně více orientovaní na kvality ubytování nemateriální povahy, zatímco uživatelé tradičních modelů ubytování na kvality materiální.

Ačkoliv se postup, který jsem zvolila, ukázal jako efektivní a skutečně odhalil relevantní rozdíly ve zkušenostech uživatelů, má své limity. I když kvantifikování informací z textů není novinkou, nové jsou rozhodně postupy pro provádění takových analýz na natolik rozsáhlých datových souborech, jaké jsou užívány v této práci a které není reálné zpracovávat ručně. Algoritmy se vyvíjí neustále a metod je celá řada, jakkoliv sofistikované ale jsou, stále se jedná o strojové automatizované zpracovávání, je tedy důležité řešit otázku validity. Neexistuje například zatím spolehlivý postup, jak by počítač rozpoznal sarkasmus, humor či si poradil se slovem, které může mít v různých kontextech jiný význam. Přesto ale jsou obecně považovány výsledky takových postupů za neméně platné než při kódování kóděrem, ani lidský čtenář není bezchybný a každá interpretace textu je v určitém smyslu jeho zploštění. V souvislosti s dolováním informací z digitalizovaných dat z webů, jako jsou právě recenze, je relevantní také otázka reliability [Weare a Lin 2000, str. 287]. Weby mají nestálou povahu, online obsah se neustále mění. Seznam ubytování, která byla aktivní v době sběru dat, či jejich URL se může lišit i den ode dne. Alespoň u Bookingu jsem toto mohla kontrolovat podle seznamu hromadných ubytovacích zařízení ČSÚ a ukázalo se, že ze všech ubytování v Praze jich bylo v době sběru dat neaktivních pouze jedno procento, což je poměrně uspokojivé. Tuto proměnlivost dat jsem se tedy podle doporučení Weare a Lina [2000, str. 287] snažila řešit co nejrychlejším shromážděním dat, aby byla co nejporovnatelnější. Data jsem následně zpracovávala v akademickém prostředí běžně používanými postupy, které jsou považovány za

spolehlivé. Omezením výsledků je ale také nemožná generalizace. Pro uchopitelnost byl cíl omezen geograficky na Prahu, jazykově na zkušenosti s ubytováními psané v angličtině a časově na poslední dva roky. Zároveň při užití recenzí je také nemožné zjistit návratnost dotazníků, které uživatelé vyplňovali. Spíše než výsledky obecně platné tak výzkum přináší poměrně detailní vhled do toho, co se dostává uživatelům ubytování přes Airbnb v Praze navíc oproti ubytování v tradičních ubytovacích kapacitách opět v Praze. Vhodná by nicméně nepochybně byla ještě validizace výsledků provedením rozhovorů s danými uživateli o obsahu jejich recenzí a přístupu k psaní recenzí, zůstává totiž stále otevřená otázka, nakolik se inspiroují recenzemi ostatních a nakolik má vliv osobnější vztah s hostitelem přes Airbnb než s majitelem hotelu na to, jak upřímní ve svých recenzích uživatelé jsou.

Výsledky této práce vzbuzují několik podnětů pro výzkum budoucí. Nabízí se prozkoumání dalších jazyků recenzí a srovnání s těmi v angličtině, analyzovanými v této práci. Dále by, domnívám se, mohlo být přínosné provést segmentaci uživatelů, tato práce s nimi pracuje jako s homogenní skupinou sdíleného a tradičního ubytování, což tak ale jistě nemusí být. Zajímavé také je, že hosté Airbnb ve svých zkušenostech hovoří o vztahu s hostitelem, o zážitcích a emocích, přičemž ale obsazenost ubytování přes Airbnb v Praze je poměrně vysoká, pouze příležitostných ubytování je jen menšina. Zajímavá by tak mohla být kvalitativní studie, jak se hostitelé vyrovnávají s napětím mezi nutností standardizace své činnosti za účelem efektivity, ale zároveň stále poskytovat unikátní zkušenost každému hostu.

Hlavní přínos své práce tak vidím především v přímé srovnatelnosti výsledků analýzy zkušeností se sdílenými a tradičními modely ubytování. Ačkoliv byly v minulosti recenze těchto typů ubytování analyzovány zvlášť, každý z autorů volil trochu jiný postup a soubor za jiné časové období, z jiné lokality, jejich výsledky tak nejsou přímo komparovatelné. V českém prostředí navíc, ačkoliv je fenomén Airbnb velmi živě řešenou problematikou, zůstávají data v podobě recenzí nezpracovaná a pozornost je zaměřena spíše na otázku regulací [Ključnikov et al. 2018; Makovský 2017] než příčin úspěchu Airbnb a stagnace tradičních ubytování. Metodologicky práci považuji za přínosnou díky využití tzv. velkých dat (celkem téměř půl milionu recenzí) a různých typů analýz (distribuční frekvence, modelování témat, sémantická analýza), které z odlišných perspektiv ukazují, že to, co Airbnb v segmentu ubytování přináší navíc oproti stávajícím modelům, jsou hodnoty nemateriální, dříve spojované výhradně se soukromým životem a vztahy. Společensky rostoucí poptávka po nemateriálních hodnotách i

na trhu tak může patrně souviset nejen s popularitou Airbnb, ale také stagnací tradičních modelů ubytování, které na tyto nové potřeby, respektive hodnoty, nedostatečně reagují.

Závěr

Hlavním zjištěním této práce je, že spotřebitelská zkušenost s modely sdílené a tradiční ekonomiky v segmentu ubytování v Praze se zásadně liší především v míře naplnění postmateriálních hodnot. Zatímco uživatelé sdíleného ubytování přes Airbnb věnovali zásadní důraz při popisu svých zkušeností na prožitky a mezilidský kontakt, tyto hodnoty se ve zkušenostech uživatelů tradičního ubytování přes Booking objevovaly jen minimálně, a to navíc s negativní konotací, tedy s nimi uživatelé nebyli spokojeni. Naopak u Airbnb jsou to hodnoty zmiňované v nejvíce pozitivním smyslu.

K prozkoumání odlišností spotřebitelských zkušeností jsem využila slovní recenze ubytování v Praze ze dvou zmíněných webů, a to v angličtině a za poslední dva roky. Analyzovala jsem tedy data v podobě tzv. velkých dat v řádu statisíců. Ke zpracování jsem využila techniky těžení z textů, a sice strukturní modelování témat a analýzu sentimentu. Teoretickým východiskem mi byly teorie pojímající úspěch sdílené ekonomiky jako vývoj sociální a hodnotový, nikoliv pouze technologický [Botsman 2018; Hyman 2018; David 2017; Sundararajan 2016; Mason 2016; Zuboff 2010]. S tímto předpokladem jsem se tedy pokusila o detekování hodnot, na které uživatelé Airbnb a tradičních ubytovacích kapacit kladou důraz a s jakými sentimenty se v jejich zkušenostech pojí, s cílem přispět k diskusi nad tím, proč podíl sdílené ekonomiky na trhu s ubytováním v Praze roste, zatímco tradiční modely spíše stagnují.

Ve zkušenostech uživatelů obou typů ubytování jsem s využitím metody *structural topic modeling* detekovala osm systematicky se opakujících témat. U Airbnb jsou to sestupně dle důležitosti Hostitel a jeho péče o hosta, Celkový dojem z ubytování, Celý dům a jeho sousedství, Dopravní dostupnost, Zážitky a emoce z pobytu, Pokoj a jeho vybavení, Lokalita ubytování a Příjezd a odjezd. Na tyto výsledky jsem navázala analýzou sentimentu, která ukázala, že nejpozitivněji uživatelé Airbnb hovoří o Celkovém dojmu, Zážitcích a emocích, Domu a sousedství a Hostiteli a jeho péči. Jediná témata, ve kterých se objevuje i negativní sentiment jsou Příjezd a odjezd a Pokoj a jeho vybavení. Z těchto výsledků tedy vyplývá, že nejen, že při popisu svých zkušeností kladou velký důraz na hodnoty nemateriální a věnují jim největší prostor, ale zároveň je také vnímají nejpozitivněji. Nemateriální sociální a emoční hodnota jsou tak evidentně pro uživatele Airbnb nejen důležité, ale také se daří je uspokojovat, a to lépe než ty materiální. Zkušenost uživatelů tradičních ubytování je v tomto ohledu téměř opačná. V jejich recenzích jsem při použití naprosto totožného postupu detekovala sestupně dle

důležitosti témata Celkový dojem z ubytování, Pokoj a jeho vybavení, (Dis)komfort na pokoji, Lokalita, Personál a jeho služby, Dopravní dostupnost, Stravování a Cena vs hodnota. Důležitost pokoje, jeho vybavení a komfortu si tedy u uživatelů tradičních ubytování stojí velmi vysoko a věnují mu dokonce čtvrtinu svého popisu celkové zkušenosti. Uživatelé Airbnb věnují tématu pokoje jen necelou desetinu prostoru svých recenzí. Naopak lidskému faktoru (personálu) uživatelé tradičních ubytování věnují o polovinu méně prostoru než uživatelé Airbnb (hostiteli). Obdobně téměř opačné výsledky oproti zkušenostem uživatelů Airbnb se ukázaly také při analyzování sentimentů. Nejpozitivněji uživatelé tradičních ubytování hovoří o Stravování, Pokoji a jeho vybavení, (Dis)komfortu na pokoji a Celkovém dojmu. S materiálními hodnotami tedy bývají zpravidla spokojeni. Aspekt, se kterým jsou naopak nejméně spokojeni, je Personál a jeho služby, tedy sociální hodnota, a také Cena vs hodnota. Emoční hodnota se pak v jejich zkušenostech systematicky neobjevuje, tedy netvoří samostatné téma jako u Airbnb. Při mikro pohledu do souboru recenzí z hlediska frekvenčních distribucí jednotlivých slov se také ukázalo, že i konkrétní výrazy, které soubory nejvíce odlišují, jsou spojeny s postmateriálními hodnotami na straně Airbnb (týkají se komunikace, pocitů, vztahu k hostiteli) a hodnotami materiálními na straně tradičních ubytování (týkají se vybavení, jídla, servisu). Ukázaly se ale ještě dva zajímavé rozdíly. Jednak že křestní jména se ve zkušenostech s Airbnb objevují více než desetkrát častěji a po sloučení jsou dokonce druhým nejčastěji užívaným slovem vůbec. To opět demonstruje mnohem osobnější povahu transakcí v sektoru sdíleného ubytování oproti tradičnímu. A pak se také, že uživatelé Airbnb užívají pojem „Airbnb“ jako koncept celého ubytování, namísto slova „byt“ či „pokoj“, nikoliv pouze ve významu pouhého zprostředkovatele, jak je tomu u Bookingu.

Zkušenosti s těmito dvěma typy ubytování mají ale i značný společný základ. Oba typy uživatelů věnují velký prostor zhodnocení celkového dojmu, je pro ně důležitá dopravní dostupnost a lokalita. U společného tématu pokoje pak oba uživatelé velmi často hovoří o čistotě, což je dokonce sedmé nejužívanější slovo vůbec v obou souborech. Z hlediska sentimentů jsou pak oba soubory výrazně více pozitivní než negativní. Jen mírně negativněji hodnotí své zkušenosti uživatelé tradičních ubytovacích zařízení. Při pohledu na konkrétní emoce v obou souborech převládá důvěra, radost, a očekávání, přičemž emoce důvěry byla jen mírně častěji detekována ve zkušenostech uživatelů sdíleného ubytování.

I přesto tedy, že většinu nabídky Airbnb v Praze představují samostatné byty, kde nedochází k fyzickému sdílení prostoru s hostitelem, z hlediska spotřebitelské zkušenosti se

zdá, že Airbnb naplňuje předpoklad sdílené ekonomiky vytvářet trh o transakcích osobnějším než na trhu tradičním. Práce tak přináší novou perspektivu na úspěšnost Airbnb v Praze, která může souviset se společensky rostoucí poptávkou po postmateriálních hodnotách na trhu, které modely sdíleného ubytování přes Airbnb evidentně uspokojují lépe než modely tradičních ubytování (hotely, hostely, penziony) přes Booking. A ačkoliv výsledky analýzy vzhledem k povaze dat nelze generalizovat, nabízí poměrně detailní a v českém kontextu unikátní vhled do dané problematiky z hlediska spotřebitelské zkušenosti.

Summary

The major result is that the consumer experience with the models of sharing and traditional economy in the segment of accommodation in Prague differs fundamentally in the degree of post-material values fulfillment. While users of sharing economy accommodation via Airbnb paid great attention to emotions and interpersonal contact in describing their experiences, these values appeared only minimally in the experience of traditional accommodation users via Booking, moreover with a negative connotation. Conversely, for Airbnb, these are the values mentioned in the most positive sense. To explore the differences in the consumer experience, I used online reviews of accommodation in Prague from the two mentioned websites from the last two years. To process the data, I used text mining techniques, such as topic modelling and sentiment analysis. The thesis is based on theories conceptualizing sharing economy as a result of social and value development, not only as a technological one [Botsman 2018; Hyman 2018; David 2017; Sundararajan 2016; Mason 2016; Zuboff 2010]. With this assumption, I tried to detect the values that Airbnb users and traditional accommodation capacities emphasize in their experiences and what sentiments are associated with them, in order to contribute to the discussion of why the share of the sharing economy in the Prague housing market had been growing, while traditional models rather stagnated.

Despite the fact that most of Airbnb accommodations in Prague are whole apartments with no physical sharing of the space with the host, in terms of consumer experience, Airbnb seems to fulfill the premise of a sharing economy to create a market of more personal transactions in compare to the traditional market. The thesis thus brings a new perspective on the success of Airbnb in Prague, which may be related to the socially growing demand for post-material values in the market, which models of shared accommodation via Airbnb seem to satisfy better than traditional accommodation models (hotels, hostels, guesthouses) via Booking. And although the results of the analysis cannot be generalized, it offers a quite unique insight in the Czech context.

Použité zdroje

- Ady, M., D. Guadri-Felitti. 2015. *Consumer Research Identifies How To Present Travel Review Content For More Bookings*. USA: TrustYou Inc.
- Airbnb, Inc. *Airbnb Newsroom* [online]. Airbnb Newsroom [cit. 18.1.2020]. Dostupné z: <https://news.airbnb.com/>.
- Airbnb. *Zásady hodnocení Airbnb* [online]. Airbnb [cit. 18.1.2020]. Dostupné z: <https://www.airbnb.cz/help/article/2673/z%C3%A1sady-hodnocen%C3%AD-airbnb>.
- Airbnb. *Zmírňování sporů Airbnb s ohledem na hodnocení* [online]. Airbnb [cit. 18.1.2020]. Dostupné z: <https://www.airbnb.cz/help/article/548/zm%C3%ADr%C5%88ov%C3%A1n%C3%AD-spor%C5%AF-airbnb-s%C2%A0ohledem-na-hodnocen%C3%AD>.
- Akbart, Y. H., A. Tracogna. 2018. „The sharing economy and the future of the hotel industry: Transaction cost theory and platform economics.“ *International Journal of Hospitality Management* 71: 91-101.
- Akhmedova, A., F. Marimon, M. Mas-Machuca. 2020. „Winning strategies for customer loyalty in the sharing economy: A mixed-methods study.“ *Journal of Business Research* 112: 33-44.
- Akhtar, N., N. Zubair, A. Kumar, T. Ahmad. 2017. „Aspect based Sentiment Oriented Summarization of Hotel Reviews.“ *Procedia Computer Science* 115: 563-571.
- BAAYEN, R. H. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. New York: Cambridge University Press.
- Bathea, V., B. L. Nguyen, N. G. Nguyen, S. Ch. Satapathy, D. Le. 2017. *Information Systems Design and Intelligent Applications: Proceedings of Fourth International Conference India 2017*. Singapore: Sringer.
- Beck, U. 2004. *Riziková společnost: na cestě k jiné modernitě*. Praha: SLON.
- Belarmino, A., E. Whalen, Y. Koh, J. T. Bowen. 2019. „Comparing guests' key attributes of peer-to-peer accommodations and hotels: mixed-methods approach.“ *Current Issues in Tourism* 22:1: 1-7.
- Bell, D. 1976. „Welcome to the post-industrial society“. *Physics Today* 29 (2): 46-49.
- Bell, D. 1976. *The Coming of Post-Industrial Society: A Venture in Social Forecasting*. New York: Basic Books.
- Benoit, S., T. Baker, R. N. Bolton, T. Gruber, J. Kandampully. 2017. „A triadic framework

for collaborative consumption (CC): Motives, activities and resources & capabilities of actors.“ *Journal of Business Research* 79: 219-227.

- Berelson, B. 1952. *Content analysis in communication research*. New York: Free Press.
- Bergh, A., A. Funcke, J. Wernberg. 2018. *Timbro Sharing Economy Index*. Stockholm: Timbro City.
- Blal, I., S. Manisha, J. Templin. 2018. „Airbnb's effect on hotel sales growth.“ *International Journal of Hospitality Management* 73: 85-92.
- Blei, D. M., A. Y. NG, M. I. JORDAN. 2003.„Latent Dirichlet Allocation.“ *Journal of Machine Learning Research*. 3: 993-1022.
- Boltanski, L., É. Chiapello. 2007. *The New Spirit of Capitalism*. New York: Verso.
- Booking. 2020. *Mohu požádat o odstranění online hodnocení?* [online]. Booking.com [cit. 18.1.2020]. Dostupné z: <https://partner.booking.com/cs/pomoc/hodnocen%C3%AD-od-host%C5%AF/mohu-po%C5%BE%C3%A1dat-o-odstran%C4%9Bn%C3%AD-hodnocen%C3%AD>.
- Botsman, R. 2010. *The case for Collaborative Consumption* [video]. Ted.com [cit. 18.1.2020]. Dostupné z: https://www.ted.com/talks/rachel_botsman_the_case_for_collaborative_consumption.
- Botsman, R. 2012. *The currency of the new economy is trust* [video]. Ted.com [cit. 18.1.2020]. Dostupné z: https://www.ted.com/talks/rachel_botsman_the_currency_of_the_new_economy_is_trust/transcript.
- Botsman, R. 2018. *Who Can You Trust?: How Technology Brought Us Together – and Why It Could Drive Us Apart*. UK: Pinguin.
- Botsman, R., R. Rogers. 2010. *What's mine is yours: The rise of collaborative consumption*. New York: HarperBusiness.
- Brown, M. 2014. *Airbnb: The Growth Story You Didn't Know* [online]. Growthhackers [cit. 18.1.2020]. Dostupné z: <https://growthhackers.com/growth-studies/airbnb>.
- Bruno, B., M. Faggini. 2020. „Sharing Competition: An Agent-Based Model for the Short-Term Accommodations Market.“ *The B.E. Journal of Economic Analysis & Policy* 20(2).
- Castells, M. 1996. *The Rise of Networked Society*. USA: Wiley-Blackwell.

- Crnojevac, I. H., J. Gugić, S. Karlovčan. 2010. „eTourism: a comparison of online and offline bookings and the importance of hotel attributes.“ *The International Journal of Artificial Organs* 34(1).
- Česko v datech. 2018. *Budoucnost ve sdílení* [online]. Česko v datech [cit. 17.2.2020]. Dostupné z: <https://www.ceskovdatech.cz/clanek/92-budoucnost-ve-sdileni-velky-prehled-toho-co-uz-v-cesku-sdilime/>.
- Český statistický úřad (ČSÚ). 2019. Hromadná ubytovací zařízení České republiky [online]. ČSÚ [vid. 30.12.2019]. Dostupné z: <https://vdb.czso.cz/huz/okhuz.jsp?k=CZ0100>.
- Český statistický úřad. 2019. *Hromadná ubytovací zařízení České republiky* [online]. *czso.cz* [cit. 18.1.2020]. Dostupné z: <https://vdb.czso.cz/huz/okhuz.jsp?k=CZ0100>.
- ČT24. 2020. *Lidé, kteří pronajímají byty turistům, to musí hlásit, shodují se politici. Časové omezení čelí kritice* [online]. ČT24 [vid. 15.4.2020]. Dostupné z: <https://ct24.ceskatelevize.cz/regiony/3046813-lide-kteri-pronajimaji-byty-turistum-musi-hlasit-shoduji-se-politici-casove-omezeni>.
- David, M. 2017. *Sharing: Crime Against Capitalism*. UK: Polity Press.
- De Vaus, D. A. 2002. *Surveys in Social Research*. London: Psychology Press.
- Deloitte. 2017. *Sdílená ekonomika. Bohatství bez vlastnictví*. Praha: Deloitte.
- Deloitte. 2019. *Prague Hospitality Report. Tourism, Hotels & P2P accommodation*. Praha: Deloitte.
- Díaz, M., I. L. Johnson, A. Piper, D. Gergle. 2018. „Addressing Age-Related Bias in Sentiment Analysis.“ *CHI '18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 412: 1-14.
- Ditz, G. W. 2014. „Protestantská etika a tržní hospodářství“ Pp. 7-50 in T. Ježek (eds.). *Liberální ekonomie*, Praha: Prostor.
- Dobusch, L. 2019. „Dynamics of the Sharing Economy between Commons and Commodification.“ *Zeitschrift für Sozialen Fortschritt* 8(2): 109-115.
- Dogru, T., M. Mody, C. Suess. S. McGinley, N. D. Line. 2020. „The Airbnb paradox: Positive employment effects in the hospitality industry.“ *Tourism Management* 77: 1-13.
- Dolnicar, S., T. Otter. 2003. Which hotel attributes matter? A review of previous and a framework for future research. *Proceedings of the 9th Annual Conference of the Asia Pacific Tourism Association (APTA) 1. University of Technology, Sydney*: 176–188.

- Downer, K., Ch. Wells, Ch. Crichton. 2019. „All Work and no Play: A Text Analysis.“ *International Journal of Market Research* 61(3).
- Duan, W., B. Gu, A. B. Whinston. 2008. „Do online reviews matter? An empirical investigation of panel data.“ *Decision Support System* 45(4): 1007–1016.
- Festila, M., S. Müller. 2017. „The impact of technology-Mediated consumption on identity: the case of airbnb.“ *Paper Presented at the Proceedings of the 50th Hawaii International Conference on System Sciences.s*
- Frenken, K., J. Schor. 2017. „Putting the sharing economy into perspective.“ *Environmental Innovation and Societal Transitions* 23: 3–10.
- Friedman, T. L. 2013. *Welcome to the 'Sharing Economy'* [online]. Nytimes.com [vid. 18.1.2020]. Dostupné z: <http://www.nytimes.com/2013/07/21/opinion/sunday/friedman-welcome-to-the-sharing-economy.html>.
- Ginterová, M. 2018. *Cestování kamkoliv dnes začíná na „sítích“*. *A kdo nemá výborné hodnocení, nemá šanci* [online]. ČT24 [vid. 15.4.2020]. Dostupné z: <https://ct24.ceskatelevize.cz/ekonomika/2459464-cestovani-kamkoliv-dnes-zacina-na-sitich-a-kdo-nema-vyborne-hodnoceni-nema-sanci>.
- Girolami, M., Kabán, A. 2003. „On an equivalence between PLSI and LDA.“ *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*: 433-434.
- Gurran, N., G. Searle, P. Phibbs. 2018. „Urban Planning in the Age of Airbnb: Coase, Property Rights, and Spatial Regulation.“ *Urban Policy and Research* 36(4): 399-416.
- Guttentag, D., S. Smith. 2017. „Assessing Airbnb as a disruptive innovation relative to hotels: substitution and comparative performance expectations.“ *International Journal of Hospitality Management* 64: 1–10.
- Guttentag, Daniel. 2015. „Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector“. *Current issues in Tourism* 18(12): 1192-1217.
- Hájek, M. 2014. *Čtenář a stroj. Vybrané metody sociálněvědní analýzy textů*. Praha: SLON.
- Harris, Brett. 2017. „Uber, Lyft, and Regulating the Sharing Economy.“ *Seattle University Law Review* 41(1): 269–86.
- He, W., S. Zha, L. Li. 2013. „Social media competitive analysis and text mining: A case study in the pizza industry.“ *International Journal of Information Management* 33: 464-472.

- Hochschild, A. 1983. *The Managed Heart: Commercialization of Human Feeling*. Berkeley: The University of California Press.
- Horn, K., M. Merante. 2017. „Is Home Sharing Driving up Rents? Evidence from Airbnb in Boston.“ *Journal of Housing Economics* 38: 14-24.
- Hyman, L. 2018. *Temp: How American Work, American Business, and the American Dream Became Temporary*. NY: Viking.
- Chen, Ch., Y. Chang. 2018. „What drives purchase intention on Airbnb? Perspectives of consumer reviews, information quality, and media richness.“ *Telematics and Informatics* 35: 1512-1523.
- Cheng, M., X. Jin. 2019. „What do Airbnb users care about? An analysis of online review comments.“ *International Journal of Hospitality Management* 76: 58–70.
- Chesky, B. 2014. *The Story No One Believed – Brian Chesky, CoFounder and CEO AirBnB* [online]. Youtube [cit. 18.1.2020]. Dostupné z: <https://www.youtube.com/watch?v=1bAT44QPPHw>.
- Inglehart, R. 2008. „Changing Values Among Western Publics from 1970 to 2006.“ *West European Politics* 31(1): 130-146.
- Insideairbnb. *Inside Airbnb. Adding Data to the Debate* [online]. Insideairbnb.com [vid. 18.1.2020]. Dostupné z: <http://insideairbnb.com/prague/>.
- Institut pro plánování a rozvoj města (IPR). 2018. *Fenomén Airbnb a jeho dopady v kontextu hl. m. Prahy*. Praha: IPR Praha.
- Ioannides, D., M. Röslmaier, E. Van Der Zee. 2019. „Airbnb as an instigator of “tourism bubble” expansion in Utrecht’s Lombok neighbourhood.“ *Tourism Geographies* 21(5): 822-840.
- Karlgren, J., M. Sahlgren, F. Olsson, F. Espinoza, O. Hamfors. 2012. „Usefulness of Sentiment Analysis.“ Conference: *Proceedings of the 34th European conference on Advances in Information Retrieval*.
- Ključnikov, A., V. Krajčík, Z. Vincúrová. 2018. „International Sharing Economy: The Case of Airbnb in the Czech Republic“. *Economics & Sociology* 11(2): 126- 137.
- Kozinets, R. V. 2010. *Netnography: Doing Ethnographic Research Online*. London: Sage.
- Kronick, J. C. 1997. „Alternativní metodologie pro analýzu kvalitativních dat.“ *Sociologický časopis* 33(1): 57-67.

- Krotov, V., L. Silva. 2018. Legality and Ethics of Web Scraping. *Proc. of 24th Americas Conference on Information Systems New Orleans*.
- Lampinen, A., C. Cheshire. 2016. „Hosting via Airbnb: motivations and financial assurances in monetized network hospitality.“ *Paper Presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*.
- Lawani, A., M. R. Reed, T. Mark, Y. Zheng. 2020. „Reviews and Price on Online Platforms: Evidence from Sentiment Analysis of Airbnb Reviews in Boston,“ *Regional Science And Urban Economics* 75: 22-34.
- Lee, E. J., S. Y. Shin. 2014. „When Do Consumers Buy Online Product Reviews? Effects of Review Quality, Product Type, and Reviewer’s photo.“ *Computers in Human Behavior* 31: 356-366.
- Liu, S. Q., A. S. Mattila. 2017. Airbnb: Online targeted advertising, sense of power, and consumer decisions. *International Journal of Hospitality Management* 60: 33–41.
- Liu, S., R. Law, J. Rong, G. Li, J. Hall. 2013. „Analyzing changes in hotel customers’ expectations by trip mode.” *International Journal of Hospitality Management* 34: 359–371.
- Liu, Y., T. Teichert, M. Rossi, H. Li, F. Hu. 2017. „Big data for big insights: investigating language-specific drivers of hotel satisfaction with 412,784 user-generated reviews.“ *Tourism Management* 59: 554–563.
- Luhmann, N., S. Fuchs. 1988. „Tautology and paradox in the self-descriptions of modern society.“ *Sociological Theory* 6(1): 21-37.
- Luo, Y., Tang, R. L. 2019. „Understanding hidden dimensions in textual reviews on Airbnb: An application of modified latent aspect rating analysis (LARA).” *International Journal of Hospitality Management* 80: 144–154.
- Makovský, P. 2017. „The Shared Economy in the Czech Republic in 2017 and Resulting Problems in Short-Term Housing Rentals“. *Economic Studies & Analyses/Acta VSFS* 11(2).
- Mason, P. 2016. *Postcapitalism: A Guide to Our Future*. UK: Pinguin.
- McDonald, D., U. Kelly. 2012. *The Value and Benefits of Text Mining*. UK: JISC.
- Mimno, D., H. M., Wallach, E. Talley, M. Leenders, A. McCallum. 2011. „Optimizing Semantic Coherence in Topic Models.” *Proceedings of the Conference on Empirical Methods in Natural Language Processing EMNLP ’11*: 262–272.

- Mody, M. A., C. Suess, X. Lehto. 2017. „The accommodation experiencescape: a comparative assessment of hotels and Airbnb.” *International Journal of Contemporary Hospitality Management* 29(9): 2377–2404.
- Mohammad, S. M., P. D. Turney. 2010. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. *Proceedings of the NAACL-HLT 2010 workshop on computational approaches to analysis and generation of emotion in Text, LA, California*.
- Pang, B., L. Lee. 2006. „Opinion Mining and Sentiment Analysis.” *Foundations and Trends in Information Retrieval* 1(2): 91-231.
- Parvez, M. S., K. S. A. Tasneem, S. S. Rajendra, R. B. Bodke. 2018. „Analysis Of Different Web Data Extraction Techniques.” *2018 International Conference on Smart City and Emerging Technology, ICSCET*.
- Perikos, I., A. Tsirtsis, K. Kostas, F. Grivokostopoulou, I. Daramouskas, I. Hatzilygeroudis. 2018. „Opinion Mining and Visualization of Online Users Reviews: A Case Study in Booking.com.” *2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA), Information, Intelligence, Systems and Applications (IISA)*: 1–5.
- PwC. 2015. *The Sharing Economy: Consumer Intelligence Series*. Delaware: PricewaterhouseCoopers LLP.
- Reinsel, D., J. Gantz, J. Rydning. 2018. *Data Age 2025. The Digitization of the World. From Edge to Core*. USA: IDC.
- Roberts, M. E., B. M. Stewart, D. Tingley. 2014. „stm: An R Package for Structural Topic Models.” *Journal of Statistical Software* 91(2): 1-40.
- Saberi, B., S. Saad. 2017. „Sentiment Analysis or Opinion Mining: A Review.” *International Journal on Advanced Science Engineering Information Technology* 7(5): 1660-1666.
- Sans, A. A., A. Q. Domínguez. 2016. „Unravelling Airbnb: Urban perspectives from Barcelona.” Pp. 209-228 in A. P. Russo, G. Richards (eds.). *Reinventing the local in tourism: Producing, consuming and negotiating place*. Bristol: Channel View Publications.
- Schor, J. B., C. Fitzmaurice, L. Carfagna, C. Will-Attwood. 2016. „Paradoxes of openness and distinction in the sharing economy.” *Poetics* 54: 66-81.
- Silge, J. 2018a. *The game is afoot! Topic modeling of Sherlock Holmes stories* [online]. JuliaSilge.com [cit. 17.2.2020]. Dostupné z: <https://juliasilge.com/blog/sherlock-holmes-stm/>.

- Silge, J. 2018b. Training, evaluating, and interpreting topic models [online]. JuliaSilge.com [cit. 17.2.2020]. Dostupné z: <https://juliasilge.com/blog/evaluating-stm/>.
- Silge, J., D. Robinson. 2017. *Text mining with R: A tidy approach*. USA: O'Reilly Media.
- Sparks, B. A., V. Browning. 2011. „The impact of online reviews on hotel booking intentions and perception of trust.“ *Tourism management* 32(6): 1310-1323.
- Statista. 2020a. *Valuation of Uber and Airbnb from 2014 to 2018* [online]. Statista [vid. 15.4.2020]. Dostupné z: <https://www.statista.com/statistics/831044/increasing-valuation-of-uber-and-airbnb/>.
- Statista. 2020b. *Global Digital Population as of April 2020* [online]. Statista [vid. 15.4.2020]. Dostupné z: <https://www.statista.com/statistics/617136/digital-population-worldwide/>.
- Stephany, A. 2015. *The business of sharing: Making it in the new sharing economy*. New York: Springer.
- Sundararajan, A. 2016. *The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism*. Cambridge, Massachusetts: The MIT Press.
- Taboada, M., J. Brooke, M. Tofiloski, K. Voll, M. Stede. 2011. „Lexicon-Based Methods for Sentiment Analysis.“ *Computational Linguistics*. 37(2): 267-307.
- Taddy, M. A. 2012. „On Estimation and Selection for Topic Models.“ *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*.
- Tang, Y., P. D. McNicholas. 2019. „Clustering Airbnb Reviews.“ *Department of Mathematics & Statistics, McMaster University arXiv, Applications*: 1-27
- Teubner, T., F. Hawlitschek, D. Dann. 2017. „Price determinants on Airbnb: how reputation pays off in the sharing economy.“ *Journal of Self-Governance and Management Economics* 5(4): 55-80.
- Toffler, A., H. Toffler. 2001. *Třetí vlna*. Praha: Dokořán.
- Trajkovska, B. 2018. *Once it was a small Dutch startup, today it's the world's largest hotel booking platform – Interview with Gillian Tans, CEO of Booking.com* [online]. Eu-startups [cit. 17.2.2020]. Dostupné z: <https://www.eu-startups.com/2018/07/once-a-small-dutch-startup-today-the-worlds-largest-accommodations-platform-interview-with-gillian-tans-ceo-of-booking-com/>.
- Tussyadiah, I. P., F. Zach. 2017. „Identifying salient attributes of peer-to-peer accommodation experience.“ *Journal of Travel & Tourism Marketing* 34: 636-652.

- Tussyadiah, L., J. Pesonen. 2016. „Impacts of peer-to-Peer accommodation use on travel patterns.“ *Journal of Travel Research* 55(8): 1022–1040.
- Úřad vlády ČR. 2017. *Analýza sdílené ekonomiky a digitálních platforem* [online]. Sekce pro evropské záležitosti Úřadu vlády ČR [cit. 17. 2. 2020]. Dostupné z: https://www.vlada.cz/assets/urad-vlady/poskytovani-informaci/poskytnute-informace-na-zadost/Priloha_4_Material_Analyza.pdf.
- Valdivia, A., M. C. Luzon, F. Herrera. 2017. „Sentiment Analysis in TripAdvisor.“ *Intelligent Systems, IEEE* 32(4): 72-77.
- Vătämanescu, E. 2017. *Knowledge management in the sharing economy: cross-sectoral insights into the future of competitive advantage*. New York: Springer Berlin Heidelberg.
- Vinodhini, G. R. M. Chandrasekaran. 2012. „Sentiment Analysis and Opinion Mining: A Survey.“ *International Journal of Advanced Research in Computer Science and Software Engineering*: 2(6): 282-292.
- Volcani, Y., D. Fogel. System and method for determining and controlling the impact of text. USA. US20020199166A1. 28.6.2001.
- Wallach, H. M., I. Murray, R. Salakhutdinov, D. Mimno. 2009. „Evaluation Methods for Topic Models.“ *Proceedings of the 26th Annual International Conference on Machine Learning*: 1105–1112.
- Weare, Ch., W. Lin. 2000, „Content Analysis of the World Wide Web. Opportunities and Challenges.“ *Social Science Computer Review* 18(3): 272-292.
- Xiang, Z., Z. Schwartz, J. H. Gerdes, M. Uysal. 2015. „What can big data and text analytics tell us about hotel guestexperience and satisfaction?.“ *International Journal of Hospitality Management* 44: 120–130.
- Xie, K. L., K. K. F. So, W. Wang. 2017. „Joint effects of management responses and online reviews on hotel financial performance: a data-analytics approach.“ *International Journal of Hospitality Management* 62: 101-110.
- Xu, F., L. La, M. Hu, J. Wang, C. Huang. 2019. „The influence of neighbourhood environment on Airbnb: a geographically weighed regression analysis.“ *Tourism Geographies* 22(1): 192–209.
- Yannopoulou, N. 2013. „User-generated brands and social media: Couchsurfing and Airbnb.“ *Contemporary Management Research* 9(1): 85–90.
- Zervas, G., D. Proserpio, J. W. Byers. 2017. „The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry.“ *Journal of Marketing Research* 54(5): 687-705.

- Zhang, J. 2019. „Listening to the Consumer: Exploring Review Topics on Airbnb and Their Impact on Listing Performance.“ *Journal of Marketing Theory and Practice* 27(4): 371-389.
- Zhou, Q., Y. Chen, Ch. Ma, F. LI, Y. Xiao, X. Wang, X. Fu. 2018. „Measurement and Analysis of the Reviews in Airbnb.“ *2018 IFIP Networking Conference (IFIP Networking) and Workshops*.
- Zuboff, S. 2010. *Creating value in the age of distributed capitalism* [online]. McKinsey [cit. 2020-01-18]. Dostupné z: <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/creating-value-in-the-age-of-distributed-capitalism>.

Teze diplomové práce

ISS, FSV UK

Projekt diplomové práce

Vedoucí práce: PhDr. Mgr. Jan Balon Ph.D.

Autorka: Bc. Tereza Svobodová

Předpokládaný rok obhajoby: 2020

Předpokládaný název: Fenomén Airbnb v kontextu postmateriálních hodnot

Název v anglickém jazyce: Airbnb Phenomenon in the Context of Post-Materialist Values

Cíl práce:

Cílem diplomové práce je zasadit fenomén sdílené ekonomiky do sociologické a axiologické perspektivy, a sice na konkrétním případu platformy Airbnb, největšího zprostředkovatele sdíleného ubytování. Teoretickým východiskem pro konceptualizaci postmateriálních hodnot je především dílo Ingleharta a Boltanského. Prvním záměrem je zmapovat zákaznickou komunikaci Airbnb, ukázat a vysvětlit její cílenost na příslib naplnění postmateriálních hodnot. Následně bude ověřen předpoklad, že k němu skutečně v očích zákazníků dochází, tedy že oni sami takové hodnoty, jaké jsou platformou komunikovány, ve svých zkušenostech s Airbnb reflektují. K identifikaci hodnot, jež zákazníci vyzdvihují, bude využit korpus uživatelských recenzí, dostupných přímo na webu Airbnb, a na nich provedena obsahová analýza. Korpus ovšem bude doplněn ještě druhým podsouborem, a to recenzemi tradičních ubytovacích zařízení (hotely, hostely, pensiony) z portálu Booking. Komparace tak umožní ověřit, zdali postmateriální hodnoty Airbnb naplňuje lépe než tradiční ubytovací zařízení, což by poskytlo jedno z možných vysvětlení jeho výrazného růstu podílu na trhu s ubytováním.

Cíl práce v anglickém jazyce:

The aim of the thesis is to put the phenomenon of sharing economy into sociological and axiological perspective, using the case of Airbnb – the biggest provider of shared accommodation. The theoretical basis for the conceptualization of post-material values is primarily the work of Inglehart and Boltanski. The first intention is to map Airbnb's customer communication and to show and explain its targeting at the promise of fulfilling post-material values. Subsequently, there will be verified the assumption that customers actually reflect the values Airbnb communicates in their experience with the platform. To identify the values that are highlighted by customers, a corpus of user reviews, available directly on Airbnb, will be used and then content analysis will be performed. However, the corpus will be supplemented with an additional subset with reviews of traditional accommodation facilities (hotels, hostels, pensions) from the Booking portal. Thus, the comparison will enable the author to verify whether the post-material needs are more often met by Airbnb than by traditional accommodation facilities. It could provide one of possible explanations for its significant growth in the housing market.

Námět práce, uvedení tématu:

Fenomén Airbnb se díky své unikátnosti, stále většímu podílu na trhu a specifické formě marketingové komunikace stal předmětem odborných studií na poli celé řady oborů. Ačkoliv samotná platforma byla založena teprve v roce 2008, již o ní vznikly desítky, ne-li stovky, publikací a celé téma sdílené ekonomiky se zdá být každým rokem aktuálnější. Mezi hojně probádané oblasti ve vztahu k Airbnb patří marketing, odkud vzešlo mnoho prací o jeho offline i online komunikaci (Liu & Mattila, 2017; Burmann & Arnhold, 2009; Yannopoulou et al., 2013), dále právo či politika zaměřující se zejména na regulace Airbnb a legislativu (Wills, 2017; Biber, 2017) a především pak ekonomie, kde nejvíce se vyskytujícími tématy jsou dopady růstu Airbnb na tradiční ubytovací kapacity (Zervas et al., 2016; Guttentag 2015; Blal et al. 2018; Varma et al., 2016), cenové strategie (Wang & Nicolau, 2017) a dopady na pracovní sílu (Fang & Ye, 2016).

Tato diplomová práce zamýšlí doplnit studie o chybějící sociologickou a axiologickou perspektivu. Z publikací, jež jsou zaměřeny na spokojenost uživatelů Airbnb (Tussyadiah, 2016; Liang et al., 2018) či segmentaci uživatelů ubytovacích kapacit (Varma, 2016; Guttentag et al., 2017), se ukázalo, že Airbnb nabízí svým uživatelům hodnoty jako zážitky, emoce nebo pocit sounáležitosti, jež většinou souvisí s přítomností hostitele či jedinečností ubytovacího prostoru. Ve stávajících pracích bývají tyto hodnoty ovšem pouze zmíněny jako jeden z faktorů při výběru ubytování vedle výhodné ceny či lokality a není jim dále věnována pozornost. Nicméně běžné faktory jako je právě cena, lokalita či servis hrají roli již dlouhé roky i při rozhodování mezi jednotkami tradičních ubytování (Sohrabi et al., 2012; Knutson, 1988; Atkinson, 1988; Lockyer, 2005; Chiang-Ming et al., 2015), nejen při rozhodování mezi Airbnb a tradičními ubytováními. To, co s Airbnb přichází nově, jsou právě zmíněné hodnoty, jež dříve byly spojované pouze s osobním soukromým životem a vztahy a do oblasti trhu a v podstatě byznysu se dostaly poměrně nedávno a jejich vzrůstající popularita může patrně souviset i s popularitou Airbnb, pokud tam tedy skutečně je potenciál, že je dokáže naplnit lépe než tradiční ubytovací kapacity. Tyto nové hodnoty, jejich potenciál a realizace jsou předmětem této diplomové práce, jejíž teoretickými východisky jsou především dílo Boltanského (Boltanski & Chiapello, 2007) a Ingleharta (1997; Inglehart & Welzel, 2005). Inglehart, Boltanski, ale i další autoři (Toffler, 2001; Ditz, 2014; Hochschild, 1983), se shodují, že nově se na trhu stávají populárními takové hodnoty, které do té doby na úrovni anonymních vztahů v rámci trhu neměly místo. Jedná se o hodnoty lidské jako jsou emoce, důvěra, sebevyjádření, sounáležitost, autenticita a podobné, jež se staly klíčovými pro dnešního zákazníka, a tím pádem i základní konkurenční výhodou na trhu (Boltanski 2007, s. 464-466). Jsou to přesně takové hodnoty, na kterých Airbnb postavilo image své značky spojenou s logem „Belong Everywhere“. Tato komunikační strategie byla zahájena od roku 2014, údajně po rozhovorech s uživateli o tom, co oni sami chtějí a v tento rok se podařilo více než zdvojnásobit počet ubytování na evropském trhu (Brown, 2014). Lze tedy předpokládat, že Airbnb nabídlo hodnoty, které na trhu do té doby k dispozici nebyly a které jsou dnešní společností žádané. Projekt tedy nově zasazuje fenomén Airbnb do kontextu celospolečenských hodnotových proměn, jakožto jednoho z podstatných faktorů jeho růstu.

Metoda:

Pro analytickou část diplomové práce bude využito bohaté množství uživatelských recenzí, které jsou volně k dispozici na platformě Airbnb, v případě tradičních ubytování na platformě Booking, a na jejich korpusu provedena komparativní obsahová analýza. Tato metoda umožní zjistit, jaké hodnoty samotní uživatelé na těchto dvou typech ubytování vyzdvihují, zdali se nějak liší a jsou-li skutečně postkapitalistické hodnoty součástí zkušeností uživatelů Airbnb více než součástí zkušeností uživatelů tradičních ubytování. Vzorek bude omezen geograficky a časově.

Klíčová slova:

postkapitalistické hodnoty, postmateriální hodnoty, sdílená ekonomika, Airbnb, zákaznická zkušenost, zákaznická komunikace, obsahová analýza

Key words:

post-capitalist values, post-materialist values, sharing economy, Airbnb, customer experience, customer communication, content analysis

Předběžná struktura práce:

Úvod

1 Teoretická východiska**1.1 Celospolečenská proměna hodnot: od tradičních k postmateriálním****1.1.1 Postmateriální nebo postkapitalistické?****1.1.1.1 Pojetí Ingleharta****1.1.1.2 Pojetí Boltanského****1.1.1.3 Vlastní konceptualizace****1.2 Důvěra, zážitky a jedinečnost: současný vývoj segmentu zákaznické péče a komunikace****1.3 Sdílená ekonomika jako nový ekonomický model postkapitalismu**

2 Fenomén Airbnb

2.1 Vznik a úspěch Airbnb

2.2 Vize, mise a funkce Airbnb

2.3 Zákaznická komunikace Airbnb

3 Metodologie výzkumu

3.1 Obsahová analýza

3.2 Design výzkumu

4 Analytická část

4.1 Výsledky obsahové analýzy

Závěry

Zdroje

Přílohy

Orientační seznam literatury:

- ATKINSON, A. Answering the eternal question: What does the customer want. *Cornell Hotel and Restaurant Administration Quarterly*. 1988. Č. 29, str. 12–14.
- BIBER, Eric; Light, Sarah E.; RUHL, J. B.; SALZMAN, James. *Vanderbilt Law Review*. 2017. Vol. 70, č. 5, str. 1561-1626.
- BLAL, Inés; SINGAL, Manisha; TEMPLIN, Jonathan. Airbnb's effect on hotel sales growth. *International Journal of Hospitality Management*. 2018. Vol. 73, str. 85-92.
- BOLTANSKI, Luc a CHIAPELLO, Éve. 2005. *The New Spirit of Capitalism*. New York: Verso. ISBN: 1859845541.
- BOTSCHAN, Rachel. *Defining The Sharing Economy: What Is Collaborative Consumption—And What Isn't?* In: fastcoexist.com [online]. 2015 [vid. 10-11-2018]. Dostupné z: <http://www.fastcoexist.com/3046119/defining-the-sharing-economy-what-is-collaborative-consumption-and-what-isnt>.

- BOTSMAN, Rachel. 2017. Who can you trust?: How Technology Brought Us Together and Why It Might Drive Us Apart. USA: Public Affairs.
- BROWN, Morgan. Airbnb: The Growth Story You Didn't Know. In: growthhackers [ONLINE]. 2014 [vid. 10-11-2018]. Dostupné z: <https://growthhackers.com/growth-studies/airbnb>.
- BURMANN, Christoph; ARNHOLD, Ulrike. User generated branding: State of the art of research. Münster: LIK, 2008. ISBN: 382581887X.
- DITZ, Gerhard. Protestantská etika a tržní hospodářství. In: Liberální ekonomie. Praha: PROSTOR, 2014. ISBN 978-80-7260-303-9
- DUMBROVSKÁ, Veronika. Konferenční příspěvek: 2017 AAG Annual Meeting, April 5 - April 9, 2017, Boston. "*The evolution of urban tourism precincts: A case study of Prague's historic centre.*"
- DUMBROVSKÁ, Veronika. Konferenční příspěvek: Atlas Annual Conference 2017, 12th - 16th September 2017, Viana do Castelo, Portugal. "*AirBnB in Prague: Sharing economy or the new way of commercialization?*".
- DVOŘÁKOVÁ, Ilona. Obsahová analýza / formální obsahová analýza / kvantitativní obsahová analýza. In: Antropoweb [ONLINE]. 2010 [vid. 10-11-2018]. Dostupné z: http://www.antropoweb.cz/media/webzin/webzin_2_2010/Dvorakova_I-2-2010.pdf.
- EROSE, Sthapit; JANO, Jiménez-Barreto. Exploring tourists' memorable hospitality experiences: An Airbnb perspective. Tourism Management Perspectives. 2018. Vol. 28, str. 83-92.
- FANG, B.; YE, Q. Effect of sharing economy on tourism industry employment. Annuals of Tourism Research. 2016. Č. 57, str. 264.
- GUEST, Greg; ARWEN, Bunce; JOHNSON, Laura. How Many Interviews Are Enough?. *Field Methods*. 2006. Vol. 18., č. 1, str. 59-82.
- GUTTENTAG, D. Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*. 2015. Vol. 18, č. 12, str. 1192-1217.
- GUTTENTAG, Daniel; SMITH, Stephen; POTWARKA, Luke; HAVITZ, Mark. Why Tourists Choose Airbnb: A Motivation-Based Segmentation Study. *Journal of Travel Research*. 2017. Č. 57.
- HOCHSCHILD, Arlie. 1983. *The Managed Heart: Commercialization of Human Feeling*. Berkeley: The University of California Press. ISBN: 9780520272941.

- CHIANG-MING, Chen; YI-CHUN, Tsai; HSIEN-HUNG, Chiu. The decision-making process of and the decisive factors in accommodation choice. *Current Issues in Tourism*. 2015. Vol. 20, č. 2, str. 111-129.
- INGLEHART, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton, N.J.: Princeton University Press. ISBN: 9780691011806.
- INGLEHART, Ronald; WELZEL, Christian. 2005. *Modernization, cultural change, and democracy: The human development sequence*. UK: Cambridge University Press. ISBN: 9780511790881.
- INGLEHART, F. Ronald. 2018. *Cultural Evolution: People's Motivations are Changing, and Reshaping the World*. UK: Cambridge University Press.
- KLJUČNIKOV, Aleksandr; KRAJČÍK, Vladimír; VINCÚROVÁ; Zuzana. International sharing economy: The case of airbnb in the Czech Republic. *Economics and Sociology*. 2018. Vol. 11, č.2., str. 126-137.
- KNUTSON, B. J.. Frequent travelers: Making them happy and bringing them back. *Cornell Hotel and Restaurant Administration Quarterly*. 1988. Č. 29, str. 82–87.
- LIANG, Exploring the relationship between satisfaction, trust and switching intention, repurchase intention in the context of Airbnb. *International Journal of Hospitality Management*. 2018. Č. 69, str. 41-48.
- LIU, Stephanie Q.; Mattila, Anna S.. Airbnb: Online targeted advertising, sense of power, and consumer decisions. *International Journal of Hospitality Management*. 2017. vol. 60. str. 33-41.
- MAKOVSKÝ, Petr. The Shared Economy in the Czech Republic in 2017 and Resulting Problems in Short-Term Housing Rentals. *Economic Studies & Analyses / Acta VSFS* . 2017. Vol. 11, č. 2, str. 144-159.
- Mason, Paul. 2015. *Postcapitalism: A Guide to Our Future*. UK: Allen Lane. David, Matthew. 2017. *Sharing: Crime Against Capitalism*. UK: Polity.
- SCHWARZ, Josef. Současný stav a trendy automatické indexace dokumentů: přehledová studie. In: Ikaros [online]. 2003 [vid. 10-11-2018]. Dostupné také z: <http://ikaros.cz/soucasny-stav-a-trendy-automaticke-indexace-dokumentu>.
- SOHRABI, B.; VANANI, I. R.; TAHMASEBIPUR, K.; FAZLI, S. An exploratory analysis of hotel selection factors: A comprehensive survey of Tehran hotels. *International Journal of Hospitality Management*. 2012. Č. 31, str. 96–106.

- TOFFLER, Alvin; TOFFLER, Heidy. *Třetí vlna*. Praha: Dokořán, 2001. ISBN 80-86569-00-4.
- TYSSYADISH, L. Factors of satisfaction and intention to use peer-to-peer accommodation. *International Hospitality Management*. 2016. Č. 55, str. 70-80.
- VARMA, , Arup; JUKIC, Nenad; PESTEK, Almir; SHULTZ; Clifford J.; NESTOROV, Svetlozar. Airbnb: Exciting innovation or passing fad? *Tourism Management Perspectives*. 2016. Vol. 20, str. 228-237.
- WANG, D.; NICOLAU, J. L. Price determinants of sharing economy based accommodation rental: a study of listings from 33 cities on Airbnb.com. *International Journal of Hospitality Management*. 2017. Č. 62, str. 120-131.
- WANG, Chuhan; JEONG, Miyoung. What makes you choose Airbnb again? An examination of users' perceptions toward the website and their stay. *International Journal of Hospitality Management*. 2018, č. 74, str. 162-170.
- WILDEMUTH, Barbara M. 2009. Applications of social research methods to questions in information and library science. In: Content Analysis. Westport, Conn: Libraries Unlimited.
- WILLS, Grant. *South Carolina Law Review*. 2017. Vol. 68, n. 4, p. 821-843.
- YANNOPOULOU, Natalia; MOUFAHIM, Mona; BIAN, Xuemei. User-Generated Brands and Social Media: Couchsurfing and AirBnb. *Contemporary Management Research*. 2003. Vol. 9, č. 1. str. 86-90.
- ZAID, Alrawadieh. Exploring entrepreneurship in the sharing accommodation sector: Empirical evidence from a developing country. *Tourism Management Perspectives*. 2018. Vol. 28, str. 179-188.
- ZERVAS, Georgios; PROSERPIO, Davide; BYERS, John W. The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research*. 2017. Vol. 54, č. 5, str. 687-705.

Seznam příloh

Příloha 1: Kampaň Airbnb „Vítej doma“ (obrázek)

Příloha 2: Kampaň Airbnb „Žij tam“ (obrázek)

Příloha 3: Dotazník zasílaný e-mailem všem ubytovaným přes platformu Airbnb (rok 2020) (obrázek)

Příloha 4: Dotazník zasílaný e-mailem všem ubytovaným přes platformu Booking (rok 2020) (obrázek)

Příloha 5: Seznam aplikovaných stop slov z lexikonu „snowball“ (tabulka)

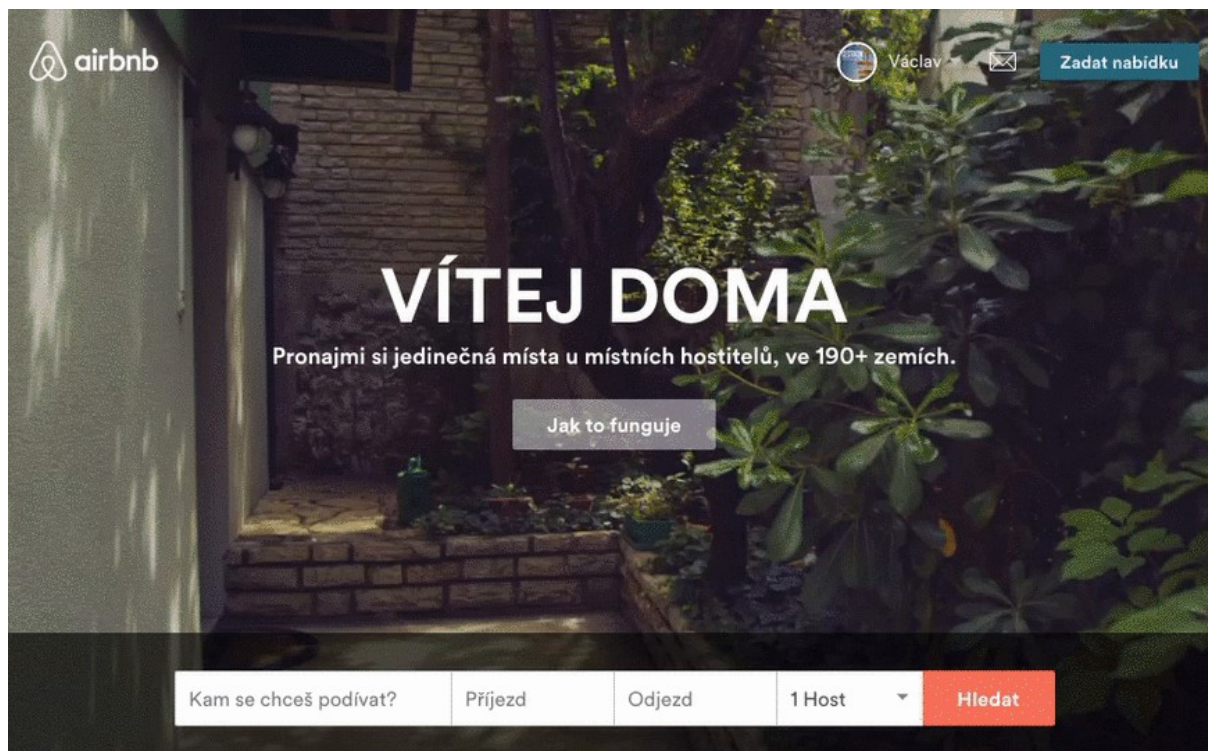
Příloha 6: Seznam ručně vytvořených a aplikovaných stop slov (tabulka)

Příloha 7: Analýza sentimentu – mrak nejčtenějších slov – emoce – Airbnb (graf)

Příloha 8: Analýza sentimentu – mrak nejčtenějších slov – emoce – Booking (graf)

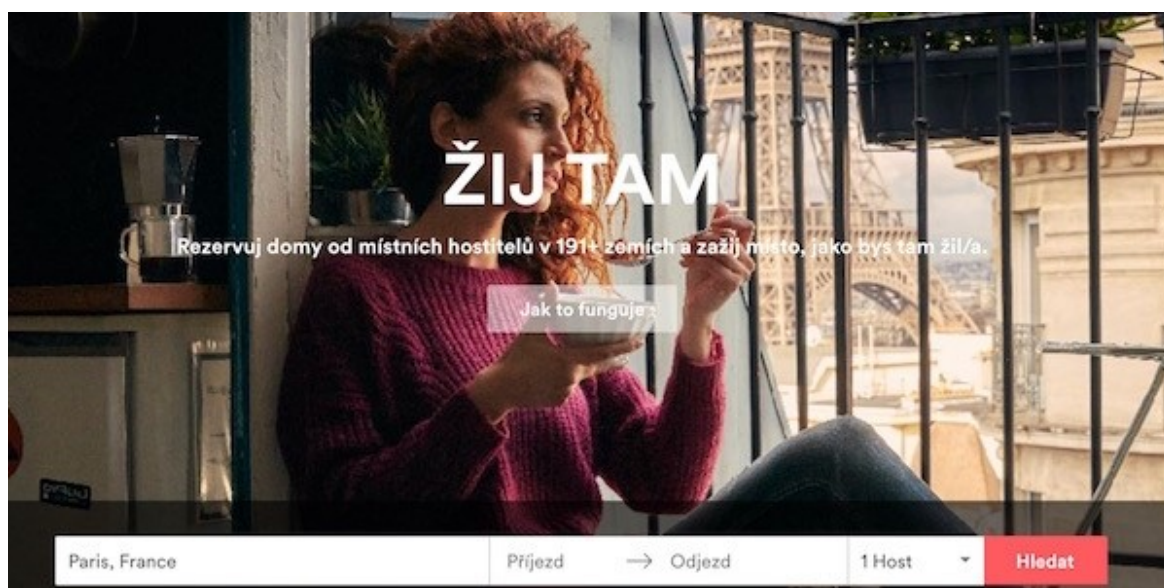
Příloha 9: Analýza sentimentu – mrak nejčtenějších slov – sentimenty pozitivní/negativní – Airbnb (vlevo) a Booking (vpravo) (graf)

Příloha 1: Kampaň Airbnb „Vítej doma“



Zdroj: <https://www.cestikon.cz/jak-se-ubytovat-s-airbnb-navod-pro-zacatecnika/>

Příloha 2: Kampaň Airbnb „Žij tam“



Zdroj:

https://www.zive.cz/getthumbnail.aspx?w=20000&h=20000&q=100&id_file=465169649

Příloha 3: Dotazník zasílaný e-mailem všem ubytovaným přes platformu Airbnb (rok 2020)

Jaký byl tvůj pobyt u hostitele [redacted]?

Tvé odpovědi nebudeme sdílet, dokud nenapišeš hodnocení také tvůj hostitel.

skvělý



Nahlásit nebezpečné chování

Další

Bylo tam něco opravdu jedinečného?

Vyber, co se ti líbilo úplně nejvíc.



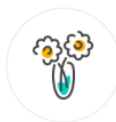
Vynikající
pohostinnost



Úžasné vybavení



Rychlé odpovědi



Promyšlená
vylepšení



Stylové místo



Zároveň čisté



Místní tipy

< Zpět

Přeskočit

Popiš svůj pobyt

Pomoz ostatním najít správné místo k pobytu. Tvůj hostitel uvidí pouze odpovědi, které vybralo více hostů.

Jak bys porovnal/a pobyt u hostitele Jan s tím, co jsi očekával/a?

☐ Mnohem lepší, než jsem čekal/a

☐ Trochu lepší, než jsem čekal/a

☐ Shodné s mým očekáváním

☐ Trochu horší, než jsem čekal/a

☐ Mnohem horší, než jsem čekal/a

< Zpět

Další

Dej uživateli [] vědět o svém zážitku v této lokalitě

Toto je důvěrné mezi tebou a tvým hostitelem.

Čistota



Přesnost



Příjezd



Komunikace



Lokalita



Cena



[< Zpět](#)

[Další](#)

Přidej neveřejnou poznámku pro []

Porad', co by mohl/a změnit nebo poděkuj za to, že byl/a skvělým hostitelem.

Přidat soukromou poznámku

[< Zpět](#)

[Přeskočit](#)

Napsat veřejné hodnocení

Řekni budoucím cestovatelům, co mohou v ubytování hostitele [] očekávat.

Napsat veřejné hodnocení

[< Zpět](#)

[Hotovo](#)



Zdroj: Airbnb

Příloha 4: Dotazník zasílaný e-mailem všem ubytovaným přes platformu Booking (rok 2020)

Ohodnoťte ubytování

nocí v destinaci

1. Byl Váš pobyt služební?

☒ Ne

☐ Ano

5 kým jste cestoval/a?

Vyberte všechny relevantní možnosti

☐ Sám/a

☐ S kamarády

☐ Partner/partnerka, manžel/manželka

☐ Rodina

☐ Kolega/kolegové

☐ S domácím mazlíčkem

Splnilo ubytování Vaše očekávání?

☐ Ne

☐ Ano

☐ Předčlo má očekávání

2. Ohodnoťte toto ubytování:

Jaký byl Váš pobyt?

1	2	3	4	5	6	7	8	9	10
Hrozně					Naprosto výjimečně				
Personál									
Zařízení									

Čistota

Pohodlí

Poměr ceny a kvality

Místo

Doplňující otázka: Rozhodlo se ubytování nepoužívat jednorázová plastová brčka?

☐ Ano ☐ Ne ☐ Nevím

3. Máte minutku? Napište nám trochu více informací.

😊 Co se Vám líbilo?

Jak Vám chutnala snídaně? Jak se Vám líbila lokalita?

☹️ Co se Vám nelíbilo?

Co by mohlo být podle Vás lepší?

Napište krátký popis k hodnocení.

Zdroj: Booking

Příloha 5: Seznam aplikovaných stop slov z lexikonu „snowball“

you'd	won't	and	above	how
he'd	wouldn't	but	below	all
she'd	shan't	if	to	any
we'd	shouldn't	or	from	both
they'd	can't	because	up	each
i'll	cannot	as	down	few
you'll	couldn't	until	in	more
he'll	mustn't	while	out	most
she'll	let's	of	on	other
we'll	that's	at	off	some
they'll	who's	by	over	such
isn't	what's	for	under	no
aren't	here's	with	again	nor
wasn't	there's	about	further	not
weren't	when's	against	then	only
hasn't	where's	between	once	own
haven't	why's	into	here	same
hadn't	how's	through	there	so
doesn't	a	during	when	than
don't	an	before	where	too
didn't	the	after	why	very

Zdroj: balíček „stopwords“, lexikon Snowball, software R, vlastní zpracování

Příloha 6: Seznam ručně vytvořených a aplikovaných stop slov

really	didn't	want	due	mind
stuff	especially	work	second	note
everything	take	see	anywhere	can't
also	still	able	couldn't	point
one	without	living	thought	told
nothing	stuff	part	week	set
us	lots	full	thinks	upon
just	overall	going	end	beyond
bit	although	everyone	different	five
can	however	fully	bigger	despite
get	etc	took	side	said
time	another	since	throughout	wasn't
even	never	let	combined	someone
quite	things	might	noone	wasn
next	highly	getting	signage	there's
will	anything	know	passed	try
much	ever	say	led	s
around	went	across	wonder	maybe
back	puts	less	real	buy
main	put	whole	four	probably
every	something	came	else	already
two	well	three	literally	along
super	little	almost	longer	certainly
away	go	may	hour	finish
lot	made	takes	others	finishes
got	within	think	you're	finished
many	short	middle	live	i.p
definitely	ish	times	must	march
extremely	absolutely	yet	via	start
always	anyone	makes	using	hand
though	sure	don't	otherwise	case
it's	everywhere	last	self	number
thing	days	coming	later	u

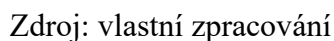
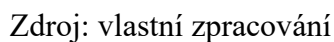
done	c	mere	turn	therefore
half	besides	ins	finally	pm
rather	basically	amd	d	doesn
completely	behind	could've	often	non
seeing	that's	additional	whilst	six
taking	instead	step	reason	cause
except	till	watch	add	maker
common	prior	j	meant	hopefully
i've	smaller	surely	pass	ones
wants	moment	you'll	didnt	everyday
together	anyway	saw	regards	started
general	4th	doesn't	daily	i'll
tried	4pm	letting	cons	moreover
year	tell	trying	prime	recently
i'd	considering	previous	3rd	seem
i'm	clearly	taken	prefer	throw
seems	round	mostly	twice	matter
least	means	relatively	lived	malo
consider	somewhere	wouldn't	potential	past
goes	fi	properly	whose	afternoon
sometimes	completed	years	accordingly	definetly
points	forthcoming	absolute	all	definatly
b	weeks	months	assume	mornings
regarding	ended	won't	assumed	advantages
parts	30pm	sound	assuming	advantage
seen	ensuite	third	n	disadvantage
opposite	ton	read	everybody	disadvantages
tiny	tons	ten	we've	ca
per	recent	wanting	dont	db
possibly	6pm	turned	example	pros
fairly	f	notice	defo	tha
comes	24hr	filled	beside	ve
seemed	daytime	month	indeed	ect
finding	iv	isn't	generally	usually

says	8pm	white	oh	depending
t	easter	unless	runs	friday
bought	hadn't	wrote	fall	likely
we'd	infront	thinking	keeping	entered
written	ear	con	passing	initially
5th	highlight	eve	1st	mr
st	perhaps	seconds	sounds	average
total	none	x	weren	he's
starting	correct	barely	wouldn	leaves
forget	weren't	onto	time	everytime
yes	lives	sort	speaks	wery
guess	wanna	hardly	seriously	compare
mile	we're	e.g	def	compared
opinion	importantly	saying	alot	comparion
hold	anybody	ist	mainly	everyting
holding	saturday	such	whether	caused
wednesday	m	appears	cant	fifteen
weekday	co	bc	re	heavily
hundret	nevertheless	realy	10pm	factor
night's	highlights	ms	hi	afterwards
day	fell	00am	fourth	v
days	slight	today	constantly	meaning
nights	you'd	tomorrow	couldn	whatsoever
night	2nd	1am	preferred	everthing
didn't	towards	andrew	red	held
continue	biggest	definetely	mid	among
continues	compared	ratio	similar	lastly
continued	forgot	we'll	she's	Pods
based	figure	definitely	reasons	sunday
whatever	possibility	ahead	thus	nobody
secondly	word	black	e	speaking
april	ran	it's	passes	august
tad	repeat	according	became	pre
stiff	furthermore	republic	constant	6th

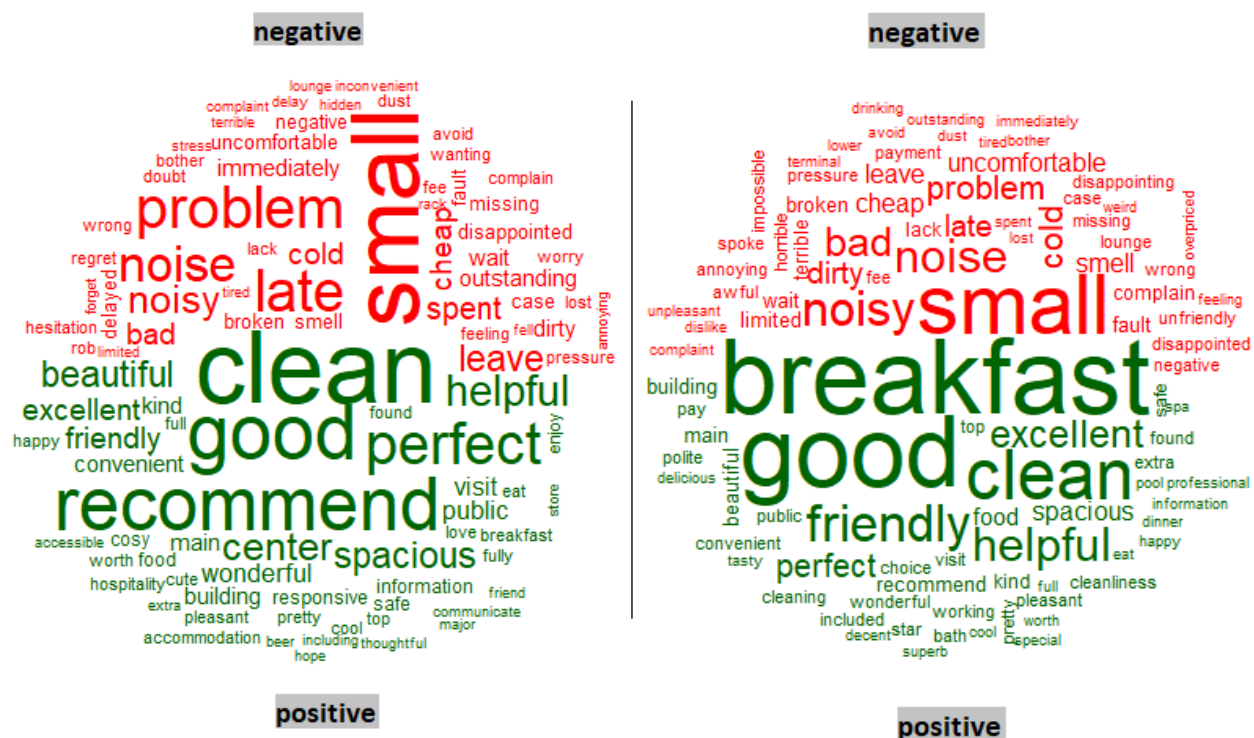
july	vše	underneath	spring	caught
lesser	et	needing	btw	could've
te	regard	november	ll	current
neither	vs	somehow	definitively	currently
june	30am	wich	haven't	previously
tho	alla	firstly	minimum	nowhere
entering	amongst	hence	they're	turning
kinda	im	wont	i.e	turns
knowing	realize	24h	zero	conclusion
regularly	fifth	pair	200m	lol
frequently	resolve	ago	starts	numbers
p	buying	begin	agree	l l am
prem	ip	final	marked	absolutly
couldnt	multi	concerning	p.s	regardless
w	autumn	10am	7am	necessary
mine	beneath	apparent	initial	necessarily
wasnt	disadvantage	downsides	affect	unnecessary
apparently	upper	haha	sq	hey
aswell	consideration	ans	unlike	cm
100m	sometime	anyways	entirely	shouldn
keeps	nonetheless	bits	related	fyi
l	und	cases	section	com
sat	fill	isnt	pro	p.m
thats	moments	appeared	result	isn
pointed	ps	h	la	christmas
43952	r	ur	somebody	republiky
december	eventually	ie	6am	thé
someday	falling	raining	forth	tlc
aren't	yellow	shouldn't	z	bcc
gonna	k	thru	definite	parc
l lpm	originally	extremly	verry	adequately
hello	sooo	albeit	blue	comms
negatives	hundred	sits	rarely	latter
seven	pieces	suche	would've	trough

everything's	nine	you've	built	na
gold	eventhough	lets	specifically	obvious
typically	ongoing	mrs	specific	october
ver	totally	what's	happens	sady
year's	gotten	explanation	become	significant
thoughts	nighttime	til	became	doesnt
colourful	perspective	9pm	marks	equally
colorful	summertime	monday	ways	indications
defenetly	stand	pod	way	january
j's	matters	compare	words	o
minuses	stood	th	aspects	soooo
minus	watched	venue	aspect	theres
plus	september	majority	considered	kept
pluses	shall	realise	consider	major
2km	eight	realize	considers	totally
500m	ultra	realizing	3pm	slightly
arround	would've	realized	word	enter
stone's	year	realised	3am	soon
yo	year's	appear	4am	happened
g	occasionally	appeared	shouldn	somewhat
midst	defenitely	appearing	5am	happen
mn	tries	2am	brown	reality
mod	depends	overly	7th	certain
nase	depend	abit	sudden	8am
pdf	depending	flow	suddenly	a.m
un	virtually	removed	9am	didn
everithing	50m	button	altogether	de
lot's	overview	stuffs	highly	2pm
partly	soo	suppose	working	frequent
whoever	twenty	supposed	works	don
everything's	xx	nil	eg	practically
february	who's	pull	farther	watching

Zdroj: vlastní zpracování



Příloha 9: Analýza sentimentu – mrak nejčtetnějších slov – sentimenty pozitivní/negativní – Airbnb (vlevo) a Booking (vpravo)



Zdroj: vlastní zpracování