

Univerzita Hradec Králové
Fakulta informatiky a managementu
Katedra managementu

Měření efektivnosti chatbota
Diplomová práce

Autor: Bc. Andrea Kotoučková
Studijní obor: Informační management

Vedoucí práce: prof. Ing. Hana Mohelská, Ph.D.

Hradec Králové

duben 2020

Prohlášení:

Prohlašuji, že jsem diplomovou práci zpracovala samostatně a s použitím uvedené literatury.

V Hradci Králové dne 25. 4. 2020

Bc. Andrea Kotoučková

Poděkování:

Děkuji vedoucí diplomové práce prof. Ing. Haně Mohelské, Ph.D. za metodické vedení, připomínky a cenné rady, které pomohly při zpracování této práce. Dále děkuji odborníkům z řad zaměstnanců společnosti Československá obchodní banka, a. s. za jejich čas a poskytnuté informace za pomoci kterých mohlo dojít k realizaci analýz potřebných k vypracování aplikační části diplomové práce.

Anotace

Diplomová práce se zabývá představením a porovnáním dostupných přístupů k měření efektivnosti chatbota a poté rozebráním jednoho vybraného. Chatbot je jedním z trendů čtvrté průmyslové revoluce a prozatím neexistuje přístup k měření jeho efektivnosti, který by se běžně používal. Z toho plyne nevědomost, které oblasti je u něj potřeba posuzovat. Cílem práce je hlouběji proniknout do této problematiky a uvažovat o jejích závěrech. Teoretická část seznamuje s pojmem chatbot, jeho spojením s umělou inteligencí, klasifikací a současným využitím. Dále vymezuje význam účinnosti a účelnosti vzhledem k IS/ICT a představuje konkrétní přístupy k měření efektivnosti chatbota. V aplikační části je podle definovaných kritérií a jejich vah vybrán jeden vhodný přístup, podle kterého je následně rozebrán postoj v reálné společnosti a na základě souhrnu získaných teoretických a praktických poznatků je navrženo jeho rozšíření.

Annotation

Title: Measurement the Effectiveness of Chatbot

The Diploma Thesis deals with introduction and comparison accessible approaches to measuring the effectiveness of chatbot and then follows analyzing selected one. A chatbot is one of the trends of Industry 4.0 and for now, there is no kind of approach that is commonly used. This leads to ignorance of knowing what is important to assess. The purpose of the thesis is a deeper understanding of this problematics and reaching conclusions. The theoretical part introduces the meaning of chatbot, its connection with artificial intelligence, classification and current use. Then there the meanings of efficiency and effectiveness concerning IS/ICT are explained and approaches to measuring the effectiveness of chatbot are characterized. In the application part, according to the defined criteria and their relative importance, one suitable approach is chosen and then follows its analysis in the real company. Based on the summary of theoretical and practical findings is proposed the extension of the selected approach.

Obsah

1 Úvod	1
2 Cíl práce a metodika	3
3 Chatbot	5
3.1 Vymezení pojmu chatbot.....	5
3.2 Spojení s umělou inteligencí.....	6
3.2.1 Strojové učení	8
3.3 Klasifikace chatbota.....	9
3.4 Současné využití chatbota	12
3.4.1 Obchod.....	14
3.4.2 Zákaznický servis	15
3.4.3 Využití uvnitř podniku	15
3.4.4 Osobní asistenti.....	15
3.4.5 Integrovaní boti	16
3.4.6 Notifikační boti.....	16
3.4.7 Hry.....	16
4 Proces měření efektivnosti chatbota	17
4.1 Efektivnost.....	17
4.1.1 Účinnost a účelnost.....	17
4.1.2 Efektivnost IS/ICT.....	18
4.2 Přístupy k měření efektivnosti chatbota	24
5 Vymezení role chatbota ve vybraném podniku	34
5.1 Představení vybraného podniku	34
5.2 Charakteristika vybraného chatbota.....	36
5.2.1 Účel chatbota.....	36

5.2.1.1 Poslání a vize	37
5.2.2 Definice chování chatbota.....	38
5.2.3 Dostupnost a obsluha	43
6 Výběr vhodného přístupu k měření efektivnosti chatbota.....	45
6.1 Stanovení kritérií.....	46
6.2 Stanovení vah jednotlivých kritérií.....	46
6.3 Výběr alternativ.....	48
6.4 Porovnání alternativ podle stanovených kritérií	49
6.4.1 Uživatelská přívětivost.....	49
6.4.2 Informační schopnost a vybavenost	50
6.4.3 Jazyková úroveň a vybavenost.....	51
6.4.4 Lidskost.....	52
6.4.5 Obchodní hledisko	52
6.5 Vyhodnocení výběru vhodného přístupu k měření efektivnosti chatbota.....	53
7 Analýza procesu měření efektivnosti chatbota ve vybraném podniku	56
7.1 Perspektiva uživatelské přívětivosti.....	57
7.2 Perspektiva vyhledávání (generování) informací.....	59
7.3 Lingvistická perspektiva	61
7.4 Technologická perspektiva	62
7.5 Obchodní (business) perspektiva.....	63
8 Shrnutí a diskuze výsledků	66
8.1 Shrnutí analýzy procesu měření efektivnosti chatbota ve vybraném podniku.	67
8.2 Diskuze vybraného přístupu	71
9 Závěry a doporučení	75
10 Seznam použité literatury	78

10.1 Tištěné zdroje	78
10.2 Elektronické zdroje	79
11 Přílohy	86

Seznam obrázků

Obrázek 1: Konverzační rámec chatbota.....	11
Obrázek 2: Možnost otevření chatovacího rozhraní na liště chatu	39
Obrázek 3: Chatovací rozhraní	39
Obrázek 4: Ukázka AHP hierarchie	45
Obrázek 5: Porovnání alternativ podle kritéria „Uživatelská přívětivost“	50
Obrázek 6: Porovnání alternativ podle kritéria "Informační schopnost a vybavenost"	51
Obrázek 7: Porovnání alternativ podle kritéria "Jazyková úroveň a vybavenost"	51
Obrázek 8: Porovnání alternativ podle kritéria "Lidskost"	52
Obrázek 9: Porovnání alternativ podle kritéria "Obchodní hledisko"	53
Obrázek 10: Výsledek výběru vhodného přístupu k měření efektivity chatbota v procentuálním vyjádření.....	53
Obrázek 11: Grafické znázornění výběru vhodného přístupu k měření efektivity chatbota.....	54

Seznam tabulek

Tabulka 1: Stanovené váhy kritérií a jejich výsledné pořadí.....	47
Tabulka 2: Souhrnný rámec vybraného přístupu k měření efektivnosti chatbota (1. část)	68
Tabulka 3: Souhrnný rámec vybraného přístupu k měření efektivnosti chatbota (2. část)	69
Tabulka 4: Saatyho matice – E1.....	89
Tabulka 5: Saatyho matice – E2.....	89
Tabulka 6: Saatyho matice – E3.....	90
Tabulka 7: Saatyho matice – E4.....	90
Tabulka 8: Saatyho matice – E5.....	90
Tabulka 9: Saatyho matice – E6.....	90
Tabulka 10: Stanovení výsledných (jednotných) vah kritérií a jejich příslušného pořadí	91

Seznam grafů

Graf 1: Účel užívání chatbota.....	14
------------------------------------	----

Seznam schémat

Schéma 1: Ukázka konverzačního stromu (úryvek ze základního scénáře).....	42
Schéma 2: Proces schvalování úpravy a tvorby textace nových odpovědí	61
Schéma 3: Dekompozice rámce vybraného přístupu k měření efektivnosti chatbota	67
Schéma 4: Návrh na rozšíření vybraného přístupu k měření efektivnosti chatbota....	74

1 Úvod

Příchod a přijetí iniciativy „Industry 4.0“, v České republice představováno pod názvem „Průmysl 4.0“ nebo zjednodušeně čtvrtá průmyslová revoluce, postupně změnil a stále mění filozofii myšlení dnešních společností. Pod tímto pojmem si lze představit transformaci výroby na plně integrované automatizované a průběžně optimalizované prostředí, kdy se nejedná pouze o digitalizaci průmyslové výroby, ale o komplexní systém změn spojený s fenomény jako je například Internet věcí, služeb a lidí, autonomní roboti a rozvoj umělé inteligence, analýza velkých dat (Big Data), digitální dvojče, virtualizace, cloud computing či rozšířená realita. Třetí průmyslová revoluce dala společnostem podobu informační, resp. informačně znalostní, která spočívá ve schopnosti účelného nakládání s informacemi, čtvrtá průmyslová revoluce poté navazuje a umožňuje pomocí nově zavedených technologií významně zefektivnit využití zdrojů.

Tato práce se věnuje technologii nazvané chatovací robot, běžněji pouze chatbot, kterou lze zařadit mezi výše uvedené trendy. Její počátky sice sahají do 60. let 20. století, své využití ovšem začala nacházet především v několika posledních letech, a to v různých formách, které jsou popsány v diplomové práci. Vzhledem k jejímu hojnému nasazování se často hovoří o faktu, zda opravdu splňuje svůj účel či se jedná pouze o utopickou představu schopnosti robota efektivním způsobem nahradit činnost člověka. Aby si podnik dokázal na tuto otázku odpovědět, potřebuje mít správně zavedený proces měření efektivnosti této technologie. Dosud ovšem nebyl vytvořen jednotný přístup, který by se běžně používal, a právě touto problematikou se práce zabývá.

Diplomová práce je rozdělená na teoretickou a aplikační část, kdy každá z nich je systematicky rozčleněna do více kapitol. V první části se lze seznámit s tím, co chatbot znamená, včetně jeho spojení s umělou inteligencí, jak je klasifikován a v současné době využíván. Na to navazuje zasvěcení do problematiky měření efektivnosti IS/ICT a poté jsou konkrétně představeny dostupné přístupy k měření efektivnosti chatbota.

Druhá část, tedy aplikační, se zabývá porovnáním přístupů k měření efektivnosti chatbota charakterizovaných v předchozí části, následným výběrem jednoho vhodného, jeho rozebráním z procesního hlediska v reálné společnosti se zaměřením na jeden konkrétní typ chatbota a návrhem na rozšíření o získané teoretické a praktické poznatky plynoucí z průběhu celé práce.

2 Cíl práce a metodika

Cílem diplomové práce je představit a porovnat dostupné přístupy k měření efektivnosti chatbota, z nich poté vybrat jeden vhodný s jehož pomocí lze analyzovat stav daného procesu v reálné společnosti a na základě získaných teoretických a praktických poznatků navrhnout rozšíření právě tohoto přístupu.

V teoretické části práce je nejdříve provedena rešerše odborné literatury zaměřená na problematiku chatovacích robotů, měření efektivnosti IS/ICT a přístupy k měření efektivnosti chatbota, která je v přímém kontextu s následnou aplikační částí. Čerpáno bylo jak z tištěných, tak i elektronických zdrojů, kdy se především jedná o zahraniční studie získané z vědeckých databází, jako je například Science Direct, Springer, Web of Science. K jejich vyhledání byla použita klíčová slova „chatbot“, „effectiveness“, „measurement“, „assessment“, „evaluation“, „metrics“, a to v různých kombinacích.

V aplikační části je využito kvantitativního i kvalitativního výzkumu, přičemž konkrétními technikami sběru dat jsou: studium dokumentů, dotazníkové šetření v podobě Saatyho metody párového porovnání, jehož výstupem je expertní stanovisko a semistrukturovaný rozhovor. Nejdříve je představena společnost Československá obchodní banka, a. s. (dále jen „ČSOB“), její dvě Klientská centra nacházející se v Hradci Králové a Náchodě a poté zainteresovanost a úspěchy v oblasti robotizace, inovace a digitální transformace. Následně je charakterizován jeden z jejích chatbotů, o jehož správu, podporu a další rozvoj se starají pracovníci z Klientského centra. Dané informace jsou získány z Výpisu z obchodního rejstříku, oficiálních internetových stránek a interní dokumentace společnosti. Tato zjištění jsou důležitá pro uskutečnění potřebných analýz vedoucích k naplnění cíle práce. Jako první je provedena vícekriteriální analýza v nástroji pro podporu rozhodování Expert Choice 2000, za pomoci které je vybrán vhodný přístup k měření efektivnosti chatbota. K tomu je mimo jiné žádoucí stanovit váhy jednotlivých kritérií. Pro jejich objektivní určení bylo využito odbornosti šesti zainteresovaných pracovníků Klientského centra, kteří byli podrobena šetření ve formě párového porovnávání definovaných kritérií. Na základě jednotlivých výsledků byly sestaveny Saatyho matice a ty vedly ke konečnému výstupu v podobě expertního stanoviska vah daných kritérií. V následujícím kroku je provedena analýza rozebírající měření efektivnosti

konkrétního chatbota v ČSOB podle vybraného přístupu, a to za pomoci semistrukturovaných rozhovorů s příslušným procesním manažerem, který je zároveň jeho gestorem, dále studia interní dokumentace a softwaru pro jeho administraci. Posledním krokem je souhrn těchto a poznatků získaných v teoretické části práce, které vedou k návrhu na rozšíření vybraného přístupu k měření efektivnosti chatbota.

K naplnění cíle diplomové práce je potřeba nalézt odpověď na následující výzkumnou otázku:

V1: Jak je možné měřit efektivnost chatbota?

3 Chatbot

Diplomová práce se věnuje přístupu k procesu měření efektivnosti jednoho z trendů čtvrté průmyslové revoluce, kterým je chatovací robot. S touto technologií je často spojován další fenomén, kterým je umělá inteligence. V rámci této kapitoly jsou pro účely práce představeny oba již zmíněné pojmy, jejich souvislost, klasifikace chatbota a jeho soudobé využití.

3.1 Vymezení pojmu chatbot

Pro pojem chatbot, někdy též uváděného pod názvem „chatterbot“ či pouze „chatovací robot“, lze najít vícero různých definic, které říkají, že se jedná o technologii automatizující, resp. simulující lidskou konverzaci či ho přímo označují za umělou inteligenci, což jak je vysvětleno v nadcházejících podkapitolách není zcela pravdivé tvrzení. Ve spojení s výše uvedeným, tuto technologii přijatelným a jasným způsobem vymezují Shawar a Atwell (2007), a to jako počítačový program zprostředkávající interakci mezi jím a živou osobou za využití přirozeného jazyka. To v překladu znamená, že by během daného procesu nemělo být poznat, že člověk komunikuje s robotem a mělo by tak docházet k naprosto bezprostřednímu přenosu informací mezi oběma stranami. Dále Brandtzaeg a Følstad (2017) upřesňují, že k interakci dochází skrz chatovací rozhraní, které společnosti často umísťují na svoje webové stránky nebo využívají existujících platforem, jako je Facebook Messenger, WhatsApp Messenger, Skype, Slack a jiné. Někteří mezi chatboty zařazují i hlasově poháněné virtuální asistenty jako je Siri, S Voice, Google Assistant, Cortana či Alexa (Rieke, 2018).

I když se chatbot stal trendem až v několika posledních letech (Accenture, 2016), tak Lokman a Zain (2010) uvádějí, že se jeho počátky datují už k roku 1966, kdy MIT profesor Joseph Weizenbaum představil prvního chatovacího robota, nazvaného ELIZA, jehož chování bylo založeno na jednoduchém principu hledání a vytváření výstupů podle klíčových slov daných vstupů, podrobených tzv. dekompozičním (transformačním) pravidlům, a stal se tak později inspirací pro ostatní chatboty. V roce 1995 nastal velký průlom této technologie, když Dr. Richard S. Wallace přišel s jeho moderní verzí ELIZA, pojmenovanou jako A.L.I.C.E., jejíž název je odvozen od Artificial Linguistic Internet Computer Entity. Ta poté dala ke vzniku novému

značkovacímu jazyku AIML (Artificial Intelligence Markup Language), založenému na XML (eXtensible Markup Language) dialektu, vytvářejícím přirozeně mluvící softwarové agenty. Dále autoři zmiňují, že s myšlenkou, zda mohou stroje myslet a jak tuto skutečnost ověřit, přišel v roce 1950 matematik Alan Turing, jehož test spočívá v rozpoznání, zda rozhodčí (opravdový člověk) rozezná, jestli konverzuje s chatbotem nebo lidským jedincem. Přesněji se vše odehrává ve dvou oddělených prostorách, v jedné z místností sedí rozhodčí a ve druhé lidský a umělý odpovídající. Tázající pokládá otázky, na které dostává odpověď. V případě, že nedojde k odhalení konverzace s chatbotem, tak stroj splnil Turingův test. I když A.L.I.C.E. tímto testem neprošla, tak třikrát získala Loebnerovu cenu, a to v letech 2000, 2001 a 2004. Loebnerova cena je od roku 1991 každoročně pořádanou soutěží založenou právě na splnění Turingova testu. Wakefield (2019) vysvětluje pravidla hry, podle kterých je každému stroji položeno stejných 20 otázek odlišených různým stupněm složitosti, konverzace probíhají 25 minut a vítězem se stává ten, který dokáže přesvědčit více než polovinu rozhodčích o tom, že je člověk. To se ovšem zatím nikdy nepodařilo, a tak si dosud cenu odnášel ten stroj, který dokázal „oblbnout“ nejvíce rozhodčích. Russel a Norvig (2010) zmiňují přísnou kritiku od Shieber (1994), co se týká užitečnosti tohoto testu v rámci soutěže Loebnerova cena.

Od ELIZY, přes A.L.I.C.E. a další úspěšné chatboty, jako je současný pětinasobný vítěz Loebnerovy ceny chatbot Mitsuku, vytvořený Stevem Worswickem (Wakefield, 2019), se v dnešním digitálním světě stávají chatboti běžnou součástí všech typů firem. Ty ho využívají k různým účelům, s nimiž se lze seznámit především v nadcházející podkapitole 3.4.

3.2 Spojení s umělou inteligencí

Pojetí umělá inteligence z anglického Artificial Intelligence (dále jen „AI“) nelze jednoznačně definovat. Jedná se totiž o interdisciplinární vědu „o vytváření strojů nebo systémů, které budou při řešení určitého úkolu užívat takového postupu, který – kdyby ho dělal člověk – bychom považovali za projev jeho inteligence“ (Minsky, 1967). Tato definice vychází z Turingova testu a lze ji volně přeložit tak, že složitost řešených úloh vyžaduje použití inteligence člověka. V tomto případě je potřeba se zaměřit na otázku co je to složitost a inteligentní řešení. Čermák (2018) uvádí,

že složitost se vyznačuje počtem všech možných řešení a druhý atribut je omezen znalostmi. Dále dodává, že umělá inteligence se zabývá hledáním hranic včetně reprezentace nabývaných poznatků a procesů, osvojováním a způsobem jejich použití při řešení problémů a využívá různých přístupů a algoritmů k nalezení východiska velmi složitých úloh.

Základní otázkou, zda mohou stroje myslet, se zabývali už v 17. století filozofové jako Descartes, Pascal, Hobbes, jejichž myšlenky se pohybovaly pouze na teoretické rovině a nebyli jej schopni přeměnit ve skutečnost. Dále Čermák (2018) zdůrazňuje rok 1950, ve kterém přišel Alan Turing se svým testem, na základě kterého shromáždil řadu důkazů proti inteligentním strojům a ty následně vyvrátil. Pod vedením Johna McCarthyho se v roce 1956 konala konference na Dartmouth College, na které se sešli odborníci zajímající se o možnosti implementace lidského myšlení na stroje (počítače). Autor zmiňuje další důležité milníky v historii vývoje umělé inteligence, mezi které patří A. Newell, H. Simon a jejich program založený na technikách heuristického vyhledávání – the Logic Theorist, systém napodobující lidské myšlení při řešení úloh ve stavovém prostoru pomocí snižování rozdílů – General Problem Solver (GPS). V roce 1958 vytvořil John McCarthy jazyk určený pro umělou inteligenci – LISP. Na počátku 70. let vznikl jazyk PROLOG, jehož autorem je A. Colmerauer. Dále v tomto období byly navrženy různé univerzální systémy, které nebyly schopny řešit vysoce specializované úlohy. Mezi ně patří například Planner. Začalo se poukazovat na zásadní problém umělé inteligence, a to reprezentace, využití a zpracování znalostí. To poté vedlo k rozvoji expertních systémů, jejichž úkolem je simulace rozhodovacího procesu specialistů při řešení komplikovaných úloh za využití báze znalostí. Úspěšnými průkopníky expertních systémů byli například MYCIN a PROSPECTOR. V roce 1981 byl v Japonsku vyhlášen Projekt počítačů 5. generace na zpracovávání nenumerních informací. Tyto počítače pracovaly s rozsáhlými znalostními bázemi a vyžadovaly učící mechanismy. Dále T. Kohonen v roce 1988 představil „elektronickou písáčku“. Ta představovala převod hlasu na anglický text prostřednictvím kohonenových neuronových sítí. V roce 1990 L. Adleman začal s vytvářením koncepce DNA počítače a prokázal následnou proveditelnost základních matematických operací a výpočtů.

Shah (2017) uvádí, že v souvislosti s chatbotem a umělou inteligencí se váže problematika zpracování přirozeného jazyka, s kterým se lze setkat pod zkratkou NLP (Natural Language Processing). Chatboti jsou schopni fungovat za pomoci předem daných pravidel založených na jazykových strukturách (tzv. rule-based) či pomoci statistického modelu zpracování přirozeného jazyka, čímž se zabývá tzv. strojové učení.

3.2.1 Strojové učení

Slouka (2019) strojové učení definuje jako jednu z podoblastí a klíčových součástí umělé inteligence a uvádí, že se jedná o dynamické přizpůsobování se novým situacím založeném na zjištěních z dostupných dat. Dále zmiňuje, že je vycházeno ze statistických metod, kdy je model tvořen z algoritmů strojového učení a tzv. trénovacích dat. Do této problematiky lze v kontextu zařadit hluboké učení a umělé neuronové sítě, jejichž inspirací je podle Čermák (2018) chování biologických struktur (lidského mozku) a je využívána především schopnost adaptace a učení se. Slouka (2019) jmenuje tři kategorie algoritmů strojového učení:

- **učení s učitelem** – v tomto případě je při trénování vyžadován dohled „učitele“ nad strojem, který rovněž zadává vstupní data, je předem známá požadovaná odezva, o jejímž hodnocení je systém okamžitě informován a po dokončení tréninku je algoritmus schopen aplikovat získané znalosti,
- **učení bez učitele** – jedná se o mnohem složitější metodu, která nevyžaduje dohled „učitele“. Algoritmy si prohlížejí data a vyhodnocují závěry samostatně, kdy po vytrénování se, může díky obrovskému množství informací docházet k vlastním interpretacím nových výstupů,
- **zpětnovazební učení** – s tímto pojmem se lze setkat i pod názvem posilované učení, kdy se stroj, stejně jako u učení s učitelem, snaží o hledání nejlepší cesty či chování k dosažení požadovaného. Rozdíl je ovšem v tom, že při trénování metody učení s učitelem je známa existující správná odpověď a u zpětnovazebního učení nikoli. Algoritmus na výstup tedy přichází sám a spočívá v maximalizaci odměny, resp. udělování pozitivní odměny za chování, čímž je podpořeno a negativní odměny za chování, které je potřeba vytěsnit.

Slouka (2019) dále rozlišuje specializovanou a obecnou umělou inteligenci. Specializovanou představují například virtuální asistenti Siri, Alexa, Cortana, kteří jsou sice svým působením velice rozšíření, ovšem registrují určitá omezení. Obecná umělá inteligence by měla být schopna kompletně uvažovat a jednat jako člověk, tu se ovšem zatím vytvořit nepodařilo.

Podle Shah (2017) je na rozdíl od chatbota založeného na pravidlech, chatbot založený na strojovém učení více robustní a dokáže tedy lépe reagovat na neznámé vstupy. Při dodávání většího množství trénovacích dat se stává generování jeho odezvy přesnější. Na druhou stranu chatboti založení na pravidlech jsou předvídatelnější, lépe kontrolovatelní a pro vykonávání rutinních úkolů dostačující.

3.3 Klasifikace chatbota

Shevat (2017) klasifikuje chatbota do několika kategorií. Pro představu o fungování a zařazení chatbota vybraného pro aplikační část a pochopení celkových výsledků práce, je vhodné ozřejmit, jak se od sebe jednotlivé typy liší.

Technologie chatovacího robota pracuje nad určitou doménou, kterou může být:

- **uzavřená doména** – jedná se o chatboty zaměřené na jeden konkrétní produkt či službu, jako je například sjednávání cestovního pojištění, půjček, koupě letenek, elektroniky (většina takových chatbotů dokáže hovořit jen o daném tématu a obecných otázkách ve stylu „Jak se máš?“ apod.),
- **otevřená doména** – v tomto případě dokáže bot hovořit v podstatě o čemkoli. K této dovednosti využívá integraci více služeb, vyhledávání na internetu a tak dále (např. virtuální asistent Siri).

Dále autor rozlišuje chatbota podle typu konverzace:

- **zaměřené na úkol** – pro tento druh chatbota je typické následování přesného konverzačního toku („konverzačního flow“), který uživatele po jednotlivých krocích dovede ke stanovenému cíli (např. objednávání jídla, sjednávání pojištění),

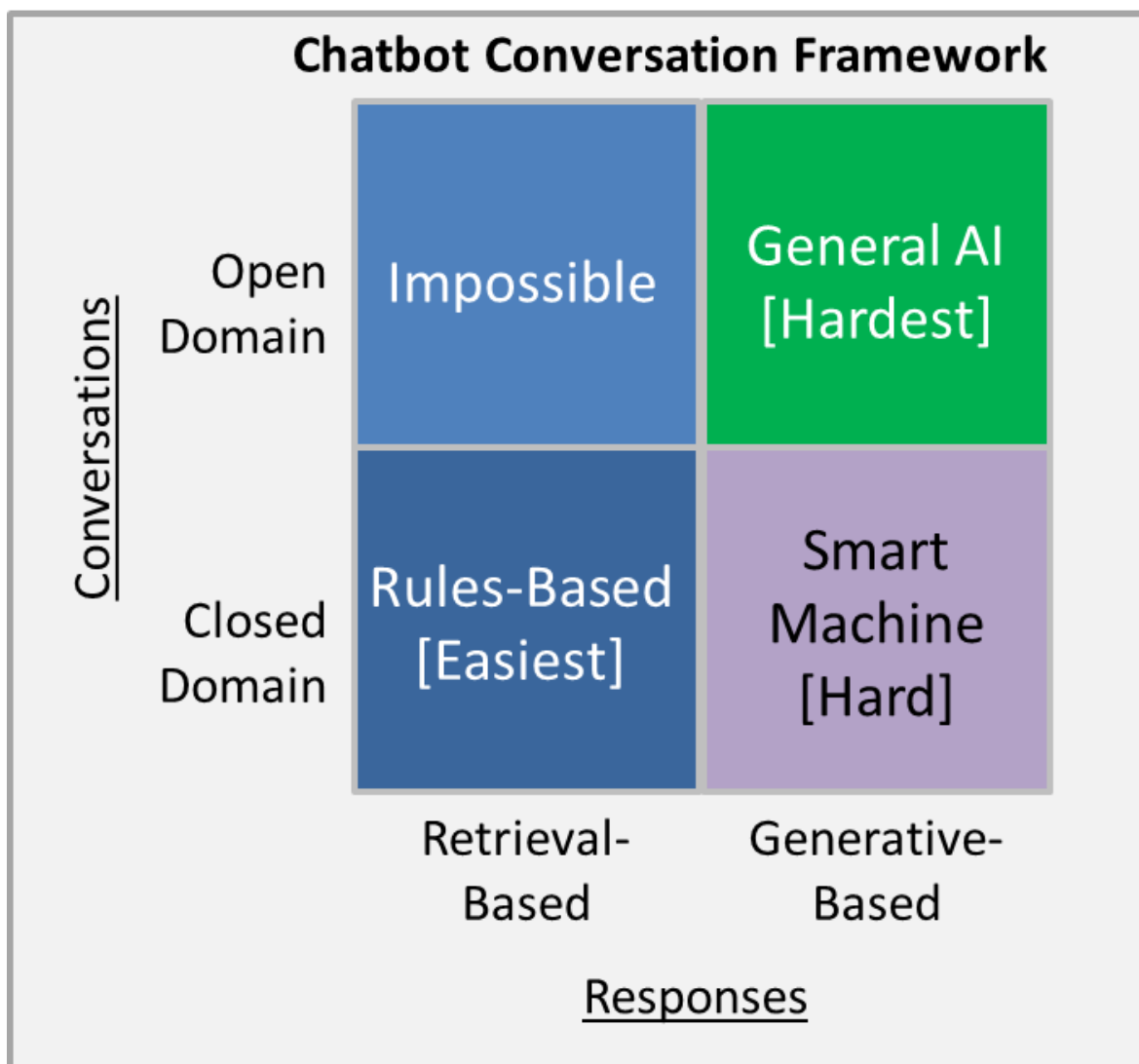
- **zaměřené na téma** – zde je chatbot uzpůsoben k výměně informací a udržování neřízené konverzace, která by měla doslova působit přirozeně (není poznat, že je na druhé straně robot a nikoli živá osoba).

U této technologie je především důležitá volba způsobu implementace tvorby odpovědi, protože na té stojí velká část ne/úspěchu jejího zavedení. Existují dvě možnosti, které Sanjeevi (2018) charakterizuje následujícím způsobem.

- **Retrieval-based model** – metoda založená na vyhledávání vhodné odezvy ze sady otázek a odpovědí, která často využívá heuristického výběru na základě vstupu a kontextu konverzace. Výběr může být omezen pouze na škálu nejlépe hodnocených odpovědí či může být různorodější, to znamená výběr náhodné reakce ze skupiny. K určité otázce se hodí více odpovědí, které chatbot využívá, aby působil přirozenějším a méně strojovým dojmem. Tento postoj je vhodnější a uživatelsky přívětivější. Výhodou tohoto modelu je gramatická a věcná správnost stavby vět, plní 100% funkci u jednodušších úkolů (hodí se implementovat v případě menšího množství dat), což znamená, že dokáže získat pozornost zákazníka a uspokojit jeho potřeby týkající se problému pro jehož řešení byl chatbot vytvořen. Nevýhodou je omezenost odpovědí, chatbot není schopen sám generovat nový text, a tudíž ve výsledku nevykazuje známky inteligence jako takové.
- **Generative-based model** – na rozdíl od první metody, zde dochází ke generování nových odpovědí na zadané otázky. Generativní modely využívají techniky strojového překladu, ale místo překladu z jazyka do jazyka dochází k „překladu“ ze vstupu do výstupu. Odezva poté působí naprosto přirozeně a chatbot není omezen pravidly, protože se samostatně učí na základě vlastní zkušenosti. Nevýhodou ovšem je obtížnost jeho kvalitní implementace. V tomto případě může docházet k chybovým reakcím, a to gramatickým i věcným. Vývojář nemá plnou kontrolu nad budoucím chováním chatbota a při jeho vývoji je vyžadováno velké množství trénovacích dat.

V dnešní době se používá hlavně první metoda, která není tak složitá na implementaci a dosahuje spolehlivých výsledků. Retrieval-based chatboti jsou vhodní pro plnění konkrétního úkolu, proto jejich služeb společnosti využívají především pro nabídku a prodej svých produktů/služeb, nahrazování FAQ dokumentů a podobně.

Pro zjednodušení představy lze na obrázku 1 vidět schéma znázorňující v jakém spojení spolu navzájem fungují jednotlivě klasifikované typy chatbota. Například za nejjednodušší je považována implementace chatbota založeného na vyhledávání podle pravidel, což je technologie fungující na uzavřené doméně. Na druhé straně stojí obecné AI, které je prozatím výzvou vývojářů.



Obrázek 1: Konverzační rámec chatbota
 Zdroj: Kojouharov (2016)

V kontextu s výše uvedenou klasifikací je zde vysvětlen jeden konkrétní způsob jednoduché implementace chatbota, a to za pomoci tzv. QnA služby, někdy též Q&A, odvozené z anglického názvu Questions and Answers. Příkladem je cloudová služba nazvaná QnA Maker od společnosti Microsoft, která ji za jistých podmínek a do určité míry nabízí zdarma (Microsoft, 2020). Také je umožněn bezplatný přístup k potřebné dokumentaci, ve které Microsoft (2019) zmiňuje typy aplikací, pro něž byl

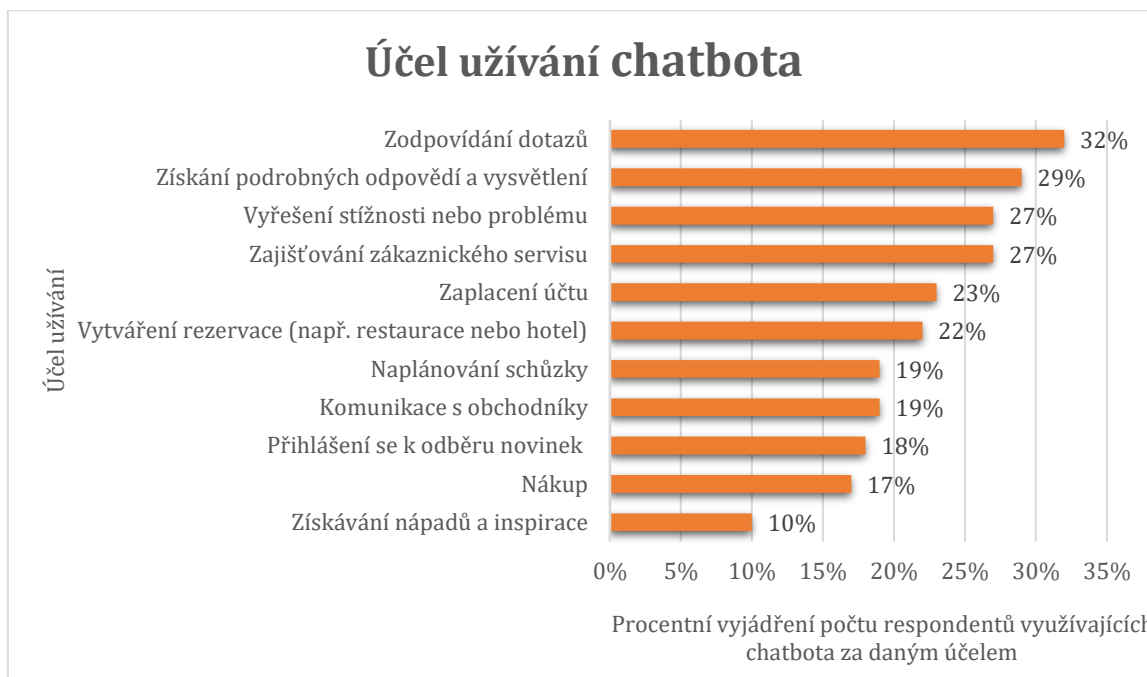
QnA Maker vytvořen a mezi které mimo jiné patří i chatbot. Dále uvádí, že ho lze použít v případě vlastnění statických informací ve znalostní bázi, zájmu společnosti filtrovat tyto informace založené na metadatech a vyžadování možnosti spravování konverzace chatbota (volba typu konverzace otázky a odpovědi či vícenásobné konverzace – vrácení odpovědi s výběrem z možností další uživatelské reakce a jiné). S touto službou se pojí znalostní báze naplněná sadou otázek a odpovědí, kdy při každém jejich importu dochází k vzájemnému provázání. Poté QnA Maker zpracovává každou otázku (vstup) a hledá co nejlepší odpověď (výstup). Důležité je ještě dodat, že služba poskytuje aktivní učení a vyhodnocuje spolehlivost (kvalitu) modelu, což znamená, že chatbot se tak stává přesnější a samostatnější.

Než se podnik pustí do vývoje chatbota a jeho znalostní báze, je potřeba si uvědomit požadované vstupy a výstupy, resp. je nutné přesné stanovení problematiky, kterou se bude zabývat. V reakci na to je potřeba vytvořit konverzační tok (conversational flow), což Sanjeevi (2018) vysvětluje jako tok navržený k přiměření spotřebitele konzumovat služby společnosti. Zjednodušeně se jedná o koncipování cesty, kterou se bude takticky ubírat daná konverzace. Počet takových flow se odvíjí podle počtu témat pro které je chatbot vyvíjen. Výběr konverzačního toku je závislý na zákazníkém zadaném inputu, ovšem jeho prostřednictvím se společnosti snaží nevědomky řídit chování poptávajících. Autor danou problematiku ilustruje na příkladu nově otevřené pizzerie, která využívá chatbota ze dvou hlavních důvodů: prodej pizzy a získání nových zákazníků. Pro tyto účely byly vytvořeny celkem tři scénáře, tedy konverzační toky, a to: okamžitá objednávka pizzy, stanovení ceny a zobrazení nejlepších nabídek a pizza trendů. Na základě těchto toků jsou zákazníci schopni zjistit veškeré potřebné informace ke koupi a zároveň získat povědomí o dalších skutečnostech, které mohou zapříčinit jejich opětovný zájem o produkty této pizzerie. Proto je potřeba si stanovit přesný cíl implementace chatbota, resp. jaké konkrétní výsledky jsou od něho očekávány. Na základě toho je podnik schopen zvolit i metriky podle kterých ho lze posuzovat.

3.4 Současné využití chatbota

Magazín Chatbots Life každý rok vydává přehled trendů a statistik a porovnává progres ve vývoji a užití chatbotů v podnicích. Statistiky zpracovává na základě dat,

které shromažďuje po celý předchozí rok a také vychází z údajů zveřejněných velkými společnostmi jako je například IBM, Facebook či Google. Na základě shromážděných dat poté přichází s různými prognózami ohledně budoucích trendů, implementace umělé inteligence, podílu firem zavádějících chatbota a další. Chatbots Life (2019) podle vlastního šetření, probíhajícího od srpna 2017 do února 2019, stanovilo žebříček TOP 5 států, sestavený na základě největší četnosti firem, které mají zavedeného chatbota. Do TOP 5 se dostala USA (36 %), Indie (11 %), Německo (4 %), Velká Británie (3 %), Brazílie (2 %), což znamená, že těchto pět velmocí tvoří celkem 56% podíl na trhu, a tudíž mají velkou převahu oproti ostatním zemím, ve kterých se společnosti snaží o úspěšnou implementaci této technologie, a to včetně České republiky. I proto se vývoj umělé inteligence stává nejvíce pokrokovým právě v USA a například v České republice se stále nasazují chatboti založení na pravidlech. Může za to i složitost českého jazyka, který je v porovnání s angličtinou pro chatbota velice těžké zpracovat. Společnost Drift ve spolupráci se Survey Monkey Audience vytvořili a vydali tzv. *2019 State of Conversational Marketing* (Drift a Survey Monkey Audience, 2019), který navazuje na *The 2018 State of Chatbots*, jehož tvůrci jsou obě již zmiňované společnosti, a to dále spolu se Salesforce a myclever (Drift, Survey Monkey Audience, Salesforce a myclever, 2018). Drift a Survey Monkey Audience (2019) uvádí, že do výzkumu pro rok 2019 bylo zapojeno přes 1000 respondentů a oproti roku 2018 byl tento report rozšířen o výzkumnou otázku, jak chatboti zapadají do širší konverzační strategie. Podle výsledků výzkumu zákazníci nejvíce u chatbota oceňují jeho fungování 24/7 a hned po komunikaci s člověkem tváří v tvář ho považují za nejrychleji reagující komunikační kanál. Dále bylo zkoumáno, pro jaké účely respondenti chatbota nejčastěji využívají, kdy každý z nich mohl zvolit více možností. Proto je potřeba na procentní vyjádření počtu respondentů u každého účelu užívání koukat zvlášť jako na procentní podíl z celkového počtu respondentů. Výsledky lze vidět na následujícím grafu 1.



Graf 1: Účel užívání chatbota

Zdroj: vlastní zpracování podle Drift a Survey Monkey Audience (2019)

Na základě grafu 1 lze vidět, že lidé chatbota nejčastěji využívají pro jednoduché úkony, jako je zodpovídání dotazů a dále získání podrobnějších odpovědí a vysvětlení, k čemuž je zapotřebí, aby měl chatbot vytvořenou rozsáhlou znalostní bázi. Dále skrz něj řeší své stížnosti či problémy a zajišťují zákaznický servis. Naopak nejméně ho využívají pro čerpání nápadů a inspirace.

Shevat (2017) uvádí výše zmíněné, ale i další oblasti, ve kterých chatbot nachází své upotřebení. Jedná se tedy o oblasti obchodu, zákaznického servisu, business neboli podnikového prostředí, osobních asistentů, integračních a notifikačních botů či herního průmyslu. Autor dále vysvětluje, jakým způsobem je chatbot v daném prostředí využit.

3.4.1 Obchod

V oblasti obchodu, resp. elektronického obchodování (e-commerce) je úkolem chatbotů usnadnit zákazníkovi proces nakupování a nahradit/zjednodušit práci zaměstnanců pracujících jako podpora. V tomto případě dokáže chatbot nejen efektivně poradit, specifikovat a odkázat, ale dokonce i vytvořit a dokončit objednávku daného zákazníka. Cui, Huang et al. (2017) uvádějí jako příklad chatbota internetového obchodu Amazon, jehož webové stránky obsahují velké množství detailních informací, které je vhodné distribuovat prostřednictvím chatovacího rozhraní.

3.4.2 Zákaznický servis

Tento typ chatbota má úzkou vazbu s chatbotem vytvořeným pro e-commerce služby. Proto se v dnešní době jedná o jednu z jeho nejčastěji implementovaných podob. Shevat (2017) ho nazývá též jako FAQ bota, což napovídá tomu, že je vyvíjen právě z důvodu poskytování informací (odpovědí na časté dotazy) externím zákazníkům a nahrazuje tak živé operátory pracující v kontaktních centrech. V současné době je nejčastěji využívána kombinace chatbota a zákaznické podpory, na kterou chatbot odkazuje v případě, že nedokáže plně vyhovět požadavku klienta. Na základě výzkumu provedeného Facebook IQ (2018) bylo zjištěno, že podle nadpoloviční většiny zúčastněných, užívání chatbota přináší lepší poradenství a péči o zákazníka v porovnání s tradičními komunikačními kanály.

3.4.3 Využití uvnitř podniku

Do této oblasti spadají chatboti působící pouze interně. Shevat (2017) do ní řadí například řízení lidských zdrojů, kde chatboti pomáhají s plánováním dovolené, otázkami týkajícími se odměňování či dalšími rutinními činnostmi, které nezbytně nevyžadují kontakt s živým personalistou. Dále jsou chatboti využíváni jako obsluha pro CRM a další podnikové systémy. Lze sem zařadit bota nazvaného Nikabot, pracujícího pro chatovací platformu Slack, který aktivně pomáhá projektovému týmu s plánováním, shromažďováním informací, vytvářením reportů, přerozdělováním úkolů a kontrolou chodu projektů.

3.4.4 Osobní asistenti

Tento typ chatbota je vytvářen pro zlehčování uživatelova každodenního života. Shevat (2017) uvádí jako příklad bota v roli osobního trenéra, který podporuje a pomáhá s hubnutím a udržováním zdravého životního stylu. Dalšími známými osobními asistenty jsou například správci financí, pomocníci při cestování či virtuální sekretářka. Zvláštním typem inteligentního osobního asistenta jsou technologie integrované do mobilních zařízení, kterými jsou Siri, S Voice, Google Assistant, Cortana či Alexa sloužící k ovládnání chytré domácnosti, kteří jsou na rozdíl od klasického chatovacího robota ovládnáni hlasem.

3.4.5 Integrovaní boti

Shevat (2017) integraci chatbota vysvětluje jako propojení se službami třetích stran, kdy uvádí jako příklad integraci Salesforce CRM do platformy Slack. Velice prospěšné je to, že uživatel může hledat informace o konkrétním zákazníkovi, přímo v konverzaci se svým kolegou.

3.4.6 Notifikační boti

Tento druh chatbota pravidelně uživateli zasílá notifikace týkající se novinek ze světa financí, sportu, počasí, nemovitostí a dalších problematik, které ho zajímají (Shevat, 2017). Do této kategorie spadá například NBC News bot, který vystupuje na platformě Facebook Messenger a informuje o novinkách podle uživatelem nastavených priorit.

3.4.7 Hry

V tomto případě se jedná o zábavní chatboty, kteří při různých hrách (např. kvíz) tvoří protivníka živé osobě. Dále Shevat (2017) zmiňuje hlasem ovládané boty, kteří jsou schopni například odvyprávět pohádku.

Chatovací robot se stává čím dál využívanější z hlediska podpory managementu všech typů podniku, ovšem svou podstatnou roli hraje i ve světě marketingu. Z toho důvodu hodně firem pro implementaci svého chatbota využívá umístění na platformě Facebook Messenger, WhatsApp Messenger a jiné.

Nutné je podotknout, že do klasifikace chatbotů často nebývají zahrnováni hlasoví asistenti jako jsou už zmiňovaní Siri, Google Assistant a další, protože se nejedná přímo o chat, ale o hlasovou interakci.

4 Proces měření efektivity chatbota

Chatbot své uplatnění nachází především v několika posledních letech, přesněji největší přelom jeho hojného nasazování přišel v roce 2016, kdy společnosti Facebook a Microsoft využití botů začaly oficiálně podporovat v rámci jejich platforem (Khorozov, 2017). Business Insider Intelligence (2016) prezentuje výsledky šetření provedeného společností Oracle, podle kterého v roce 2016 přibližně 80 % dotazovaných amerických firem vlastnilo nebo plánovalo chatbota zavést nejpozději do roku 2020. Nicastro (2018) uvádí, že společnost Inbenta na základě svého průzkumu využití a vnímání chatovacího robota z pohledu spotřebitelů a podniků vyhodnotila 50% preferenci komunikace s chatbotem při nakupování online oproti volání na zákaznickou podporu, přičemž 72 % nakupujících shledává jeho služby za bezchybné a velmi nápomocné. V České republice se několika případy úspěšné implementace chatbota do firemního prostředí pyšní agentura Feedyou, která se od roku 2014 snaží ukázat, jaký může mít tato technologie přínos pro dnešní podniky. V oblastech jako je HR, zákaznická podpora, prodej či dokonce GDPR pomohla společnostem jako je ČEZ, a. s., STRV, s. r. o., Fincentrum, a. s., Knorr-Bremse, s. r. o. a další (Feedyou, 2019).

Aby bylo možné reálně považovat zavedení technologie za úspěšné, respektive účinným způsobem splňující svůj účel (nejen z důvodu, že se jedná o trend), je potřeba její efektivnost pravidelně měřit. V tomto ohledu nebyl prozatím vytvořen jednotný přístup, podle kterého by firmy mohly svého chatbota posuzovat, což je předmětem zkoumání této diplomové práce. Tato kapitola se věnuje efektivnosti IS/ICT obecně a dále podrobně představuje přístupy k měření efektivnosti chatbota, které vycházejí ze zahraničních studií zabývajících se touto problematikou.

4.1 Efektivnost

4.1.1 Účinnost a účelnost

Vzhledem k tomu, že často dochází ke ztotožňování či zaměňování pojmů „Efficiency“ a „Effectiveness“, je potřeba vysvětlit jejich význam a vzájemný rozdíl. Řešení tohoto problému by mohl zjednodušit známý citát od P. F. Druckera (1992):

“Efficiency is doing things right; effectiveness is doing the right things.”, který lze přeložit jako *„Účinnost je děláním věcí správně; účelnost je děláním správných věcí“* (překlad vlastní).

Podobně se k tomuto tématu vyjadřují Učeň a kol. (2001), kteří toto stanovisko transformovali přímo na problematiku měření. Podle nich je potřeba klást důraz na účelnost, resp. měřit správné věci a až v dalším kroku tzv. měřit věci správně, a to za pomoci vhodně zvolené metody.

Jak ze zmíněného citátu vyplývá, tak „Efficiency“ představuje účinnost, přesněji účinnost vložených prostředků a jimi získaný užitek a jinými slovy také produktivitu, kterou lze charakterizovat jako poměr výstupů a vstupů nějaké činnosti či systému, což Mohelská a Pitra (2012) z hlediska managementu organizace popisují jako hodnocení úrovně podnikatelské výkonnosti. Podrobněji tento jev vysvětlují jako *„procentuální podíl ročního provozního zisku na celkovém kapitálu organizace – o kolik (haléřů) se zhodnotila jedna investovaná koruna za rok díky výnosům z provedených podnikatelských aktivit“*.

Na druhé straně stojí anglické „Effectiveness“, které podle Molnár (2000) představuje účelnost a je používáno ve spojitosti s posuzováním schopnosti vytvářet/uspokojovat požadovaný efekt/užitek. Mohelská a Pitra (2012) uvádějí, že účelnost je hodnocena na základě míry naplnění strategických záměrů. Proto je při měření efektivnosti vycházeno především z cílů, které se vztahují k monitorovanému procesu, službě či produktu. Učeň a kol. (2001) dále zmiňují, že je účelnost vyhodnocována v rámci plnění procesních auditů, katalogů služeb či nastavení systémů a pravidel vztahů pro měření výkonnosti podniku, kdy se například jedná o známou metodu Balanced Scorecard. Bloom (2016) také povahu účelnosti považuje za strategickou, což plyne především z otázek, které si klade. Jedná se například o:

- „Jaký je obchodní cíl?“,
- „Jaký je nejlepší možný výsledek?“.

4.1.2 Efektivnost IS/ICT

Molnár (2000) z hlediska efektivnosti informačních systémů (dále jen „IS“) a informačních technologií (dále jen „IT“), resp. informačních a komunikačních technologií (dále jen „ICT“) uvádí, že má smysl zkoumat pouze takové systémy,

u kterých je možné definovat účel. Ty poté nazývá jako systémy s cílovým chováním. Dále také vysvětluje vztah mezi informačním systémem a informační technologií, kdy informační systém reprezentuje potřebu informací, zatímco informační technologie představuje uspokojení této potřeby a odtud pramení zkratka IS/IT, resp. ICT.

Dále autor dodává, že hodnocení efektivnosti se nezabývá pouze otázkou potřeb a jejich naplňováním, ale také očekáváním zainteresovaných stran. Z celopodnikového hlediska se může jednat o:

- majitele, kteří v zavedení IS/ICT vidí trvalé zhodnocování majetku vloženého do společnosti,
- manažery, kteří jsou na základě IS/ICT schopni efektivně řídit podnik, a to bez vynaložení nadbytečných zdrojů, které jim byly svěřeny do správy,
- zaměstnance, kterým IS/ICT nabízí výhody v podobě lepšího, výkonnějšího a zcela integrovaného pracovního prostředí,
- zákazníky, kterým využití IS/ICT přináší vyšší přidanou hodnotu požadovaného produktu či služby.

Molnár (2000) podle Dolanský, Němec a Měkota (1996) definuje efektivnost jako *„účinnost prostředků vložených do nějaké činnosti hodnocenou z hlediska užitečného výsledku této činnosti“*. Za vložené prostředky jsou považovány výdaje (investice) do IS/ICT, jejichž účinnost je posuzována na základě přínosů z IS/ICT. Proto je z hlediska zjišťování efektivnosti kladen důraz na účel, resp. co by měl daný systém čili technologie přinést firmě, jinými slovy, jaká je jeho přidaná hodnota. Častým problémem bývá, že efekty (dosažení stanovených cílů) ze zavedení IS/ICT se nedostavují ihned, v některých případech to trvá i několik let, za což mohou například faktory jako je adaptace zaměstnanců na zavádění inovativních změn, špatně provedená analýza potřeby daného IS/ICT, unáhlená a nedostatečně podložená rozhodnutí, ztráta konkurenční výhody při zavádění globálně aplikovaných trendů. I proto autor uvádí podle vlastní zkušenosti v souladu s literaturou, že ve většině případů výsledky standardních finančních ukazatelů efektivnosti (ROI, TCO, NPV) indikují na nevýhodnost investic do IS/ICT. Je tomu tak i z důvodu, že ne vždy jsou známé přesné hodnoty všech vstupů a výstupů, protože k výpočtu

jsou vyžadována podrobná číselná data, kdežto přínosy a náklady jsou složeny i z dat kvalitativní povahy a při jejich transformaci na tvrdé ukazatele může poté dojít k nesrovnalostem, resp. ke zkreslení jednoho ukazatele, který indikuje na opačný výsledek než ukazatel druhý (Švík, 2009). Čech a Bureš (2009) se zmiňují o tzv. paradoxu produktivity a řetězovém efektu investic do IS/ICT, což dále vysvětlují jako nepodílení se IS/ICT na zlepšení hospodářského výsledku podniku, kdy s každou novou investicí se váže potřeba další a další. Podle Molnár (2000) je tedy potřeba se zaměřit na metriky, které jsou určeny v reakci na stanovený cíl a nemá smysl se zabývat absolutními částkami, které samy o sobě mají velmi nízkou vypovídací schopnost. Efektivnost také ovlivňuje hledisko pořizování IS/ICT, resp. jakou formou byla daná technologie do společnosti zaváděna – bylo využito interního vývoje či outsourcingu? Oba způsoby mají své výhody i nevýhody a nelze obecně určit, který z nich je lepší, vždy záleží na konkrétním případě a na dalších skutečnostech, jako je velikost a hodnota firmy, co upřednostňuje vedení a tak dále.

Pokud chce společnost zavádět efektivní IS/ICT, musí mít jasnou informační strategii, kterou je podle Molnár (2000) myšlena soustava cílů a způsob jejich dosažení. Pomocí takové informační strategie by měla být organizace schopna najít odpovědi na otázky, jak pomáhá daný IS/ICT podniku zvyšovat výkonnost pracovníků, podporovat dosahování strategických cílů, získávat pro něj konkurenční výhodu či vytvářet další strategické příležitosti rozvoje.

Autor uvádí, že pro sledování a vyhodnocování efektivnosti daného IS/ICT je stěžejní, aby při vytváření informační strategie podniku, tzn. v etapě plánování, byly:

- jasně definované cíle, kterých má být daným projektem (aplikací IS/ICT) dosaženo,
- uvažovány všechny jeho přínosy,
- určeny ukazatele (metriky) k dosažení stanovených cílů,
- určen zodpovědný manažer,
- stanoven systém časového a organizačního sledování a vyhodnocování dosahování stanovených cílů,
- všechna očekávání, včetně příslušných organizačních opatření, komunikována se všemi pracovníky dané organizace.

Dále je dodáno, že by se společnosti měly soustředit na informace nefinanční povahy, které jsou orientované na budoucnost, resp. jsou založeny na realistických prognózách a neměly by tudíž vycházet pouze z minulých zkušeností.

V návaznosti na výše uvedené je potřeba vymezit pojem metrika, který Učeň a kol. (2001) definují jako přesně vymezený způsob, nástroj a rozsah měření, který se používá k hodnocení úrovně efektivnosti konkrétní oblasti, procesu či projektu – seskupení ukazatelů za určitým cílem se poté nazývá „portfolio metrik“. Každou z nich by měly určovat atributy, kterými podle autorů jsou:

- název a identifikace,
- algoritmus (vzorec) – týká se tvrdých metrik,
- definice – týká se měkkých metrik,
- vlastník,
- dimenze (měrná jednotka, organizační jednotka, časové období atd.),
- výchozí a cílová (chtěná) hodnota,
- zdroj dat pro měření,
- postup, způsob, periodičita, harmonogram, odpovědnost, vykazování měření a jeho verifikace.

Aby bylo dosaženo celkové efektivnosti IS/ICT, je potřeba si odpovědět na otázku co se od daného systému či technologie očekává, resp. hned na začátku životního cyklu IS/ICT je nutné definovat příslušné ukazatele přínosů, stanovit jejich vyhodnocování i konkrétní odpovědnosti za jednotlivé výsledné hodnoty. Molnár (2000) takové ukazatele klasifikuje na:

- finanční (měřené v peněžních jednotkách) a nefinanční (měřené v jednotkách času, počtech kusů atd.),
- kvantitativní (měřeno na kardinální stupnici – intervalové a poměrové proměnné) a kvalitativní (měřeno na ordinální stupnici (klasifikační, bodovací) či pomocí nominálních proměnných) (Záhora, 2015),
- přímé (lze jednoznačně prokázat příčinu dosaženého přínosu) a nepřímé (je zapotřebí stanovit zástupné ukazatele vyjadřující změnu),

- krátkodobé (indikují na přínosy projevující se do půl roku od zavedení IS/ICT) a dlouhodobé (dosažení přínosů až v rozmezí několika let),
- absolutní (vyjádřené na základě nějaké měřitelné hodnoty) a relativní (vyjádřené poměrovým číslem).

Dále autor podotýká, že volba a kombinace ukazatelů záleží vždy na konkrétním případě a zdůrazňuje, že nejrelevantnějším vztahem pro hodnocení efektivnosti je sledování hlediska účelnosti, které je přímo měřitelné mírou dosažení cílů, a to tedy jako:

$$\text{Účelnost} = \frac{\text{dosažená hodnota cíle}}{\text{plánovaná hodnota cíle}}$$

Co se týká ukazatelů nefinanční měřitelné povahy, tak mezi ně lze zařadit produktivitu, která jak už bylo uvedeno v podkapitole 4.1.1 značí poměr mezi vytvořenými výstupy a vynaloženými vstupy za určitou časovou jednotku. Pro výpočet této metriky je zapotřebí správně definovat čítec a jmenovatel a zároveň tyto hodnoty systematicky sledovat a vyhodnocovat. Výsledkem by mělo být zvyšování produktivity, resp. zvyšování výstupů a snižování vstupů (např. zvyšování počtu obslužených zákazníků jedním pracovníkem/strojem za den). Proces posuzování produktivity je dlouhodobou záležitostí, je potřeba brát v potaz sezónní vlivy, a tudíž nelze komparovat navzájem nesouvisející časové úseky. Dalšími ukazateli tohoto charakteru jsou například vyšší počet zákazníků, kratší doba obsluhy zákazníka, nižší počet reklamací. Ty lze po zevrubnější analýze vyjádřit finančně.

V rámci kvantitativních a kvalitativních ukazatelů, autor rozlišuje tzv. „měkké“ a „tvrdé“, kdy „tvrdý“ ukazatel působí jako zástupný, jehož změna odráží změnu zmiňovaného „měkkého“. Podle Učeň a kol. (2001) mezi tvrdé metriky patří např. zkrácení průběžné doby vývoje a výroby, snížení počtu reklamací, zvýšení počtu zákazníků, snížení doby prostoje výrobního zařízení, zkrácení doby obsluhy zákazníka. Mezi „měkké“ metriky zařazuje např. zlepšení dobrého jména podniku (aplikování různých průzkumů), spokojenost zákazníků (též existují různé typy průzkumů – dotazníky, ankety, recenze, lze odhadnout ale i podle počtu stálých zákazníků, celkového růstu počtu zákazníků atd.), přizpůsobení se trhu (rychlá reakce podniku na nově vzniklé trendy), zlepšení pracovního prostředí (ankety hodnotící

zaměstnaneckou spokojenost), přidaná hodnota produktu či služby (zvýšení počtu zákazníků, vyšší spokojenost zákazníků a jejich ochota zaplatit více – vždy záleží na charakteru produktu/služby).

Molnár (2000) dále přisuzuje velký význam vyhodnocování kvality produktu/služby IS/ICT, kterou považuje do jisté míry za rovnocennou pojmu efektivnost. To je však pochopitelné, protože nekvalitní produkt/služba neplní požadovanou funkci a tím pádem není dosaženo zamýšleného účelu/efektu. K posouzení kvality je tedy navrženo sledovat následující parametry:

- funkčnost (např. přesnost, spolehlivost, robustnost a předvídatelnost chování daného IS/ICT, a dále škálovatelnost, propojitelnost a ověřitelnost jeho vlastností),
- vzhled, komfortní využívání (např. příjemné působení, flexibilita, jednoduchost, intuitivnost, přehlednost, robustnost),
- spolehlivost (např. stálý výkon, dosažitelnost a neustálá použitelnost aplikací, obnovitelnost, spolehlivá kvalita informací),
- provoz a údržba (např. úroveň podpory pro průběžné monitorování provozních charakteristik, řešení požadavků a incidentů, nároky na rutinní obsluhu, kvalita a úplnost dokumentace, možnosti testování)
- dostupnost (např. 24/7, možnost přístupu k datům),
- integrita a komplexnost,
- trvanlivost (např. dlouhá životnost, schopnost dalšího vývoje),
- bezpečnost (např. bezpečný a kontrolovatelný přístup a klasifikace a uchování dat).

Dále je dodáno, že kvalitu nelze změřit jako kvantitu, ale jedná se o souhrn vlastností, podle kterých je uživatelem usouzeno, zda je daný produkt/služba kvalitní. Kvalita IS/ICT bývá často vyhodnocována podle obsahu (zda jsou informace relevantní pro daný účel, dostatečně přesné, úplné, detailní, důvěryhodné) a formy prezentace informací (zda dochází k vhodnému předání srozumitelných informací správným osobám, a to včas, resp. v náležitý čas). Tyto i metriky umístěné v podbodech výše se zjišťují pomocí přímého měření (např. doba odezvy, počet výpadků), dotazníkového šetření u uživatelů, provedení interního či externího auditu výkonnosti a funkčnosti

IS/ICT. Pro standardizované řízení jakosti lze využít norem z řady ISO 9000 (ISO, 2019).

Zajištění kvality, resp. kvality služeb IS/ICT lze v podniku sledovat i pomocí Service Level Agreement (dále jen „SLA“), která definuje úroveň, rozsah a kvalitu té konkrétní služby. Overby, Greiner a Paul (2017) uvádí, že se jedná o smlouvu mezi zákazníkem (uživatelé) a dodavatelem (poskytovatelem), ve které jsou stanoveny přesné metriky, podle kterých je daná služba měřena a dále nápravná opatření (např. způsob a rychlost odstranění nestandardního stavu), sankce v případě, že ze strany dodavatele nedochází k udržení/dosažení sjednané úrovně, podpora služby, odpovědnosti, komunikační kanály a další ujednání. Autoři též dodávají, že obvykle se SLA uzavírá mezi společnostmi a externími dodavateli, ale může tak být učiněno i mezi dvěma odděleními v rámci jednoho podniku.

Molnár (2000) tedy shrnuje následující poznatky (postup), kterých by se měl podnik při snaze vyhodnotit efektivnost IS/ICT držet:

1. stanovit účel – definovat cíle, kterých má být zavedením IS/ICT do podniku dosaženo a určit příslušné metriky, na základě kterých bude prokázáno, zda bylo stanovených cílů dosaženo nebo k nim alespoň přiblíženo,
2. uvažovat nad přínosy a vytvářením nových podnikatelských příležitostí,
3. uvažovat nad všemi možnými příjemci užitku z IS/ICT a jejich očekáváním,
4. definovat systém hodnocení dosahování cílů v časové (kdy a jak často) a personální (osoby zodpovědné za plnění, resp. dosahování cílů) rovině, brát na vědomí délku časového horizontu vedoucího k tomuto cíli (až v řádech let) a uvažovat všechny relevantní ukazatele.

4.2 Přístupy k měření efektivnosti chatbota

Jak už bylo zmíněno výše, měření efektivnosti chatbota postrádá jednotný přístup, což vede k nevědomosti firem zavádějících tuto technologii, jaké ukazatele je důležité sledovat a vycházejí tak z obecných standardů, které nejsou pro tento případ dostatečně konkrétní. Autoři studií, kteří se touto problematikou již v minulosti zabývali, se svými přístupy odlišují a důležitost přikládají různým oblastem.

Nicastro (2018) na základě rozhovorů s technickými a výkonnými řediteli ze společností Ameex Technologies a Inbenta uvádí, že i v případě chatovacího robota lze použít některou z tradičních metrik, využívaných při posuzování výkonnosti a účelnosti technologií obecně. V tomto případě je myšleno stanovení tzv. KPI. Tento pojem je akronymem z anglického Key Performance Indicators, v překladu klíčové indikátory výkonnosti. Pomocí těchto ukazatelů lze měřit celou organizaci, jednotlivé útvary, pracovníka, ale také pouhý proces či službu. Učeň a kol. (2001) uvádí, že KPI lze použít na všech úrovních řízení a upřesňují, že „*jsou odvozeny z dekomponovaných cílů jednotlivých procesů (resp. funkčních oblastí) a kritických faktorů úspěchu podnikové strategie*“. Dále je zmiňováno, že v užším pojetí tyto ukazatele slouží jako doplněk pro tzv. KGI (Key Goal Indicators), které jsou přímo orientované na dosahování procesních cílů. Oba tyto případy by měly splňovat podmínky SMART, což je analytická technika sloužící pro navrhování specifických cílů, jejíž název je složený z anglických adjektiv, kterými jsou Specific, Measurable, Acceptable/Achievable, Relevant/Realistic, Time Specific/Trackable, podle Sokolová a kol. (2015) znamenající v překladu specifický, měřitelný, akceptovatelný, relevantní a trvajících. Například cíl by měl být dostatečně konkrétní, aby byl měřitelný a mohl tak tedy vzniknout důkaz o jeho dosažení. Nicastro (2018) navrhuje konkrétní ukazatele, které by měly být schopny správně posoudit výkonnost a efektivnost chatovacího robota. Jedná se o kvantitativní i kvalitativní metriky vhodné pro hodnocení chatbota orientovaného na zákazníky, ovšem po přizpůsobení se konkrétnímu cíli jsou svým obecným charakterem použitelné i pro ostatní typy. Autor dané ukazatele nijak systematicky neseřadil a soustředil se pouze na jejich výčet. Jedná se tedy o následující:

- celková hodnota prodeje,
- míra konverze – podíl počtu provedených konverzí (neboli akce návštěvníka splňující cíl pro který byly vytvořeny webové stránky, kdy v případě chatbota se jedná například o prodej zboží/ produktu/služby, sjednání či rezervace služby, odeslání formuláře, zanechání telefonního čísla) a počtu návštěv/zahájených interakcí za sledované období,
- podpora a úspora času zákazníka,
- zvyšování NPS (Net Promoter Score – míra spokojenosti/loajality zákazníků) – metrika založená na měření spokojenosti/loajality zákazníků prostřednictvím

odpovědi na jednoduchou otázku: „*Jak pravděpodobné je, že byste doporučili (organizaci/produkt/službu) příteli nebo kolegovi?*“, kdy jsou poté respondenti rozděleni, na škále od 0 do 10, na promotéry, pasivní, kritiky a konečný výsledek NPS je dán rozdílem procentuálního zastoupení promotérů a kritiků (Qualtrics, 2019),

- náklady na pořízení,
- náklady na provoz a údržbu,
- počet nových uživatelů vytvořených denně/týdně/měsíčně,
- počet aktivních uživatelů,
- celkový a průměrný počet zahájených interakcí denně/týdně/měsíčně,
- průměrný denní počet dokončených návštěv, resp. s výsledkem poskytnutí reálné služby,
- samostatnost (do jaké míry, resp. kolik případů je chatbot schopen úspěšně sám vyřešit, bez toho, aniž by musel být jeho úkol dořešen lidským činitelem),
- míra spokojenosti uživatele (dotazníkové šetření, anketa, bodovací (hvězdičkový) systém),
- počet řešených neočekávaných situací (chatbot se může dostat do situace, kdy neví co odpovědět a je potřeba mít efektivní způsob, jak zjistit kdy, jak a proč se to stalo),
- sledování učení chatbota v průběhu času,
- sledování chování (reakcí) uživatelů skrz analýzu sentimentu, na jejímž základě lze pochopit, jak se uživatel cítí ohledně průběhu konverzace vzhledem k relevantnosti obsahu.

Peras (2018) ve své studii zmiňuje, že téma týkající se hodnocení chatbota je velmi často zanedbáváno, respektive metriky, které by k tomu měly být použity nejsou systematizované ani sjednocené. Tento problém se snažila vyřešit pomocí spojení již existujících metrik, které získala na základě rozboru patnácti odborných článků zabývajících se touto tematikou a ty poté rozdělila do pěti perspektiv, kterými jsou:

- perspektiva uživatelské přívětivosti,
- perspektiva vyhledávání (generování) informací,
- lingvistická perspektiva,
- technologická perspektiva,

- obchodní (business) perspektiva.

Peras (2018) perspektivu uživatelské přívětivosti rozděluje do čtyř kategorií, kterými jsou: použitelnost, výkon, působení (vliv) a spokojenost. Použitelnost měří na základě intuitivnosti chatovacího rozhraní a času (uživatelova úsilí) potřebného k dokončení úkolu. Námaha vynakládaná na práci v daném prostředí je ovlivňována uživatelskými znalostmi a zkušenostmi, což znamená, že v porovnání s druhou a další návštěvou je při první návštěvě, úplně neznámého prostředí, vždy vyžadována větší soustředěnost a snaha ze strany dané osoby. Výkonem se rozumí dokončení úkolu z hlediska úplnosti, pohotovosti a vhodnosti, kdy výsledek je ovlivňován charakterem, resp. náročností jeho zadání. Kategorie působení, jinými slovy vlivu je spojená s posouzením uživatelské zkušenosti a percepce daného systému. Spokojenost je vyjádřena potěšením vyplývajícím ze srovnání uživatelského očekávání a reálného výkonu chatbota. Autorka uvádí, že v některých dalších studiích je tento atribut hodnocen na základě vztahu délky a četnosti chatu, respektive delší a častější interakce značí vyšší spokojenost s obsluhou chatbota. Vliv a spokojenost je obtížné numericky vyjádřit pomocí reportů, a proto se často zkoumají prostřednictvím dotazníkového šetření a anket.

Perspektivu vyhledávání (generování) informací autorka charakterizuje jako schopnost společnosti naplnit chatbota informacemi, které uspokojí potřeby uživatele. Toto hledisko se skládá z přesnosti, srozumitelnosti, účinnosti a je založeno na hodnocení transferu relevantních a vhodných informací ze znalostní báze chatbota. Dané metriky jsou kvantitativní povahy a posuzují úroveň přesnosti výstupu vzhledem k danému vstupu, a to bez vynaložení nadměrných zdrojů.

Lingvistická perspektiva analyzuje úroveň jazykové přesnosti chatbota a schopnost vytvářet adekvátní odpověď. V tomto případě je zkoumána slovní zásoba, pravopis a gramatická úroveň výstupů chatbota, a to prostřednictvím vyhodnocení pěti kategorií, kterými jsou: kvantita, kvalita, relace, způsob a gramatická přesnost. První čtyři aspekty jsou založené na teorii tzv. konverzačních maxim vytvořených britským filozofem H. P. Grice. V tomto pojetí podle Grice (1975) kvantita představuje množství informací, které mají být poskytnuty a v rámci nichž definuje dvě maxima říkající, že by informace měla poskytovat to co je od ní v daný okamžik

žádáno a zároveň by neměla dávat více než je od ní očekáváno. Kvalitou informací je myšleno sdělovat pouze to co je pravdivé a k čemu existují dostatečné důkazy. Kategorie relevance obsahuje pouze jedno maximum, a to důležitost informace, což je podle autora v některých případech těžké odhadnout. Posledním aspektem je způsob jednání, který hovoří o vyvarování se nejasnostem výrazu, dvojznačnosti a ukládá důležitost stručnosti a uspořádanosti. Peras (2018) uvádí, že tyto čtyři kategorie lze hodnotit pomocí Likertovy škály. V poslední řadě autorka přidává pátou metriku, která znázorňuje přijatelnost chatbota z gramatického a významového hlediska. Ke zhodnocení této perspektivy je potřebný tým expertů.

Technologická perspektiva obsahuje pouze jeden atribut, a to lidskost. Tato metrika je posuzována na základě chování chatbota, který by měl vystupovat jako člověk. To, aby bylo možné, musí být chatbot schopen nejen zpracovávat přirozený jazyk, ale také mu porozumět, generovat a samostatně se učit. Nejčastějším způsobem, jak zhodnotit lidské chování této technologie je Turingův test, o němž bylo zmiňováno v podkapitolách 3.1 a 3.2. Problémem je, že Turingovým testem dosud žádný chatbot neprošel, a tak tato metrika z hlediska efektivnosti komerčních botů ztrácí svůj smysl. Dále autorka navrhuje využití alternativních možností testování lidskosti za použití hodnotící škály a procentního vyjádření míry (ne)porozumění chatbota.

Jako poslední byla brána v úvahu obchodní neboli business perspektiva, která zkoumá, jakou přidanou hodnotu chatbot přináší pro společnost. Peras (2018) uvádí, že tato hodnota se vypočítá jako rozdíl mezi účelností (produkcí) a náklady na chatbota. Pro měření jeho účelnosti je navrhováno se zaměřit na počet uživatelů, dobu trvání konverzace mezi uživatelem a chatbotem a počet takových konverzací. Tyto metriky by se měly poté porovnávat s počtem účastníků konverzace (například, zda konverzace probíhala pouze mezi chatbotem a zákazníkem či bylo zapotřebí aktivity dalšího lidského činitele), délkou interakce mezi lidskými účastníky (např. mezi zákazníkem a obsluhou help linky), počtem neúspěšných rozhovorů, nevhodných odpovědí a opakovaných dotazů. Chatovací roboti jsou nejčastěji implementováni pro plnění snadných a rutinních úkolů, například zodpovídají často kladené otázky (FAQ), což vede k úspoře zdrojů společnosti.

V poslední řadě Peras (2018) dodává, že využití konkrétních perspektiv a kombinace metrik, záleží na analyzovaném typu chatbota a není tedy vždy potřeba aplikovat všech pět perspektiv. Jedná se totiž o komplexní balíček metrik pro zjišťování efektivnosti chatbota obecně, a proto je potřeba vyhodnotit, co je pro daný výzkum relevantní a naopak.

Maroengsit et al. (2019) berou v úvahu podobné oblasti jako Peras (2018), akorát v menší míře a navrhují jiné způsoby jejich implementace. Autoři chatbota měří pomocí tří hlavních kategorií, kterými jsou hodnocení obsahu, spokojenost uživatelů, funkční aspekty.

Hodnocení obsahu je zaměřeno na kontext odezvy, kterou chatbot vytváří. V tomto případě byla tato kategorie posuzována ze dvou pohledů, a to jako automatické a expertní hodnocení.

V rámci automatického hodnocení je považována za nejdůležitější přesnost a schopnost vytváření výstupů. Přesností je myšleno množství konverzací, ve kterých jsou poskytovány informace relevantní pro diskutované téma. Schopnost vytváření je posuzována podle odezvy, jejíž obsah chatbot označuje za vhodný. Výstupy obou těchto metrik jsou vyjádřeny v procentech a jsou v podstatě základem pro metody jako je BLEU (Bilingual Evaluation Understudy), kdy se jedná o algoritmus vyhodnocující kvalitu textu (překladu), který byl za pomoci strojového překladu převeden z jednoho přirozeného jazyka do druhého a ROUGE (Recall-Oriented Understudy for Gisting Evaluation), což je sada metrik a softwarový balíček vyhodnocující automatické určení kvality souhrnu slov vytvořených systémem ve srovnání s člověkem. BLEU hodnotí výstup v rozsahu od 0 do 1, kdy strojový překlad identický s lidským (kvalitním) obdrží skóre 1, ale podle Maroengsit et al. (2019) jen málo překladů dosáhne takového výsledku. Dále lze jednodušeji přesnost hodnotit pomocí komparace dvou modelů se stejným významem, a tedy schopností vrátet přibližně stejné odpovědi.

Druhou cestu pro posouzení přesnosti a vytváření odezvy nabízí odborné či expertní hodnocení. Qui et al. (2017) jejich chatbota vyhodnotili pomocí business analytiků. Ti rozhodovali, zda je odpověď nevhodná, neutrální či vhodná, k čemuž použili sadu 600 otázek. Pro zjištění účinnosti tohoto přístupu vytvořili vzorec,

na základě kterého poté určili nejvyšší procento přesnosti nazvané P_{top1} . Jedná se tedy o následující vzorec: $P_{top1} = \frac{(N_{suitable} + N_{neutral})}{N_{total}}$.

Dále se Maroengsit et al. (2019) zabývají měřením spokojenosti uživatelů. Tu navrhují zkoumat za pomoci Likertovy škály na dvou úrovních: relace a obratu.

Při měření úrovně relace, jsou uživatelé požádáni o hodnocení celé relace (průběhu) chatu. Autoři například zmiňují chatbota, který byl vyhodnocen na základě dotazníkového šetření, v němž se posuzovalo celkem šest faktorů, a to: pozornost, pozitivnost, vzhled, pocit z jednání chatbota, vnímání, zveřejňování informací. Xu et al. (2017) tyto aspekty zužují pouze na vhodnost, empatii a vstřícnost.

Na rozdíl od hodnocení celé relace chatu, jsou u úrovně obratu respondenti požádáni o posouzení každé odpovědi vytvořené chatbotem. V tomto případě se hledá odpověď na otázku, zda chatbot dokáže svými výstupy uspokojit uživatelova očekávání. Tento aspekt lze zkoumat i za pomoci odborníků, kdy se například jedná o porovnání odpovědí na sadu otázek u dvou chatbotů se stejným zaměřením. Výsledkem je poté procentuální vyjádření spokojenosti s jejich odezvou.

Jako poslední autoři zkoumali metody hodnocení, které zařadili do kategorie funkčních aspektů. Prvním je posouzení na základě cíle (úkolů), což je vhodné pro chatbota zaměřeného na cíl (úkol), o němž bylo diskutováno v podkapitole 3.3. Dané metriky zahrnují počet dokončených cílů (úkolů), průměrnou délku dialogu, průměrnou délku projevu uživatele apod. Dalším aspektem je sledování statistických údajů, které znázorňují využití daného chatbota. Pomocí těchto ukazatelů lze dobře zhodnotit fungování a celkový důvod (motivaci) pro jeho udržení ve firmě.

Radziwill a Benton (2017) se ve své studii zaměřují především na atributy hodnocení kvality chatbota, jejíž zajištění považují za stěžejní nejen z hlediska efektivnosti, ale i bezpečnosti (např. při boji proti sociálnímu inženýrství). Při sestavování frameworku bylo autory povšimnuto, že jednotlivé atributy jsou v souladu s konceptem použitelnosti podle normy ISO 9241, resp. ISO 9241-11. ISO (2018) v rámci této normy použitelnost nedefinuje jako atribut produktu, ale jako výsledek interakce se systémem, produktem či službou a považuje ji

za prostředek vedoucí k efektivnímu dosažení cílů. Proto i v návaznosti na to definuje tři hlavní pojmy:

- účelnost – přesnost a úplnost, s níž uživatelé dosahují stanovených cílů,
- účinnost – zdroje vynaložené ve vztahu k dosaženým výsledkům, mezi které typicky patří čas, lidské úsilí, náklady a materiál,
- spokojenost – reakce a míra uspokojení potřeby uživatele vyplývající z používání systému.

Radziwill a Benton (2017) dále tedy rámeček měření rozdělili právě do těchto tří kategorií, v rámci kterých posuzovali výkon, funkčnost, lidskost, působení, způsobilost chování a přístupnost.

Zajímavý je poznatek ohledně posuzování lidskosti, kterou výše zmínění autoři dosud navrhovali hodnotit pomocí projití či neprojítí Turingovým testem. Ovšem Radziwill a Benton (2017) na základě podnětu od Ramos (2017) poznamenávají, že výsledek tohoto testu nerozhoduje o kvalitě chatbota, protože lidé jsou si vědomi, že reálně nekomunikují s člověkem, a i tak jsou schopni si konverzaci užít a odnést si z ní informace naplňující jejich potřeby. I Russel a Norvig (2010) zmiňují přísnou kritiku od Shieber (1994) a Ford a Hayes (1995), kteří ho nepovažují za užitečný.

Jak už bylo uvedeno v podkapitole 4.1.2, tak řízení kvality je důležitou součástí zajištění celkové efektivnosti, a proto se atributy měření jmenované Radziwill a Benton (2017) téměř opakují s těmi již zmíněnými v předešlých přístupech, akorát v jiném seskupení. I Učeň a kol. (2001) uvádí, že k úspěšnému dosahování cílů vede neustálé, nejlépe kontinuální, zlepšování, kterým se zabývá řízení jakosti. Dále je dodáno, že tento soubor ukazatelů je pouze orientační a vždy je potřeba brát v úvahu, který z nich se hodí použít na daný typ chatbota, resp. jaká zjištění ohledně výsledků chatbota jsou pro danou společnost nejdůležitější.

Zajímavý přístup k hodnocení chatbota (přesněji pouze B2C chatbota) nabízí i Kuligowska (2015), ve kterém se zaměřila na metriky týkající se kvality, výkonu a použitelnosti. Ty podrobněji rozdělila do následujících deseti podoblastí:

- vizuální stránka chatbota,
- forma implementace na webové stránce,
- syntéza řeči (pouze u virtuálních asistentů),
- znalostní báze,
- prezentace znalostí a dodatečné funkcionality,
- konverzační a jazykové schopnosti a citlivost kontextu,
- osobnostní rysy,
- možnosti personalizace,
- reakce chatbota na neočekávané situace,
- možnost hodnocení (chatbota a webové stránky) uživatelem.

Vizuální stránka chatbota je posuzována podle toho, zda jeho „podobu“ zakomponovanou v chatovacím rozhraní představuje obrázek živé osoby (nejčastěji živého operátora), animovaná postavička či nic. Implementace na internetové stránce je hodnocena podle umístění chatovacího prostředí. Namísto vestavěného a plovoucího chatovacího okna, je preferována flexibilní kombinace vestavěného okna a výsuvné boční karty. Vybavenost znalostní báze je podle autorky základním předpokladem správného fungování chatbota a tvoří tak tedy jeho podstatu. Chatbot by měl být schopen konverzovat o obecných tématech, resp. reagovat na otázky typu „Jak se jmenuješ?“, „Jak se máš?“ a podobně a dále o specializovaných tématech, kvůli nimž byl vytvořen. Forma prezentace výstupů a dodatečných funkcionalit je posuzována na základě schopnosti chatbota nabízet přímé přesměrování na požadované odkazy otevírající se na nové kartě prohlížeče, konverzační strom provázející uživatele výběrem z možností podle přesného flow, tlačítka s volbou „zpět“ a uchovávání historie chatu, „pomoc“, „domů“ a vrácení na úvod chatu. Lingvistické dovednosti a kontextovou citlivost lze měřit na základě schopnosti vést soudržný dialog a porozumět složitým výrazům. Dále se studie zabývala osobnostními rysy chatbota, které vyjadřuje tím, že dokáže mluvit sám o sobě jako by byl člověk, například sdělí jeho věk, pohlaví, jaké má zájmy. Hodnocení možností personalizace zahrnuje měření schopností velmi inteligentního bota, který umožňuje změnu pohlaví, resp. vizuální podobu, dále je schopen si zapamatovat jméno uživatele, dovoluje prohlížet historii konverzace během chatu, vyhledávat ji zpětně, tisknout či přeposílat e-mailem a rozpoznat internetovou stránku,

kterou si uživatel právě prohlíží. Podle výsledků výzkumu této studie žádný z chatbotů není schopen plnit více než dvě z těchto pěti jmenovaných funkcí a lze je tedy zatím považovat za nadstavbu pro dnešní chatboty. Co se týká reakcí na neočekávané situace, tak těmi je myšleno chování bota při nezpůsobilém jednání ze strany uživatele, čímž jsou například myšleny vulgární a sexuální nadávky. Kuligowska (2015) uvádí, že podle Brahnem (2005) za neefektivní způsob, jak řešit takovou situaci, je útočit nazpět či ukončit konverzaci a znemožnit uživateli klást další otázky, protože je to neuctivé a může to poškodit image podniku. Dále do této podoblasti spadají uživatelem špatně formulované věty, překlepy a gramatické chyby na které musí být smysluplně reagováno. Jako poslední autorka zmiňuje důležitost hodnocení spokojenosti uživatelů s chatbotem a webovou stránkou, na které je umístěno jeho chatovací rozhraní. V rámci toho u vybraných chatbotů zkoumala, zda mají zavedený pětihvězdičkový bodovací systém, hodnocení v rámci dialogu či oboje. Všechny aspekty byly ohodnoceny na stupnici 1–5 bodů, kdy 1 znamená velmi špatný a 5 velmi dobrý výsledek. Úroveň efektivnosti chatbota poté udal průměrný počet bodů ze součtu celkového hodnocení vybraného chatbota.

5 Vymezení role chatbota ve vybraném podniku

K naplnění cíle diplomové práce je potřeba vybrat vhodný přístup k měření efektivnosti chatbota, použít ho k rozebrání tohoto procesu ve zvoleném podniku a nabýt praktických poznatků potřebných k jeho rozšíření. Proto je záměrem této kapitoly představit danou společnost a charakterizovat chatbota, kterého se daná analýza týká. Informace, zde obsažené, pochází z externích zdrojů, jako jsou např. oficiální internetové stránky společnosti a dále jimi poskytnuté interní dokumentace.

5.1 Představení vybraného podniku

ČSOB je stoprocentní dceřinou společností belgické KBC Bank NV, jejíž mateřskou společností je KBC Group NV, kdy obě zmíněné sídlí v Bruselu (ČSOB, 2019a). Klíčovými trhy bankopojišťovací skupiny KBC je Belgie, Česká republika, Slovensko, Maďarsko, Bulharsko, Irsko a v omezené míře své služby nabízí i v jiných zemích (KBC, 2019).

Jak uvádí Ministerstvo spravedlnosti (2019), společnost ČSOB vznikla k 21. 12. 1964 zápisem do obchodního rejstříku v Praze, kde zároveň sídlí její centrála. ČSOB (2019a) zmiňuje, že banka byla založena státem a jejím záměrem bylo zajišťovat financování zahraničního obchodu a volnoměnových operací na československém trhu. Na konci roku 1997 se vláda rozhodla instituci prodat a učinila tak v červnu 1999, kdy se jejím majoritním vlastníkem stala KBC Bank. Hned dvanáct měsíců na to ČSOB převzala Investiční a poštovní banku a o sedm a půl let později, tedy k 1. 1. 2008, se stala slovenská pobočka samostatnou právní osobou.

V současnosti ČSOB na trhu vystupuje jako skupina ČSOB, jejíž součástí je banka ČSOB a dceřiné společnosti, které jsou přímo či nepřímo ovládané ČSOB, popřípadě KBC a mezi ně podle ČSOB (2019b) patří:

- Hypoteční banka, a. s.,
- ČSOB Pojišťovna, a. s.,
- Českomoravská stavební spořitelna, a. s. (ČMSS),
- ČSOB Penzijní společnost, a. s.,

- ČSOB Leasing, a. s.,
- ČSOB Asset Management, a. s.,
- ČSOB Factoring, a. s.,
- Patria Finance, a. s.

Společnost tedy nabízí svým zákazníkům mnoho služeb, resp. spolehlivých řešení, a to od financování jejich potřeb spojených s bydlením až po leasing a factoring. Banka svou pozornost zaměřuje na všechny segmenty trhu, to znamená, že je k dispozici fyzickým osobám (dále jen „FO“), malým a středním podnikům (dále jen „SME“), korporátním a institucionálním klientům (ČSOB, 2019a).

Podle ČSOB (2019b) společnost uplatňuje dualistický systém řízení, a tudíž se skládá z valné hromady, představenstva, dozorčí rady a výboru pro audit. Dále dodává, že funkci valné hromady zastává pouze jediný akcionář společnosti, kterým je KBC Bank NV.

Skupina ČSOB má na území České republiky zřízena dvě Klientská centra (dále jen „KC“), která sídlí v Hradci Králové a Náchodě. Obě call centra nabízí komplexní péči o zákazníky v oblasti vzdáleného bankovního poradenství, přičemž jak ČSOB (2019c) uvádí, využívá mnoho komunikačních kanálů, kterými jsou telefonní hovory, e-mail, SMS, chat, WhatsApp, videohovor, Skype, Facebook, Twitter a dále vytváří podporu pro elektronické bankovníctví a platební karty. Kromě útvarů komunikujících přímo s klienty, se v KC v Hradci Králové nachází i podpůrný a rozvojový tým zabezpečující technologii využívanou, jak ze stran kontaktního centra, tak i ze stran pobočkové sítě.

ČSOB dále zaměstnává tým odborníků působících v útvarech robotizace, inovace, digitální transformace, kteří se zabývají automatizací, robotizací a rozvojem umělé inteligence. V bance je momentálně nasazeno přes 80 plně funkčních robotů (Sikorová, 2019), kteří pomáhají „lidským“ pracovníkům především s rutinními činnostmi a zefektivňují tak hlavně transakční a kontrolní procesy, s jejichž implementací pomáhá např. společnost Deloitte (Deloitte, 2019). ČSOB se v této oblasti může pyšnit i oceněním, které získala v mezinárodní soutěži The Efma-Accenture Customer Insight & Growth Banking Innovation Awards. V té totiž získala 3. místo z celkem 635 přihlášených inovací v kategorii „Workforce Experience“

a účastnila se jí s týmovým robotem Ró, který kombinuje technologii workflow, robotizaci a automatickou analýzu textu (ČSOB, 2019d).

5. 2 Charakteristika vybraného chatbota

V návaznosti na představení společnosti ČSOB, se tato podkapitola věnuje charakteristice jednoho z jejích chatbotů, který nese jméno Radim. V dané bance je zaveden od května roku 2019, jeho vývojem se zabývá pražský ČSOB tým specializovaný na Big Data a AI platformu, ale jeho správa, podpora a další rozvoj je zastřešován pracovníky již zmiňovaného KC. Jak už bylo poznamenáno v podkapitole 5.1, tak hlavní úlohou KC, jakožto kontaktního centra je obsluha klientů prostřednictvím hlasových i nehlasových interakcí, kdy mezi nehlasové mimo jiné patří právě služby prováděné chatbotem. Radimovým úkolem je sloužit pro zákaznický servis a jako podpora prodeje půjček na cokoli, s nímž lze komunikovat za pomoci odpovědních tlačítek (předdefinovaný scénář konverzačního stromu) a formou free textu (využití Q&A modulu).

5.2.1 Účel chatbota

Chatbot Radim byl do banky zaveden především z důvodu rozšíření a zkvalitnění služeb pro stávající a potenciální klienty v oblasti půjček, přesněji půjček na cokoli. Tato technologie podporuje retailovou strategii banky „Digitálně s lidmi“, která v sobě kloubí lidský prvek v digitálním prostředí. Jeho úkolem je odpovídat na kladené dotazy ve formě volného textu, nabízí odpovědní tlačítka, přeměrovává na důležité dokumenty (např. obchodní podmínky pro ČSOB spotřebitelské úvěry či povolené přečerpání účtu) a nabízí sjednání půjčky online a po telefonu. Podle Revenda a kol. (2014) je úvěr bankovním produktem, který kryje finanční potřeby klientů dané banky. Dále dodávají, že dlužník musí ve sjednané době splatit jistinu úvěru a zaplatit připsané úroky, které jsou pro věřitele odměnou za její poskytnutí. V odborné literatuře se lze setkávat se ztotožňováním pojmů úvěr a půjčka, které i pro potřeby práce nabývají stejného významu. Jedná-li se o půjčku peněžních prostředků, která je klientovi poskytnuta v bance a klient je vázán povinností hradit vzniklé úroky, znamená úvěr to samé. Jak bylo uvedeno výše, chatbot Radim se zabývá půjčkou na cokoli, která se řadí do kategorie neúčelových úvěrů/půjček. Tu lze sjednat velice snadno, protože žadatel nemusí dokládat účel

použití půjčených peněz a v případě ČSOB se jedná aktuálně o půjčku do 800 000 Kč bez zajištění, kterou lze rozložit do menších splátek až na osmileté období (ČSOB, 2020b). Dále daná banka uvádí, že v případě klientů lze sjednat tento typ půjčky z pohodlí domova skrz internetové a mobilní bankovníctví, k čemuž je zapotřebí mít u sebe pouze občanský průkaz a telefon na který je zaslán SMS kód pro podpis příslušných dokumentů. Žadatelé, kteří nejsou klienty banky, poté mají možnost získat půjčku vyplněním online žádosti bez nutnosti přihlašování se do internetového bankovníctví či Smartbankingu. S tím vším je chatbot Radim schopen poradit a přesměrovat na žádoucí místa (dále podrobněji viz 5.2.2).

5.2.1.1 Poslání a vize

Obecným posláním společnosti ČSOB je poskytovat co nejširší spektrum bankovních produktů a služeb všem segmentům trhu, tj. FO, SME, korporátním a institucionálním klientům a stát se tak největší bankou působící v České republice. ČSOB (2019b) uvádí, že podle celkových aktiv skupina ČSOB patří mezi jednu ze tří nejvýznamnějších bankovních skupin na území státu. K tomuto poslání kromě lidských pracovních sil bezesporu dopomáhají i technologie zavedené v bance a vzhledem k probíhající čtvrté průmyslové revoluci, společnost neustále usiluje o udržování, rozvoj a přizpůsobování se digitálnímu světu tak, aby stávajícím i potenciálním klientům umožňovala, co nejkomfortnější a nejefektivnější využití všech jejích služeb.

Jednou z těchto technologií je i chatbot Radim, jehož účelem je přispívat k plnění poslání banky a k jehož nasazení vedla vize vytvořit chatbota, který bude schopen nahradit živé operátory na úrovni plnohodnotné konverzace týkající se půjček na cokoli. Tato vize je podmíněna především samotným vývojem chatbota, který s sebou nese až roky práce a vzhledem k tomu, že se jedná o bankovní technologii, tak chtěným stavem je být schopen regulovat výstupy, které chatbot Radim produkuje. Jeho konverzační tok je tedy plně kontrolován a řízen, i proto byl implementován formou konverzačního stromu (odpovědních tlačítek) a free textu, v rámci kterého je uživateli vracen výstup založený na vyhledávání nejrelevantnějšího kandidáta ze sady předem vytvořených Q&A. Kdyby byl chatbot typu generative-based a samostatně se učil, nebylo by možné jej řídit, což je nechtěný stav, a proto banka jeho učení ponechává pouze na odpovědných osobách (administrátorech). S touto hlavní vizí je

spojena další, a to v podobě úspor, mezi které patří odhadované ušetření 0,5 – 1 FTE. K tomu vedou následující skutečnosti, které lze považovat za hlavní výhody zavedení chatbota:

- chatbot nikdy nespí, není nemocný, nebere si dovolenou, pracuje o víkendech a svátcích – tzn. funguje 24/7 a může tak být neustále v kontaktu se stávajícími i potenciálními klienty, uspokojit jejich informační potřeby a pomoci jim se sjednáním půjčky (nepřímo přispívá k výnosům plynoucích z jejího poskytnutí),
- jeho odezvy jsou okamžité,
- svojí činností šetří čas živého operátora, který se tak může věnovat jiné, důležitější práci – zvyšuje se tak produktivita daných zaměstnanců, resp. dochází k lepšímu využití pracovních sil,
- na rozdíl od lidí je učení chatbota téměř okamžité, co se naučí už nezapomene a dělá jenom to co ho naučí administrátor,
- lze regulovat a eliminovat chybovost jeho výstupů,
- lze ho rychle nasadit do produkčního prostředí a dále customizovat podle potřeb.

5.2.2 Definice chování chatbota

Pro uspokojení potřeb a motivování využití služeb chatbota Radima stávajícími či potenciálními klienty, jsou nabízeny dva základní scénáře, fungující jako jeden komunikační celek. Těmi jsou:

- servisní – chatbot Radim odpovídá na otázky dotazujících, poskytuje jim informace/servis, udržuje plynulou zdvořilou konverzaci a v případě, že není schopen klientovi pomoci, poprosí ho o zanechání telefonního čísla, na které se mu ozve živý operátor v příslušné provozní době či pokud je projevem zájem o úvěr, dochází ke kontinuálnímu přechodu na prodejní scénář,
- prodejní – nabízí klientovi sjednání půjčky online či po telefonu, kdy v případě, že má zájem o sjednání online, tak ho chatbot Radim směřuje na přihlašovací stránky internetového bankovníctví, mobilní aplikaci Smartbanking

či odkazuje na online formulář a pokud raději preferuje jednání po telefonu, chatbot ho požádá o telefonní číslo, které dále předává kompetentnímu živému operátorovi a ten ho v provozní době kontaktuje.

Chatbota Radima lze tedy nalézt na oficiálních internetových stránkách ČSOB v sekci *Půjčky* -> *Půjčka na cokoliv*, jehož chatovací prostředí se nachází vpravo v rohu, kde nabízí dvě možnosti zahájení konverzace:

- stisknutí ikony **+**, která otevře chatovací prostředí (viz obrázek 2),




Obrázek 2: Možnost otevření chatovacího rozhraní na liště chatu
Zdroj: ČSOB (2020a)

- po definované době 30 sekund dochází k automatickému vyskočení pop-up okna.

Jakmile se uživatel prvně nachází v chatovacím prostředí, chatbot Radim ho osloví, zdvořile se představí a nabízí spolupráci, což lze konkrétně vidět na následujícím obrázku 3. Tato textace je vždy stejná.



Obrázek 3: Chatovací rozhraní
Zdroj: ČSOB (2020a)

Chatbot Radim nabízí možnost aktualizace pomocí ikony , která se nachází vpravo nahoře a vrací první rozcestník konverzace (viz obrázek 3) na úroveň již rozběhlého chatu, a tudíž si může jeho účastník nerušeně vybrat jednu z dalších ze tří dostupných možností a zahájit tím nové konverzační flow za pomoci tlačítkového stromu a podle potřeby v kombinaci s free textem. Přitom se historie chatu nemaže, takže se lze pohodlně vrátit zpět k rozebraným tématům a navázat na ně. V případě, že by se snad někdo v konverzaci již nevyznal a chtěl by vytvořit úplně nové vlákno, je potřeba aktualizovat celou webovou stránku.

Jak už bylo uvedeno výše, tak tento chatovací robot je schopen interakce s živou osobou dvěma způsoby, kterými jsou:

- konverzační strom (odpovědní tlačítka),
- free text.

Konverzační strom se řídí přesně definovaným scénářem. Výhodou této implementace je, že eliminuje chybovost, na druhou stranu je dialog velice omezen svými možnostmi. Největším přínosem je pro lidi, kteří neví, na co přesně se ptát, a tak i navzdory neznalosti svého konkrétního dotazu, se mohou pomocí nabízených možností proklikat k odpovědi, kterou hledají.

Po otevření chatovacího rozhraní lze jako první zprávu od Radima vidět již zmiňovaný uvítací pozdrav (viz obrázek 3). Dále chatbot nabízí výběr z možností: *Chci půjčku*, *Nezávazně spočítat*, *Chci informace*, nebo lze využít free textu napsáním zprávy do kolonky *Zadejte zprávu*. První tři východiska se poté větví do dílčích uzlů. Vzhledem k tomu, že je větvení stromu velmi rozsáhlé, tak se v následujícím schématu (viz schéma 1) nachází vizualizace pouze základního scénáře konverzačního flow po prokliknutí *Chci půjčku*, *Nezávazně spočítat*, *Chci informace*, kdy ukázané terminální uzly *Online žádost*, *Internetové bankovníctví* a *Smartbanking* se chovají následujícím způsobem:

- online žádost – chatbot Radim přesměrovává k příslušné dlaždici umístěné na webu ČSOB, ve které mu znovu nabízí možnost sjednání přes internetové bankovníctví (kdyby se uživatel v konverzaci s chatbotem náhodou překlikl) a poté mu správně nabízí možnost vyplnit online formulář,

- internetové bankovníctví – chatbot Radim přímo odkazuje na přihlašovací stránku internetového bankovníctví,
- Smartbanking – chatbot Radim popisuje, jak sjednat půjčku v mobilním bankovníctví a poté poskytuje odkaz, na základě kterého je uživatel přesměrován na URL stránku ČSOB Smartbanking, kde se dozví vše potřebné o aplikaci, včetně QR kódu, po jehož načtení se dostane na svém chytrém zařízení do příslušného obchodu, kde může program zdarma okamžitě získat.

Co se týká informací o alternativách *Podmínky poskytnutí, Parametry půjček, Příjmy* až *Sazebník, obchodní podmínky*, tak ty dále chatbot Radim upřesňuje či odkazuje na celé jejich znění. Pokud uživatel klikne na tlačítko *Zajímá mě něco jiného*, poté je vyzván napsat dotaz formou volného textu.

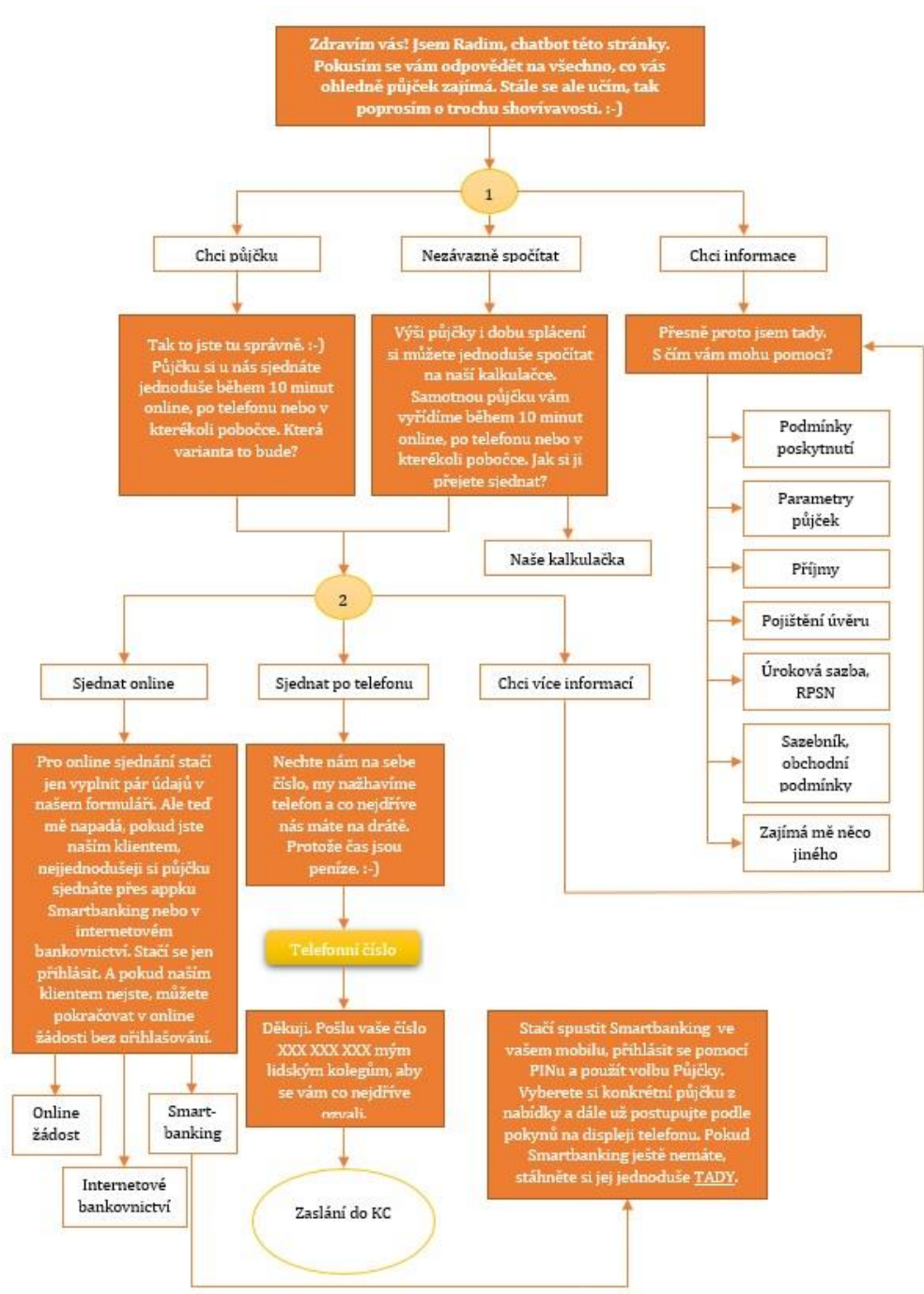


Schéma 1: Ukázka konverzačního stromu (úryvek ze základního scénáře)
 Zdroj: vlastní zpracování podle ČSOB (2020a)

Během interakce formou free textu, chatbot Radim, podle vstupu od uživatele, vyhledává ve znalostní bázi nejvhodnější výstup (nejvhodnějšího kandidáta). Jeho znalostní báze je naplněná sadou relevantních a obecných Q&A, v nichž podle daného algoritmu (pravidla) hledá nejlepší shodu. Výběr kandidáta je podmíněn významností (váhou) vzorových otázek a odpovědí, jejichž obsah je přímo ovlivněn odpovědnou osobou (více viz 7.3). Pokud robot uzná za vhodné, tak i během volné

konverzace, klientovi nabízí možnosti, z kterých si může vybrat a navázat tak na konverzační strom. Tyto dva způsoby tedy lze během interakce kombinovat, což vede k nejlepšímu výkonu chatbota, který je tak schopen poskytnout co nejširší spektrum informací. V případě, že chatbot správně neporozumí otázce tazajícího či nezná odpověď, omluví se a požádá ho, aby na sebe zanechal kontakt v podobě telefonního čísla, které dále distribuuje oprávněnému agentovi na příslušný útvar k vyřízení (dokončení).

5.2.3 Dostupnost a obsluha

Jednou z hlavních výhod zavedení chatbota je jeho dostupnost 24/7, což znamená, že na rozdíl od živých osob, chatbot nikdy nespí, není nemocný, nebere si dovolenou, slouží o víkendech, svátcích, v noci a nevyžaduje za to žádné příplatky, nepotřebuje nikým nahradit, a tudíž za něj firma platí jedny „mzdové“ náklady. Ve skutečnosti se ale jedná o přímé i nepřímé náklady spojené s pořízením, provozem, údržbou atd. Dalším důležitým přínosem této technologie je úspora času zaměstnance, který před implementací chatbota musel obsluhovat veškeré dotazy, které se neustále opakovaly a zdržovaly ho od jiné důležité práce, kterou nelze zautomatizovat, jako je tomu právě u rutinních úkolů prováděných chatbotem.

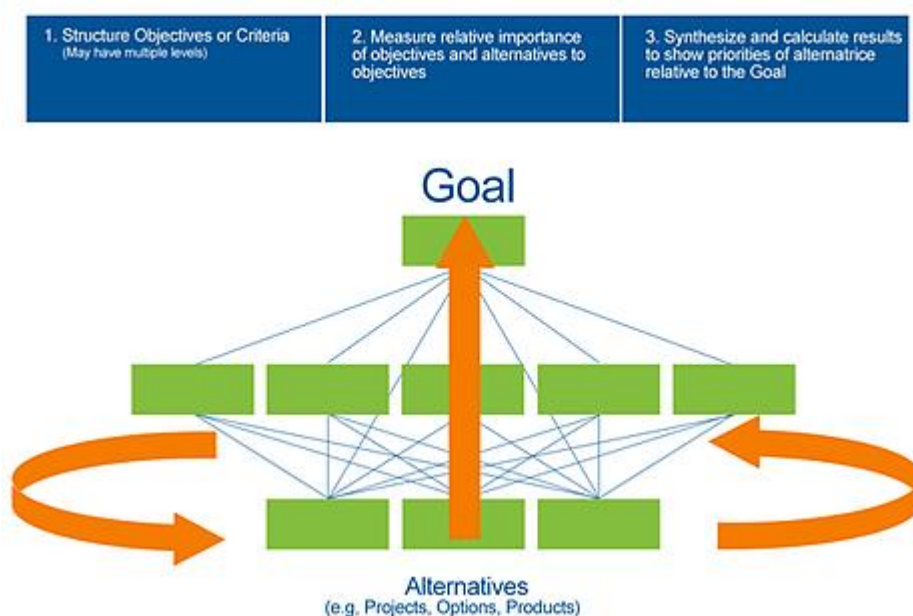
Holzman (2019) na základě rozhovoru se spoluzakladatelem pardubického startupu Feedyou, Vojtěchem Dlouhým, uvádí, že 85 % lidí s chatboty komunikuje mimo klasickou pracovní dobu v čase mezi šestou a desátou hodinou večer. Banka se snaží být pro veřejnost co nejvíce dostupná, a právě s touto technologií dokáže kdykoli uspokojovat informační potřeby stávajících i potenciálních klientů v oblasti poskytování neúčelových půjček, což je vnímáno jako největší benefit. I proto, jak uvádí Dittrichová a kol. (2014), hlavní zdroje příjmů bank tvoří úroky z půjček peněz a různé druhy poplatků, je tedy potřeba se soustředit na zvyšování přidané hodnoty pro zákazníky právě v této oblasti.

V případě, že chatbot Radim nedokáže nějaký požadavek zcela odbavit, je schopen posbírat informace a předat telefonní číslo klienta na kompetentního živého operátora, který ho v provozní době svého útvaru kontaktuje a už ví, co s ním bude řešit. I toto znamená značnou úsporu času pro daného pracovníka.

Interakce, které chatbot Radim nedokáže z důvodu nedostatečných znalostí dokončit sám, obstarávají pracovníci KC, kteří mají na starosti servis, a to ve všední dny od 7 do 22 hodin a o víkendech a svátcích od 8 do 22 hodin. Ve chvíli, kdy je ze strany tazajícího projevem zájem o půjčku a zanechává na sebe kontakt, tak je požadavek směrován na útvar zabezpečující spotřební financování, který standardně pracuje ve všední dny od 8 do 19 hodin, v sobotu od 9 do 17 a neděli od 10 do 18 hodin. Požadavky produkované chatbotem k příslušným pracovníkům směřují podle pořadí, ve kterém byly vytvořeny a ti s nimi pracují skrz aplikaci WDE (Workspace Desktop Edition).

6 Výběr vhodného přístupu k měření efektivity chatbota

Aby byl splněn cíl diplomové práce je potřeba zvolit vhodný přístup k měření efektivity chatbota. Ten byl vybrán pomocí analytického nástroje pro podporu rozhodování, který vystupuje pod názvem Expert Choice 2000. Možnost volně dostupného použití tohoto softwaru poskytuje Fakulta informatiky a managementu Univerzity Hradec Králové, kde autorka studuje. Tento program pro vícekritériální analýzu vyvinula společnost Expert Choice, jejímž zakladatelem byl E. Forman a T. L. Saaty. Dále Expert Choice (2020) uvádí, že to právě T. L. Saaty dal ke vzniku The Analytic Hierarchy Process (dále jen „AHP“), na základě kterého Expert Choice 2000 pracuje. Společnost strukturu jednoduché AHP hierarchie ilustruje na následujícím obrázku.



Obrázek 4: Ukázka AHP hierarchie

Zdroj: Expert Choice (2020)

V rámci této kapitoly byla nejdříve určena relevantní kritéria a jejich jednotlivé váhy. K tomu bylo využito expertního stanoviska vytvořeného za pomoci Saatyho metody párového porovnání, ve kterém celkem šest odborníků KC ČSOB, zainteresovaných v této problematice, hodnotilo vybraná kritéria a určovalo jim jimi vnímanou důležitost. Výsledky komparace, spolu s vypočítanými váhami definovaných kritérií, byly zaznamenány do dílčích a jedné souhrnné Saatyho matice, které jsou poskytnuty k nahlédnutí v příloze č. 2. Získaná a sesumarizovaná data byla přenesena do analytického softwaru Expert Choice 2000, za pomoci kterého poté v závislosti

na kritériích a jejich vahách, došlo k porovnání a výběru jedné z alternativ, tedy vhodného přístupu k měření efektivity chatbota.

6.1 Stanovení kritérií

Podle Expert Choice (2020) je prvním krokem, při vícekritériálním rozhodování, definování cíle, kterého by mělo být dosaženo. Ten byl v tomto případě nazván jako „*Výběr vhodného přístupu k měření efektivity chatbota*“. Následně bylo zapotřebí stanovit kritéria pro výběr. Ta byla zvolena na základě pečlivého prostudování odborných studií zabývajících se problematikou hodnocení chatbota, které byly rozebrány v podkapitole 4.2 a rad expertů z ČSOB. Jedná se tedy o pět nejdůležitějších oblastí, které by podle již zmíněných teoretických východisek a expertního stanovení, neměly ve správném přístupu k měření efektivity chatbota chybět:

- **uživatelská přívětivost** – vyhodnocování, jakým dojmem chatbot (vč. chatovacího rozhraní) působí na uživatele a zda jsou s jeho službami během interakce spokojeni,
- **informační schopnost a vybavenost** – zjišťování a doplňování stavu znalostní báze chatbota tak, aby produkované výstupy dokázaly uspokojit potřeby uživatele,
- **jazyková úroveň a vybavenost** – hodnocení schopnosti chatbota vytvářet z hlediska pravopisu a gramatiky správné a slovně rozmanité výstupy,
- **lidskost** – posuzování, zda se chatbot chová jako člověk,
- **obchodní hledisko** – zjišťování, jakou přidanou hodnotu přináší chatbot pro společnost a s tím metriky spojené.

6.2 Stanovení vah jednotlivých kritérií

I navzdory tomu, že lze všechna stanovená kritéria považovat za důležitá, tak nikoli stejnou měrou. Proto bylo zapotřebí každému z nich určit přiměřené váhy. K tomu bylo využito výstupů z dotazníkového šetření (viz příloha č. 1), v podobě metody párového porovnání, provedeného uvnitř podniku. Přesněji se výzkumu účastnilo dohromady šest vybraných manažerů a analytiků, zainteresovaných v této problematice, a tudíž schopných relevantního posouzení, kteří pracují v ČSOB v oddělení Rozvoj a provoz technologie KC v Hradci Králové. Právě KC zajišťuje správu, provoz, podporu

a veškeré procesy týkající se analyzovaného chatbota Radima. Cílem šetření bylo zjistit, jakou důležitost daní jednotlivci přisuzují definovaným kritériím, na základě čehož došlo ke stanovení jejich výsledných vah, tedy expertnímu stanovisku. Dotazník byl vytvořen podle Saatyho metody párového porovnání a podoba jeho zpracování byla inspirována samotným programem Expert Choice 2000. Data získaná z provedeného výzkumu byla zpracována do dílčích Saatyho matic (viz příloha č. 2). Nejdříve byla pro každého experta zvlášť vytvořena matice, jejímž výstupem byl výpočet dílčího a celkového geometrického průměru a následných vah pro první až páté kritérium. Tyto vypočítané váhy byly přeneseny do společné matice (viz tabulka 10) znázorňující jejich vztah vůči prvnímu až šestému respondentovi. Poté byl pro každé kritérium z jednotlivých vah vypočítán aritmetický průměr, udávající hodnoty výsledných (jednotných) vah, které lze vidět v tabulce 1.

Kritérium	Váha	Pořadí
Uživatelská přívětivost	0,214	3.
Informační schopnost a vybavenost	0,345	1.
Jazyková úroveň a vybavenost	0,099	4.
Lidskost	0,045	5.
Obchodní hledisko	0,297	2.

Tabulka 1: Stanovené váhy kritérií a jejich výsledné pořadí
Zdroj: vlastní zpracování

Jak si lze povšimnout, tak ze všech pěti kritérií, respondenti považují za nejdůležitější informační schopnost a vybavenost. Chatbot Radim slouží pro zákaznický servis a podporu, co se týká poskytování informací ohledně půjček na cokoli a možnostech jejich sjednání (více viz 5.2), do jisté míry ho lze nazvat jako FAQ bota, a proto banka především dbá na vybavenost jeho znalostní báze a schopnost produkovat výstupy vedoucí k uspokojení potřeb stávajících či potenciálních klientů. Zajišťování informační schopnosti a vybavenosti tedy představuje přidanou hodnotu pro zákazníka a hned za ní se umístilo obchodní hledisko, které je tvořeno metrikami vyhodnocující přidanou hodnotu pro společnost. Váhy obou kritérií zvlášť se pohybují okolo třetinového podílu ze všech pěti kritérií a představují tak oproti nim dohromady téměř dvoutřetinovou preferenci. Ve vhodném přístupu k měření efektivnosti chatbota by tedy rozhodně neměly chybět metriky zaměřené na tyto dvě hlediska.

Jako třetí nejdůležitější kritérium je vnímána uživatelská přívětivost, která vyžaduje vyhodnocování, jakým dojmem chatbot působí na uživatele a jak jsou spokojeni s jeho službami. Je tedy potřeba se zaměřit i na zajišťování možnosti získat co nejpříjemnější a nejefektivnější zkušenost, které se uživateli během interakce s chatbotem dostává.

Na čtvrtém místě, co se týká důležitosti, se nachází jazyková úroveň a vybavenost. Tou je myšleno posouzení schopnosti chatbota vytvářet z hlediska pravopisu a gramatiky správné a slovně rozmanité výstupy, díky čemuž je uživatel schopen zachytit přesnou interpretaci předávaných informací.

Za nejméně důležité kritérium je považována lidskost. V případě umělé inteligence testování lidskosti hraje podstatnou roli, ovšem chatbot Radim výstupy vrací na základě předdefinovaného scénáře a vyhledávání shody v sadě Q&A vytvořených odpovědným zaměstnancem a je konstruován tak, aby byl schopen splnit konkrétní úkol. Uživatel už dopředu ví, že konverzuje s chatbotem a nikoli živým operátorem, a tak společnost svou pozornost věnuje více ostatním, v tomto případě podstatnějším, kritériím.

6.3 Výběr alternativ

V odborných vědeckých databázích jako je například Springer, Science Direct, Web of Science bylo podle kombinací klíčových slov „chatbot“, „effectiveness“, „measurement“, „assessment“, „evaluation“, „metrics“, po pečlivém prostudování, nalezeno malé množství publikací, které by svým obsahem byly pro účely diplomové práce použitelné, resp. dostatečně zainteresované v řešené problematice. Po prvním vyloučení se do druhého kola dostalo celkem sedm studií, které se zabývaly alespoň dvěma oblastmi měření, resp. kritérii pro výběr (viz 6.1) a obsahovaly užitečné metriky. V podkapitole 4.2 byla provedena diskuze rozebírající nejrelevantnější poznatky získané právě z těchto sedmi analyzovaných studií a do nejužšího výběru se poté dostaly celkem čtyři z nich, které nabízely nejkomplexnější přístup k měření efektivnosti chatbota, resp. se zabíraly alespoň čtyřmi z pěti požadovaných oblastí měření.

Pro výběr vhodného přístupu k měření efektivnosti chatbota byly zvoleny následující alternativy:

- Kuligowska (2015),
- Maroengsit et al. (2019),
- Peras (2018),
- Radziwill a Benton (2017).

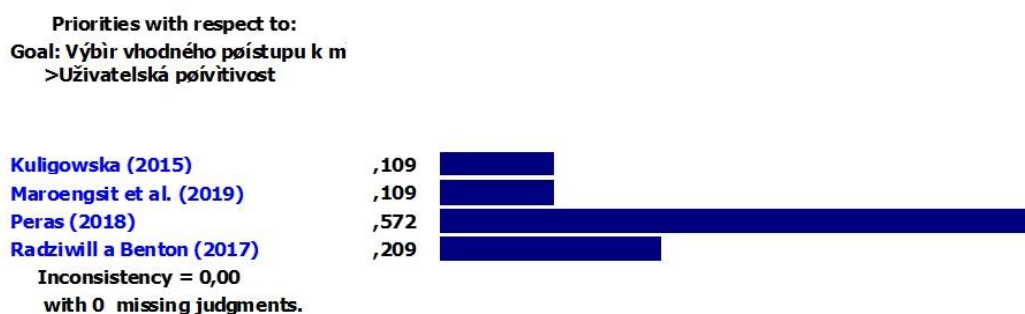
6.4 Porovnání alternativ podle stanovených kritérií

Po stanovení kritérií a jejich náležitých vah následovalo porovnání vybraných alternativ. K tomu byl využit analytický software Expert Choice 2000, který byl představen na začátku kapitoly 6. Jednotlivé přístupy k měření efektivnosti chatbota mezi sebou byly komparovány na základě míry zahrnutí/nezahrnutí definovaných kritérií z podkapitoly 6.1. Dále během porovnávání hrála podstatnou roli možnost proveditelnosti navržených metrik z pohledu podniku. Program poté vyhodnotil z hlediska každého kritéria nejvhodnější přístup.

6.4.1 Uživatelská přívětivost

Vybrané přístupy pohlízejí na uživatelskou přívětivost ze dvou úhlů. Jedním z nich je externí hodnocení, které navrhuje implementovat pomocí dotazníkového šetření provedeného na uživatelích (klientech) či nezávislých osobách, otázky zobrazující se při zahájení, v průběhu či ukončování chatu a (pěti)hvězdičkového systému hodnocení. Na druhé straně se jedná o vnitropodnikové expertní hodnocení, kdy je podle přesného scénáře na určité hodnotící stupnici posuzováno chatovací rozhraní, jeho působení, funkcionality a další atributy. To ovšem ztrácí možnost zhodnotit spokojenost uživatelů během interakce s chatbotem, která je do tohoto kritéria též zařazována. Na základě porovnání alternativ z hlediska uživatelské přívětivosti, byl za nejvhodnější zvolen přístup od Peras (2018). Oproti ostatním možnostem Peras (2018) nabízí komplexní a uspořádaný souhrn metrik, který nazvala jako „perspektiva uživatelské přívětivosti“, k čemuž dodává upřesňující slovní vyjádření. Radziwill a Benton (2017) nabízí pro měření určitý balíček seřazených atributů, ovšem bez další specifikace. V ostatních případech je popsán postup měření, ale vždy pouze z jednoho úhlu a soubor

ukazatelů je tedy v porovnání s alternativou od Peras (2018) a Radziwill a Benton (2017) omezený.



Obrázek 5: Porovnání alternativ podle kritéria „Uživatelská přívětivost“

Zdroj: vlastní zpracování v programu Expert Choice 2000

6.4.2 Informační schopnost a vybavenost

Jak bylo uvedeno v podkapitole 6.2, tak informační schopnost a vybavenost byla kompetentními hodnotiteli z ČSOB shledána nejdůležitějším kritériem. Z tohoto hlediska by tedy neměla být vybraným přístupem pro měření efektivity chatbota opomíjena. Porovnávané studie se opět odlišují svým přístupem, resp. některé jsou přímo orientované na obsah znalostní báze chatbota a posuzují ho na základě konkrétně sestavených testovacích/hodnotících scénářů, komparují výstupy dvou stejně zaměřených chatbotů, či se zabývají kvantitativními metrikami, jako je např. měření doby, za kterou je chatbot schopen vytvořit odpovídající reakci. V tomto případě je soustředěno především na posuzování naplnění znalostní báze a následného vyhledávání či generování adekvátních outputů. Toto kritérium nejlépe splňuje přístup od Kuligowska (2015), která považuje znalostní bázi a její následné použití za hlavní náplň existence chatovacího robota. I proto nabízí způsob, jak zhodnotit obecné a specializované znalosti, kterými chatbot disponuje. Za ní se umístil přístup od Peras (2018), která využívá patřičné kvantitativní metriky. Zbývající dva postoje nejsou dostačující, z důvodu malého zohlednění tohoto kritéria, přesněji Maroengsit et al. (2019) se zabývají pouze hodnocením vytvářené odezvy a Radziwil a Benton (2017) zmiňují jen jeden atribut řešící dostatek znalostí chatbota.

Priorities with respect to:
 Goal: Výběr vhodného přístupu k m
 >Informační schopnost a vyb...



Obrázek 6: Porovnání alternativ podle kritéria "Informační schopnost a vybavenost"

Zdroj: vlastní zpracování v programu Expert Choice 2000

6.4.3 Jazyková úroveň a vybavenost

Toto kritérium se v porovnání s ostatními, co se týká důležitosti, umístilo na předposledním místě, avšak ve všech alternativách mu je věnován určitý prostor. S jazykovou úrovní a vybaveností se pojí schopnost chatbota vytvářet smysluplné odezvy, což znamená výstupy, které budou gramaticky korektní a tedy srozumitelné. Aby byla odezva kvalitní, je poté potřeba spolupráce s předchozím kritériem, které zajišťuje informace uspokojující uživatelskou potřebu. Na základě komparace se jako nejvhodnější jeví třetí přístup, jehož autorkou je Peras (2018), která se v tomto případě zaměřila především na kvalitativní metriky, které rozdělila celkem do pěti oblastí. Maroengsit et al. (2019) se věnují již existujícím algoritmům měřícím kvalitu strojového překladu z jednoho jazyka do druhého a vyhodnocují atributy jako je přesnost a relevance použitých výrazů. Dále i napříč tomu, že posouzení přímo uživateli je podle nich nejpřesnější, tak navrhují využití expertního hodnocení, které je levnější a rychleji proveditelnou metodou. Kuligowska (2015) navrhuje sledovat tři aspekty během interakce s chatbotem a na základě toho vyhodnotit závěry. Jako poslední skončil přístup od Radziwill a Benton (2017), který poskytuje pouze výčet atributů, které dále nespecifikuje.

Priorities with respect to:
 Goal: Výběr vhodného přístupu k m
 >Jazyková úroveň a vybaven...

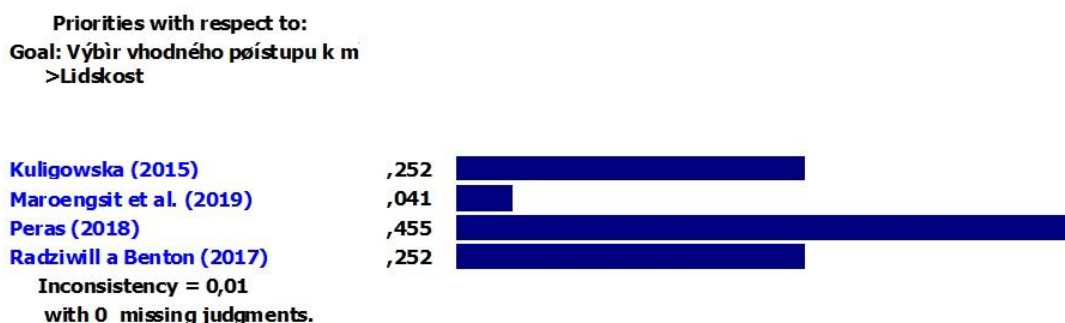


Obrázek 7: Porovnání alternativ podle kritéria "Jazyková úroveň a vybavenost"

Zdroj: vlastní zpracování v programu Expert Choice 2000

6.4.4 Lidskost

V případě specializované a obecné AI či chatbota založeného na generování vlastních výstupů hraje posuzování lidskosti významnou roli, ovšem jak bylo již v diplomové práci zmíněno, chatbot Radim funguje na základě definovaných pravidel a vyhledává nejrelevantnější odezvu z předem vytvořené sady Q&A a nelze ho označit za AI, jak se někteří domnívají. I proto tomuto kritériu byla přisouzena nejnižší váha, ostatní kritéria jsou pro společnost důležitější. Peras (2018) a Radziwill a Benton (2017) nabízí, pro prokázání lidskosti, aplikovat Turingův test, o kterém bylo zmiňováno v podkapitolách 3.1 a 3.2 a jak bylo již uvedeno v podkapitole 3.1, prozatím tímto testem žádný chatbot neprošel. Peras (2018) se dále soustředí na metriky kvalitativního i kvantitativního charakteru. Radziwill a Benton (2017) se poté zabývají spíše schopností udržet diskuzi na určité téma či odpovídat na specifické dotazy. Kuligowska (2015) navrhuje sledovat lidské rysy vykazované chatbotem během konverzace, a to za pomoci pokládání osobních otázek mívajících na věk, pohlaví, zájmy a podobně a v poslední řadě Maroengsit et al. (2019) toto hledisko specificky vůbec nezohledňují.



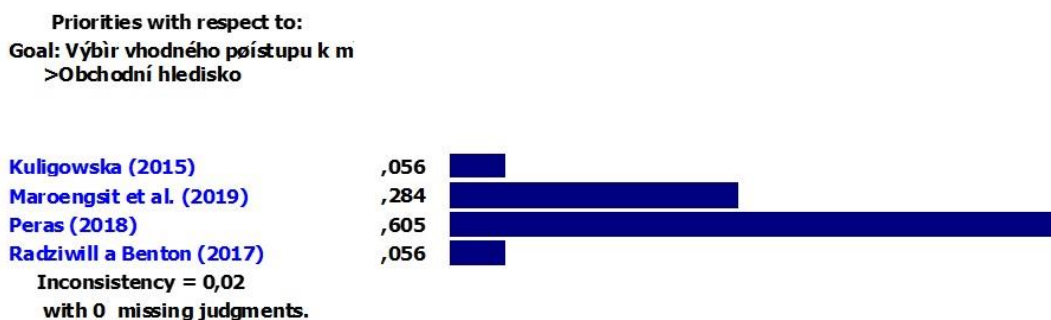
Obrázek 8: Porovnání alternativ podle kritéria "Lidskost"

Zdroj: vlastní zpracování v programu Expert Choice 2000

6.4.5 Obchodní hledisko

Vybraní posuzovatelé z ČSOB pokládají vyhodnocování obchodního hlediska za druhé nejdůležitější kritérium, resp. je podle nich klíčovým faktorem hned po vytváření přidané hodnoty pro zákazníka, vytvářet přidanou hodnotu pro společnost. V tomto ohledu je nejlépe zaměřený přístup od Peras (2018), která rozebírá vyhovující kvantitativní metriky. Podobně Maroengsit et al. (2019) navrhuji sledování určitých

statistik, které ale dále nekonkretizují. Zbývající přístupy se tomuto kritériu nikterak nevěnují.

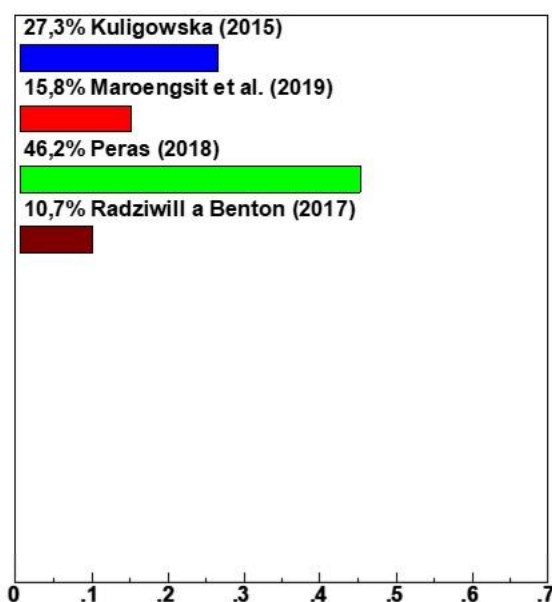


Obrázek 9: Porovnání alternativ podle kritéria "Obchodní hledisko"

Zdroj: vlastní zpracování v programu Expert Choice 2000

6.5 Vyhodnocení výběru vhodného přístupu k měření efektivity chatbota

Za pomoci procesu vícekritériálního rozhodování zrealizovaného v AHP softwarovém produktu Expert Choice 2000 byly vytvořeny následující dva grafy (viz obrázek 10 a 11) znázorňující vhodnost vybraných přístupů pro měření efektivity chatbota.

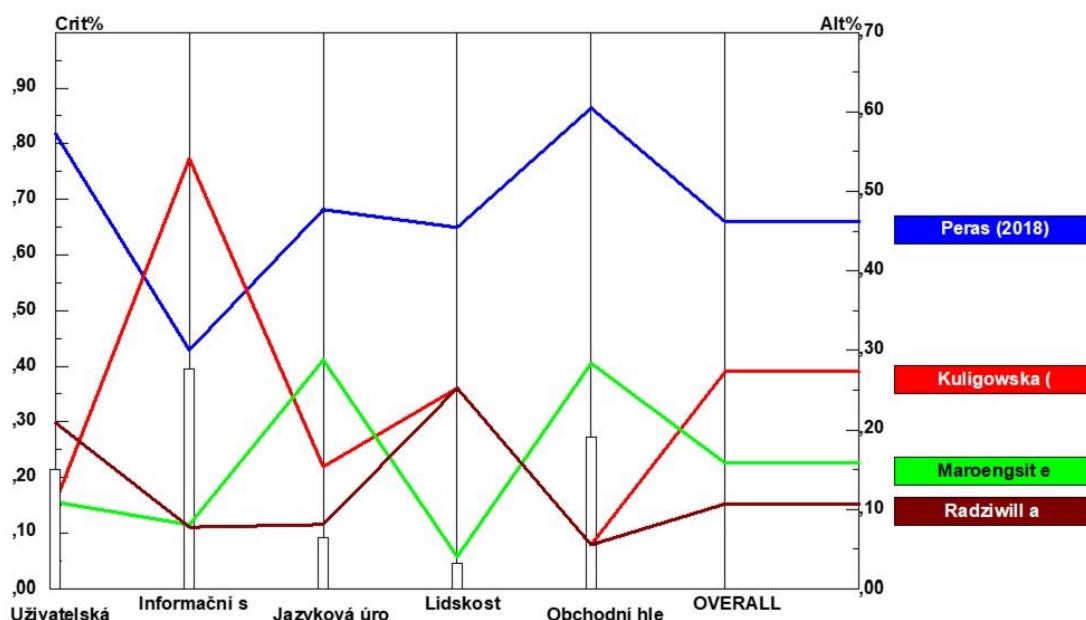


Obrázek 10: Výsledek výběru vhodného přístupu k měření efektivity chatbota v procentuálním vyjádření

Zdroj: vlastní zpracování v programu Expert Choice 2000

Na obrázku 10 lze vidět rozdíly ve vhodnosti jednotlivých přístupů v procentuálním vyjádření. Na základě výsledků provedené multikritériální analýzy je zřejmé,

že nejlépe hodnocený je přístup od Peras (2018), který dosahuje téměř 50% převahy oproti ostatním alternativám. Naopak pouhé desetiny dosáhl postoj k měření od Radziwill a Benton (2017), který byl v porovnání s ostatními přístupy z hlediska tří z pěti kritérií označen za nedostačující, a tudíž nevyhovující.



Obrázek 11: Grafické znázornění výběru vhodného přístupu k měření efektivity chatbota

Zdroj: vlastní zpracování v programu Expert Choice 2000

Pomocí druhého grafu (viz obrázek 11) lze na rozdíl od předchozího (viz obrázek 10) sledovat zařazení vhodnosti vybraných přístupů k měření efektivity chatbota vzhledem k jednotlivým kritériím. Lze si také povšimnout znázornění dílčích vah, které vypovídá o jejich relevanci a je poté tedy možné usoudit, ve kterých hlediscích vynikají dané přístupy.

Za nejvhodnější přístup ze čtyř analyzovaných byl zvolen způsob měření vytvořený autorkou Peras (2018), která proces hodnocení efektivity chatbota rozdělila celkem do pěti perspektiv. Jí definovaná hlediska v podstatě plně vystihují zvolená kritéria, a proto oproti ostatním alternativám, až na jednu výjimku, vždy získala lepší skóre. Zmíněnou výjimkou je viditelný výkyv zaznamenaný u kritéria „Informační schopnost a vybavenost“, kterému byla přisouzena nejvyšší důležitost ze strany účastníků tvořících expertní stanovisko jednotlivých vah. V tomto případě ji předčila Kuligowska (2015), která se přímo zabývá vyhodnocováním vybavenosti znalostní báze a z hlediska celkových výsledků skončila na druhé příčce. V ostatních hlediscích

se její způsob měření vyrovnával zbývajícím dvěma přístupům a největší propad byl registrován u obchodního hlediska z důvodu absence jeho zohlednění. Třetí místo získal přístup od Maroengsit et al. (2019), který zaujmul přibližně jednu třetinu toho, co preferovaný přístup od Peras (2018) a u dvou kritérií byl označen za nevyhovující. Nejméně použitelným v porovnání s konkurencí se stal přístup od Radziwill a Benton (2017), který je nedostačující z hlediska celkem tří kritérií a jeho největší problém tkví v nepřítomnosti způsobu implementace jím navržených atributů a podrobnějšího rozpracování do dílčích metrik.

Pro provedení analýzy přístupu k měření efektivnosti chatbota byl tedy vybrán přístup od chorvatské autorky D. Peras publikovaný v roce 2018, která působí na Fakultě organizování a informatiky Univerzity Záhřeb. Ten byl spolu s ostatními studii rozebrán v podkapitole 4.2 a jeho další upřesnění lze nalézt v následující kapitole 7.

7 Analýza procesu měření efektivnosti chatbota ve vybraném podniku

Záměrem tohoto úseku aplikační části práce je zanalyzovat proces měření efektivnosti chatbota Radima v ČSOB podle přístupu, který byl vybrán v kapitole 6 a získat praktické poznatky vedoucí k jeho následnému rozšíření a vytvoření uceleného návrhu v souladu s klasickými postupy měření efektivnosti IS/ICT, které byly rozebrány v podkapitole 4.1.2.

Data byla nasbívána pomocí semistrukturovaných rozhovorů, jejichž návod byl konstruován podle zvoleného přístupu, uskutečněných s procesním manažerem, který má v současnosti chatbota Radima v gesci a byl schopen sdělit většinu potřebných informací. Dále bylo čerpáno z interní dokumentace společnosti a softwaru pro administraci daného chatbota.

V kapitole 6 byla provedena komparace celkem čtyř uvažovaných alternativ, z nichž jako vhodný přístup k měření efektivnosti chatbota, byl vybrán pohled od chorvatské autorky D. Peras zveřejněný v roce 2018.

Peras (2018) celý postup hodnocení rozděluje do pěti perspektiv nazvaných:

- perspektiva uživatelské přívětivosti,
- perspektiva vyhledávání (generování) informací,
- lingvistická perspektiva,
- technologická perspektiva,
- obchodní (business) perspektiva.

Každou z nich poté rozčleňuje do kategorií, atributů a vybírá k nim metriky, díky čemuž je rámec měření přehledný a usnadňuje jeho použití.

Daná analýza probíhala z pohledu jednotlivých perspektiv, kdy zjištěná fakta byla přenesena do následujících podkapitol 7.1–7.5.

7.1 Perspektiva uživatelské přívětivosti

Perspektiva uživatelské přívětivosti podle Peras (2018) obsahuje celkem čtyři kategorie, a to: použitelnost, výkon, působení (vliv) a spokojenost. V případě použitelnosti je odkazováno na účinnost a účelnost chatbota ve smyslu jeho snadného použití a času potřebného k dokončení úkolu. Záměrem přiřazených atributů a metrik je vyhodnotit úroveň dokončení úkolu a množství úsilí vynaloženého na jeho dosažení, kdy většina z nich závisí na osobnosti, zkušenosti a nápadech posuzovatele. Výkon chatbota je měřen na základě dokončení úkolu z hlediska úplnosti, rychlosti a vhodnosti. Hodnocení působení (vlivu) chatbota závisí na zkušenostech a emocích, které uživatel při interakci zažívá. Proto je jako vhodný způsob hodnocení voleno dotazníkové šetření nebo anketa. Spokojenost uživatele je založena na porovnání jeho očekávání a reálného výkonu chatbota, k čemuž je doporučeno sledovat dobu trvání a počet interakcí v rámci jedné konverzace a aplikovat hodnotící stupnici.

V ČSOB se hodnocením uživatelské přívětivosti chatbota Radima zabývají pouze interně, a to z toho důvodu, že je do provozu zaveden od května roku 2019 a potřebě externího hodnocení zatím nebyla přikládána taková důležitost.

Před nasazením úplně první verze bylo k internímu hodnocení vybráno několik zaměstnanců, kteří se chatbotem zabývali nebo mohli svými zkušenostmi vhodně přispět. Ti sdíleli veškeré svoje postřehy a dopomohli tak ke konečné podobě chatbota Radima. Z hlediska uživatelské přívětivosti se jednalo o posuzování použitelnosti, výkonu – měření pomocí zátěžových testů, působení (vlivu) – tedy vzhledu chatovacího rozhraní, který je do jisté míry možné konfigurovat ve webové administraci chatbota (např. zobrazení barev) a k provedení větších změn je potřeba zásahu webařů.

Od nasazení chatbota do provozu, probíhá interní hodnocení za pomoci sledování statistik. Je tedy reportován počet dokončených úkolů chatbotem a kolik z nich musel předat (poté je rozlišována servisní a prodejní větev), počet interakcí uživatele v rámci jedné konverzace, počet a délka proběhnutých konverzací, rozložení počtu konverzací v týdenním čase (v jaké dny a časy s chatbotem lidé nejčastěji konverzují) a reakce chatbota na neočekávané situace, kdy úkolem administrátora je, aby zkontroloval vhodnost této odezvy a případně ho naučil přívětivější a výstižnější odpověď

a dále označil barevným štítkem, kterým oddělí neočekávanou situaci od běžné. Tyto metriky jsou užitečné i pro vyhodnocení ostatních perspektiv, a proto se s nimi lze setkat též v rámci nadcházejících podkapitol.

Během výzkumu bylo zjištěno, že daná banka, mimo jiné, pro posuzování, resp. efektivní zjišťování zpětné vazby, i když tento způsob na hodnocení chatbota Radima aplikován nebyl, využívá tzv. metodu Sounding Board, která spočívá v zapojení vybrané skupiny lidí. Walter et al. (2017) píše o velikosti skupiny padesáti lidí, kdy jeden z nich je nezávislý moderátor, který provází celým průběhem setkání, dále odpovědné osoby za hodnocený projekt a účastníci, kteří se k danému projektu vyjadřují. Získané otevřené názory poté mohou vést k velmi dobrým námětům pro změny, i proto, jak autoři uvádějí, byla tato metoda vytvořena v souvislosti s change managementem.

Společnost se tedy svým přístupem k posuzování uživatelské přívětivosti přibližuje návrhu od Peras (2018), kromě nezahrnutí externího hodnocení, jak chatbot na uživatele působí a jak jsou spokojeni s jeho službami. K tomu jsou navrženy následující způsoby od Peras (2018) doplněné o vlastní náměty vytvořené autorkou diplomové práce, kdy se konkrétně jedná o třetí a čtvrtý podbod:

- dotazníkové šetření,
- anketa,
- položení otázky na konci relace chatu týkající se spokojenosti uživatelů,
- hodnotící stupnice – v tomto případě by se přesněji jednalo o systém udělování hvězdiček, jako je tomu u recenzí. Na konci relace chatu by v daný spouštěcí moment (např. když bude chtít uživatel ukončit konverzaci tím, že zavře chatovací okno) vyskočilo: „Ohodnot' mě“ a uživatel by poté chatbotovi udělil počet hvězdiček, který mu podle jeho názoru přísluší. Stejný způsob hodnocení by se dal použít i u systému přisouzení náležitě emotikony podle toho, jaký má uživatel dojem z konverzace s chatbotem.

Z hlediska toho, že se jedná o chat, tak je podle autorky diplomové práce nejúčinnější zavést jednu ze dvou posledních možností. Položení otevřené otázky na konci relace chatu může společnosti přinést širší zpětnou vazbu od uživatelů, na druhou stranu oproti čtvrté možnosti, kde stačí pouze kliknout na hvězdičku či emotikonu, může být

pro uživatele vypisování mírně obtěžující, a tak svou odpověď odbydou neurčitým a jednoslovným vyjádřením či raději vůbec. Kdežto kliknutí je rychlé, jednoduché a okamžitě vyjadřující míru spokojenosti.

7.2 Perspektiva vyhledávání (generování) informací

Tato perspektiva vyhodnocuje, jak dobře společnost naplňuje informační potřeby uživatelů, resp. jak efektivně chatbot vyhledává či generuje outputy a je rozdělena do tří kategorií, kterými jsou: přesnost, srozumitelnost a účinnost. Pozornost je tedy zaměřena na získávání relevantních a vhodných informací ze znalostní báze chatbota. Dané metriky jsou kvantitativního charakteru a hodnotí úroveň přesnosti a patřičnosti poskytovaných informací a schopnosti dokončit úkol bez vynaložení nadměrných zdrojů.

Jak bylo již zjištěno v kapitole 6, tak kritérium informační schopnost a vybavenost bylo shledáno nejdůležitějším, a právě toto kritérium je obsahem perspektivy vyhledávání (generování) informací. Zájmem podniku tedy je, aby chatbot plnil především svoji informační funkci a byl co nejefektivnější v tomto směru.

Odpovědnou osobou za vybavenost znalostní báze chatbota Radima je její administrátor, který denně kontroluje výstupy produkované chatbotem a učí ho novým slovním spojením a výrazům tak, aby byl co nejpřívětivější pro uživatele (viz 7.1.1) a hlavně uspokojoval jejich informační potřeby. Jak bylo uvedeno v podkapitole 5.2, tak chatbot Radim nabízí dvě možnosti, jak s ním vést dialog, a to formou konverzačního stromu a free textu. Implementace konverzačního stromu je založená na výběru z možností v podobě odpovědních tlačítek, na základě kterých probíhá interakce mezi jím a uživatelem, který se postupně proklikává k tomu, co hledá. Konverzační strom je zatím plně v kompetenci vývojáře chatbota Radima, a tudíž do něj administrátor nezasahuje. Ten ovlivňuje vybavenost znalostní báze, ze které chatbot čerpá znalosti při produkování výstupů formou free textu. Ta se skládá z relevantních Q&A (zaměřené na úkol – půjčky a vše s tím spojené) a obecných Q&A (doplňují plnohodnotnou lidskou konverzaci (např. pozdravy) a základní informace o společnosti). Databáze se doplňuje velice jednoduše, a to importem souboru s příponou .xls a .xlsx či vytvořením jednotlivých záznamů přímo v administraci. Způsob vložení závisí na množství nových Q&A, kdy v případě, že jich je hodně

(např. více než 10), tak je efektivnější zpracování a import souboru s .xls či .xlsx příponou a naopak.

Přesností, srozumitelností a účinností výstupů vytvářených chatbotem se denně zabývá jeho administrátor, který, jak už bylo uvedeno výše, je jeho učitelem. Z tohoto hlediska je Radim schopen porozumět jednoduchým překlepům (záměna jednoho písmene), se složitějšími zkomoleninami si neporadí a slangovým výrazům je potřeba ho naučit zvlášť v tzv. slangovém slovníku. Aby chatbot reagoval na vstup od uživatele co nejpřesněji, nejpřirozeněji a nepředával chybné informace, vyhledává nejvhodnější výstup podle shody a vah slov umístěných v uživatelem zadaném inputu (v případě free textu), resp. čím častěji je učené slovo (nachází se v jeho znalostní bázi ve vzorových otázkách), tím je pro něj méně významné (stává se běžným) a je schopen si s ním lépe poradit. S tím je spojena metrika míra jistoty chatbota, která se odvíjí od vybavenosti jeho znalostní báze, což znamená, že čím více Q&A bude umět (v souvislosti s významností učených slov), tím bude vyhledávat úspěšněji. Přesněji tato metrika značí průměrnou míru jistoty chatbota plynoucí ze všech zpráv a těch které se nachází nad prahovou hodnotou (tj. když neodpověděl, že neví) v daný den. Hodnota se pohybuje v rozmezí intervalu od 0 do 1 (nejistý-jistý). V případě nevědomosti se omluví a odvětví, že neví (aby konverzace působila co nejpřirozeněji, tak bot vybírá z více různých odpovědí tohoto typu) či nabídne zanechání telefonního kontaktu živému operátorovi, který za něj konverzaci (úkol) dokončí. Pro posuzování vybavenosti znalostní báze je tedy vhodné pozorovat tuto hodnotu, kdy cílem by mělo být její zvyšování. Dále je v rámci srozumitelnosti sledován poměr nevědomosti, resp. kolikrát chatbot nedokázal odpovědět vůči celkovému počtu volání Q&A modulu. Účinnost chatbota, v rámci této perspektivy, určuje rychlost, za kterou je schopen zpracovat vstup od uživatele a vyprodukovat mu adekvátní výstup, počet interakcí od uživatele v rámci jednoho vlákna a délka této konverzace, což při bližším zkoumání ukazuje, za jak dlouho byl chatbot schopen uspokojit jeho informační potřebu.

Společnost při vyhodnocování a zajišťování schopnosti chatbota efektivně vyhledávat a produkovat výstupy využívá metriky navržené od Peras (2018) a vzhledem k tomu, že se této oblasti hojně věnuje, což je z důvodu zvoleného způsobu implementace chatbota vyžadováno, tak v rámci práce nevznikají žádné další náměty a doporučení.

7.3 Lingvistická perspektiva

Lingvistická perspektiva se zabývá analýzou úrovně přesnosti a schopnosti chatbota vytvářet vhodnou odezvu z pohledu lingvistiky a skládá se z pěti kategorií, kterými jsou: kvalita, kvantita, relevance, způsob jednání a gramatická přesnost. Peras (2018) k hodnocení prvních čtyř kategorií navrhuje použít Likertovu škálu a změřit tak úroveň dosahování Griceových maxim (vysvětlených v podkapitole 4.2), kterou je vhodné použít u posuzování více lidmi. Pátá kategorie hodnotí přijatelnost odezvy chatbota z gramatického a významového hlediska a k jejímu změření je kompetentní tým expertů.

S chatbotem Radimem lze komunikovat dvěma způsoby, a to za pomoci odpovědních tlačítek (předem vytvořený scénář v podobě konverzačního stromu) a formou free textu (vyhledávání nejrelevantnějšího výstupu podle klíčových slov ze sady možných Q&A) a negeneruje tedy úplně nové vlastní výstupy. To vede k tomu, že není potřeba měřit efektivnost daného chatbota z hlediska lingvistiky, tak jako se tomu ve svém přístupu věnuje Peras (2018), protože už je dopředu známo, že bot vyprodukuje takové odezvy, jaké ho naučí lidský činitel, tedy administrátor a ty jsou před samotným učením už dopředu ošetřené.

Role administrátora je sice odpovědná za vybavenost znalostní báze chatbota Radima, ovšem úprava a tvorba nových odpovědí podléhá následujícímu schvalovacímu procesu (viz schéma 2), který přenáší odpovědnost na více kompetentních osob.

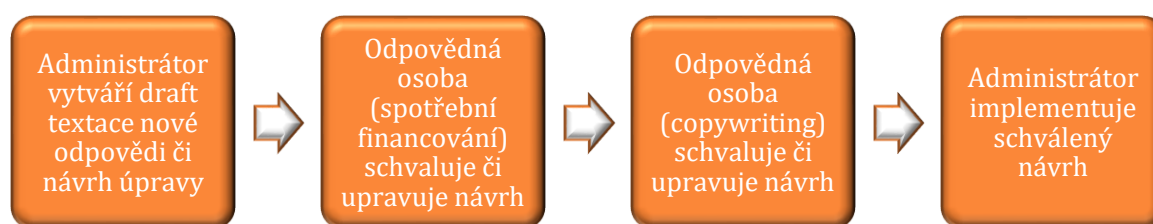


Schéma 2: Proces schvalování úpravy a tvorby textace nových odpovědí

Zdroj: vlastní zpracování podle interní dokumentace podniku

Z hlediska lingvistiky v tomto procesu hraje hlavní roli copywriter, jehož náplní práce je psaní textů. Základním předpokladem výkonu této činnosti je perfektní znalost pravopisu a gramatiky jazyka, ve kterém daná osoba texty píše (zde český jazyk).

Další zásadou je, aby byl dostatečně kreativní a dokázal odhadnout, kdy se hodí použít jaký typ mluvy. Chatbot Radim působí na public webu ČSOB a má na klienty působit příjemně, přirozeně a zároveň poutavě a zábavně. Cílem je, aby jeho odpovědi vyhovovaly každé věkové kategorii, a proto produkuje výstupy, které se nachází na pomezí formální a neformální češtiny. O tom, jak moc je formální chatbotova reakce rozhoduje konkrétní situace. Výstupy jsou ale vždy gramaticky korektní a bez pravopisných chyb, za což se schvalovací proces zaručuje. V případě, že by ale přece jenom k nějaké chybě došlo (např. překlep), tak ji odhalí tzv. spellchecker či později administrátor.

Peras (2018) v rámci této perspektivy řeší, zda jsou vytvářené odezvy hodnoceny jako správné, relevantní a dochází k jejich kategorizaci. Jak už bylo uvedeno v předchozí podkapitole 7.2, tak znalostní báze chatbota se skládá z relevantních a obecných Q&A, kdy nově vytvořené před jejich importem prochází právě již zmiňovaným schvalovacím procesem (viz schéma 2). Dále je sledována adekvátnost informací, které chatbot v rámci jeho výstupu produkuje, kdy se Peras (2018) odvolává na teorii Griceových maxim, podle kterých by informace měla poskytovat to co je od ní v daný okamžik žádáno a zároveň by neměla dávat více než je od ní očekáváno. Tato úloha též přímo spadá pod schvalovací proces zobrazený ve schématu 2. S tím souvisí i jednoznačnost odpovědí, která je také zajišťována při jejich tvorbě a odsouhlasení.

Peras (2018) doporučuje měřit kvalitu, kvantitu, relevanci odezvy a způsob jednání chatbota za pomoci Likertovy škály. Tato metrika se v bance nepoužívá, protože její implementace je vhodná při posuzování více lidmi, a to ve formě dotazníků, anket či vytváření stanovisek. Uvedený schvalovací proces (viz schéma 2) se dotýká pouze jednotlivých osob.

7.4 Technologická perspektiva

Technologická perspektiva řeší pouze jednu kategorii, a to lidskost, resp. zda se chatbot chová jako člověk. Tím je myšleno, že by neměl být schopen pouze zpracovávat přirozený jazyk, ale také dokázat porozumět a sám generovat odezvu na základě předchozí zkušenosti. Navrženými metrikami je Turingův test, hodnotící stupnice či sledování statistických ukazatelů.

Jak už bylo v práci uvedeno, tak chatbot Radim se řídí předem definovanými pravidly a vhodnou odezvu hledá v sadě dopředu vytvořených a odpovědnými osobami ošetřených Q&A. Je závislý na administrátorovi, který mu vytváří jeho znalostní bázi a přeučuje ho. Jednou z navržených metrik je tzv. Turingův test, o němž bylo zmiňováno v podkapitole 3.1 a jímž dosud žádný chatbot neprošel. Chatbot Radim nedokáže generovat svoje vlastní výstupy, je orientovaný na konkrétní úkol a chtěným stavem je možnost plně regulovat jeho chování. I proto při stanovování vah jednotlivých kritérií pro výběr vhodného přístupu k měření efektivnosti chatbota (viz 6.2) došlo k zjištění, že vybrané zainteresované osoby, v této záležitosti zastupující společnost, vyhodnocování lidskosti přikládají nejnižší důležitost. Podnik se jeho jednání snaží přiblížit k lidskému neustálým zajišťováním vybavenosti znalostní báze. V tomto případě platí, že čím více toho chatbot ví, tím více bude působit jako člověk. S tím souvisí jeho slovní zásoba a hlavní metrika, která je v tomto směru v ČSOB sledována – míra jeho jistoty. Tato metrika byla rozebírána již v souvislosti s perspektivou vyhledávání (generování) informací (viz 7.1.2) a vyjadřuje tedy míru jistoty chatbota na základě jeho schopnosti vyhledávat nejrelevantnější shodu ze sady Q&A, resp. jak si byl chatbot průměrně ze všech zpráv a těch které se nachází nad prahovou hodnotou (tj. když neodpověděl, že neví) v daný den jistý. Což znovu vede k tomu, že čím rozšířenější znalostní bázi chatbot má (v součinnosti s významností daných klíčových slov), tím je zvyšována jistota korektnosti jím zvolených odezev. Dále je jeho znalostní báze rozdělena na relevantní a obecné Q&A, aby jeho reakce působily co nejpřirozenějším dojmem a v rámci toho je sledován poměr nenalezených odezev vůči celkovému počtu volání Q&A modulu.

7.5 Obchodní (business) perspektiva

Aby byl podnik schopen registrovat, zda mu chatbot přináší nějakou přidanou hodnotu, je potřeba pravidelně vyhodnocovat obchodní (business) perspektivu. Tu Peras (2018) definuje jako rozdíl mezi produkcí a časem, který je vnímán jako náklad. Autorka pro měření produkce chatbota navrhuje použít metriky typu: počet uživatelů, celková a průměrná doba trvání konverzace, počet proběhnutých konverzací a pro zjištění nákladů: počet agentů zahrnutých do konverzace (chatbot + živý operátor), celková a průměrná doba trvání konverzace s agentem, počet neúspěšných konverzací, počet nevhodných reakcí a opakovaných dotazů.

Na základě výzkumu provedeného v kapitole 6, přesněji 6.2, došlo k určení expertního stanoviska, podle kterého je vyhodnocování obchodního hlediska (v tomto přístupu obchodní (business) perspektivy) považováno za druhou nejdůležitější oblast při zjišťování efektivity chatbota.

V podniku jsou sledovány následující metriky, na základě kterých lze tuto perspektivu podle Peras (2018) vyhodnotit:

- počet konverzací za den,
- celkový počet konverzací (od prvního dne zavedení do provozu až do vybraného dne),
- průměrná délka konverzace za den,
- počet konverzací, které vedly ke sjednání půjčky,
- počet konverzací, které chatbot nedokázal dokončit sám (předal uživateli telefonní kontakt živému operátorovi),
- doba trvání konverzace s živým operátorem.

Možnost konverzace s živým operátorem je omezena časem, kdy v případě, že chatbot není schopen uspokojit potřebu uživatele z důvodu nedostatečných znalostí, tak od něj předává telefonní číslo živému agentovi, který se mu ozve v otevírací době útvaru zajišťující servis (7–22 h, 8–22 h). Pokud se jedná o sjednání půjčky, tak je distribuován kontakt agentovi, který zabezpečuje spotřební financování (8–19 h, 9–17 h, 10–18 h). Z hlediska měření efektivity chatbota je důležité brát v úvahu pouze čas daného pracovníka, který je vnímán jako náklad a lze na základě toho poté podle rozdílu produkce a nákladů určit přidanou hodnotu, kterou chatbot znamená pro společnost. Počty neúspěšných, nevhodných i opakovaných dotazů lze zjistit za pomoci štítkování a následného filtrování v administraci chatbota Radima.

K vyhodnocování obchodní (business) perspektivy, podle které je určena přidaná hodnota chatbota pro společnost, využívá ČSOB stejné metriky, které Peras (2018) navrhuje. Mimoto je pro banku navíc relevantní rozlišovat podíl konverzací vznikajících v a mimo bankovní síť, aby věděla, které konverzace proběhly z důvodu testování na produkčním prostředí a které inicioval stávající či potenciální klient. Dále je sledováno rozložení počtu konverzací v týdenním čase, které indikují na dny a časy, v nichž je chatbot ze stran uživatelů nejvíce vytížen. Konkrétně chatbot Radim

podle reportu za poslední rok čili od jeho nasazení do provozu, nejvíce konverzuje v pondělí, úterý a středu v odpoledních hodinách (tzn. 12–18 h) a nejméně o víkendu. Tyto všechny statistiky mohou být ovšem ovlivněny sezónními vlivy, kdy např. v období letních prázdnin či před Vánoci bylo jeho využití vyšší než v průběhu ostatních měsíců, což celkově koresponduje se zvýšením sjednávání neúčelových půjček. Zda se ale skutečně jedná o sezónní vliv je předběžné tvrzení, které lze potvrdit až po analýze dlouholetých dat s periodicitou jednotlivých sledování kratších než jeden rok (měsíc, kvartál apod.). Peras (2018) ve svém přístupu nezmiňuje sledování statistik v čase, což lze na základě této analýzy považovat za negativum. Aby mohla společnost určit, zda je pro ni chatbot opravdovým přínosem, je zásadní sledovat jeho vývoj od začátku zavedení, k čemuž je vhodné vytvořit časové řady a následně je analyzovat. Z chatbota Radima se výše zmíněné statistiky přímo sbírají a propisují do softwaru pro jeho administraci, ve které jsou vytvářeny časové řady v grafickém vyjádření. Hodnoty daných ukazatelů závisí na délce časového intervalu, a proto je při jejich analýze nutné odstraňovat kalendářní variace (musí být srovnávány stejně dlouhé časové úseky).

8 Shrnutí a diskuze výsledků

S cílem diplomové práce bylo přímo spjata hledání odpovědi na stanovenou výzkumnou otázku V1: *Jak je možné měřit efektivnost chatbota?*

Aby bylo možné na tuto otázku odpovědět, byly v teoretické části v podkapitole 4.2 představeny dostupné přístupy k měření efektivnosti chatbota, získané z odborných vědeckých databází. Problémem těchto přístupů je, že se jedná o málo řešenou problematiku, a tak ztrácejí na komplexnosti. Z tohoto důvodu byla provedena vícekritériální analýza v kapitole 6, v rámci které byla definována:

- kritéria, resp. oblasti, jejichž zohlednění by nemělo ve vhodném přístupu k měření efektivnosti chatbota chybět,
- váhy (stanovené za pomoci Saatyho metody párového porovnání a vnímané důležitosti expertů v oboru z KC ČSOB),
- alternativy (přístupy k měření efektivnosti chatbota).

Použitelnými, resp. zohledňujícími alespoň čtyři z pěti definovaných kritérií, byly shledány čtyři přístupy, které udávají, *jak je možné měřit efektivnost chatbota*. Autory těchto přístupů jsou: Kuligowska (2015), Maroengsit et al. (2019), Peras (2018), Radziwill a Benton (2017). Ty byly poté porovnány v analytickém nástroji Expert Choice 2000 sloužícím pro podporu rozhodování, a to v závislosti na daných kritériích a jejich vahách, jimž byla prisouzena významnost dle následující posloupnosti (seřazeno od nejvyšší po nejnižší důležitost): informační schopnost a vybavenost, obchodní hledisko, uživatelská přívětivost, jazyková úroveň a vybavenost, lidskost. Na základě výsledků této analýzy byl vybrán nejvhodnější a nejkomplexnější přístup od chorvatské autorky D. Peras publikovaný v roce 2018. Ta efektivnost chatbota měří z pohledu pěti perspektiv, kterými jsou: perspektiva uživatelské přívětivosti, perspektiva vyhledávání (generování) informací, lingvistická perspektiva, technologická perspektiva a obchodní (business) perspektiva. Každou z nich poté rozčleňuje do kategorií, atributů a vybírá k nim metriky, díky čemuž je rámec měření přehledný a usnadňuje jeho použití. Tento postup byl poté autorkou diplomové práce přenesen do následujícího schématu 3 v podobě sestupného procesu.

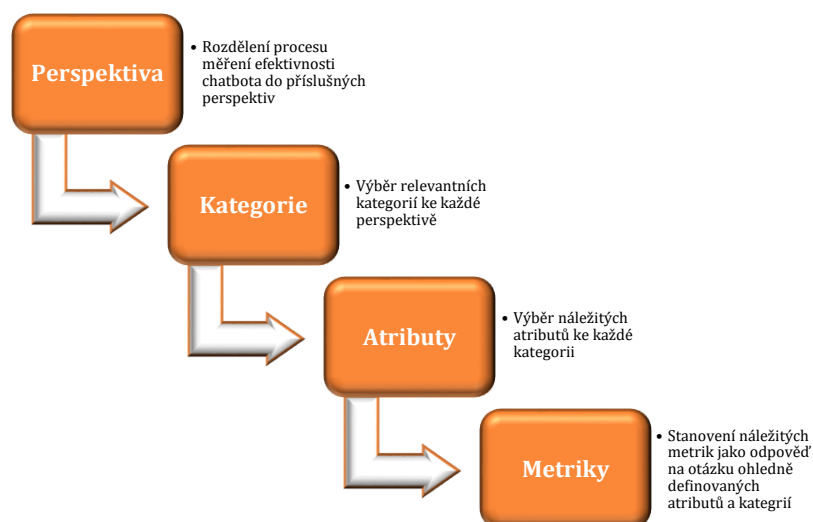


Schéma 3: Dekompozice rámce vybraného přístupu k měření efektivity chatbota

Zdroj: vlastní zpracování podle Peras (2018)

S tímto přístupem bylo dále zevrubněji pracováno, načež daná zjištění jsou shrnuta a diskutována v nadcházejících podkapitolách.

8.1 Shrnutí analýzy procesu měření efektivity chatbota ve vybraném podniku

Podle vybraného přístupu od Peras (2018) byla provedena analýza procesu měření efektivity chatbota Radima v ČSOB zaměřeného na konkrétní úkol, kterým je poskytování zákaznického servisu a podpory pro sjednávání neúčelových půjček.

Ve zkoumané společnosti nepřístupují k dílčím částem (perspektivám) stejně komplexním způsobem jako Peras (2018). Při rozebírání jednotlivých navržených metrik bylo zjištěno, že daná banka největší úsilí vkládá do vybavování znalostní báze chatbota, což má vést k jeho stále se zvyšující samostatnosti. Uživatelskou přívětivost vyhodnocuje pouze interně a nikoli externě. Podnik prozatím nezjišťuje spokojenost uživatelů se službami, které chatbot poskytuje a nedostává tak zpětnou vazbu z vnějšího prostředí. Vzhledem k tomu, že chatbot Radim negeneruje vlastní výstupy a obsah jeho odezev prochází jistým schvalovacím procesem (viz schéma 2), není potřeba věnovat pozornost měření jeho efektivity z hlediska lingvistiky, protože je zde spoléháno pouze na lidského činitele. Technologická perspektiva se zabývá schopností chatbota chovat se jako člověk, kdy je sledována míra jeho jistoty a poměr nenalezených odezev vůči celkovému počtu volání Q&A modulu. Poslední

analyzovanou perspektivou byla tzv. obchodní (business), na jejímž základě je zjišťována přidaná hodnota, kterou tato technologie vytváří. Banka v rámci ní pozoruje všechny ukazatele, které daný přístup nabízí, spolu s některými dalšími, pro ni relevantními, jako je například podíl konverzací vznikajících v a mimo bankovní síť. K jejich sledování využívá časových řad přenesených do grafického znázornění.

Na základě všech získaných poznatků byly podle vybraného přístupu zpracovány souhrnné tabulky 2 a 3.

Perspektiva	Kategorie	Atributy	Metriky	Charakter metrik
Perspektiva uživatelské přívětivosti	Otázka: <i>„Jak je chatbot uživatelsky přívětivý, resp. použitelný, výkonný, působící a uspokojující?“</i>			
	Použitelnost	Úroveň dokončení úkolu Využívání chatbota k asistenci či získání informací Četnost výskytu daného typu odezvy	Průměrný čas, za který je chatbot schopen splnit svůj úkol (počet minut a sekund) Celkový a průměrný počet interakcí uživatele v rámci jedné konverzace Relativní četnost výskytu daného typu odezvy Pravděpodobnost výskytu daného typu odezvy Procentní vyjádření shody (vytvářených odezev) Hodnotící škála (5tíhvězdičkový systém, hodnocení pomocí emotikony) Dotazníkové šetření, ankety Fokusní skupiny (hodnocení založené na názoru určité skupiny lidí)	Kvalitativní, kvantitativní
	Výkon	Robustnost Odezvy v nečekaných situacích Koherence Efektivní alokace úkolu	Procentní vyjádření úspěšné odezvy v nečekaných situacích (jiná odpověď než, že neví apod.), tzn. udržování souvislé konverzace Okamžité předání úkolu na živého agenta (počet minut a sekund) Hodnotící škála (5tíhvězdičkový systém, hodnocení pomocí emotikony)	Kvalitativní, kvantitativní
	Působení (vliv)	Osobnost Emoce Zábava Oddanost Osobnostní rysy Potřeba lidské asistence Důvěryhodnost Ochrana soukromí	Hodnotící škála (5tíhvězdičkový systém, hodnocení pomocí emotikony) Dotazníkové šetření, ankety Fokusní skupiny (hodnocení založené na názoru určité skupiny lidí) Kontrola klíčových slov Celkový počet konverzací Celkový a průměrný počet interakcí v rámci jedné konverzace Celková a průměrná doba trvání konverzace (počet minut a sekund) Celkový a průměrný počet konverzací, ve kterých byl předán osobní údaj uživatele	Kvalitativní, kvantitativní
	Spokojenost	Očekávání Dojem Znalost Navigace Oddanost Zábava Zvědavost Sociální vztahy Schopnost učit se Schopnost pomoci	Celková a průměrná doba trvání konverzace (počet minut a sekund) Počet interakcí v rámci jedné konverzace Dotazníkové šetření, ankety Hodnotící škála (5tíhvězdičkový systém, hodnocení pomocí emotikony)	Kvalitativní, kvantitativní

Tabulka 2: Souhrnný rámec vybraného přístupu k měření efektivnosti chatbota (1. část)

Zdroj: vlastní zpracování podle Peras (2018)

Perspektiva	Kategorie	Atributy	Metriky	Charakter metrik
Perspektiva vyhledávání (generování) informací	Otázka: <i>„Jakou přidanou hodnotu (užitek) přináší chatbot pro uživatele, resp. jak jsou jeho výstupy přesné, vhodné a účinné?“</i>			
	Přesnost	Schopnost pracovat s různými jazykovými variacemi	Přesnost výstupů Celkový a průměrný počet volání znalostní báze Schopnost rozeznat překlepy, pravopisné chyby a synonyma	Kvantitativní
	Srozumitelnost	Schopnost detekovat význam a záměr vstupu a vhodně reagovat	Citlivost na kontext Procentní vyjádření úspěšné detekce a reakce Celkový a průměrný počet nevhodných odpovědí (reakcí) Poměr počtu opravovaných výstupů <i>Míra jistoty chatbota</i>	Kvantitativní
	Účinnost	Jak dobře jsou zdroje použity k dosažení cílů (dokončení úkolu)	Celkový a průměrný počet shod nalezených ve znalostní bázi chatbota Měření času volání znalostní báze chatbota (počet sekund) Celková a průměrná doba trvání konverzace (počet minut a sekund) Celkový a průměrný počet interakcí uživatelů Celkový počet interakcí na jeden úkol (v rámci jedné konverzace)	Kvantitativní
Lingvistická perspektiva	Otázka: <i>„Jaká je kvalita, kvantita, relevance, jednoznačnost a gramatická přesnost výstupů, které chatbot produkuje?“</i>			
	Kvalita	Správnost odpovědí Kategorizace odpovědí	Likertova škála – hodnotí tým expertů	Kvalitativní
	Kvantita	Adekvátnost (přiměřenost) informace	Likertova škála – hodnotí tým expertů	Kvalitativní
	Relevance	Relevance odpovědí vzhledem ke kontextu konverzace	Likertova škála – hodnotí tým expertů	Kvalitativní
	Způsob jednání	Jednoznačnost odpovědí	Likertova škála – hodnotí tým expertů	Kvalitativní
Gramatická přesnost	Přijatelnost z gramatického a významového hlediska	Celkový počet chyb v jedné periodě chatu Analýza slovní úrovně (rozsah slovní zásoby, pravopis, velká/malá písmena) – hodnotí tým expertů Analýza gramatické úrovně (podstatná jména, zájmena, slovesa, slovosled) – hodnotí tým expertů	Kvalitativní, kvantitativní	
Technologická perspektiva	Otázka: <i>„Chová se chatbot jako člověk?“</i>			
	Lidskost	Přirozenost Udržování téma diskuze Odpovídání na specifické dotazy Míra (ne)porozumění	Turingův test Hodnotící stupnice Procentní vyjádření (ne)porozumění <i>Míra jistoty chatbota</i>	Kvalitativní, kvantitativní
Obchodní (business) perspektiva	Otázka: <i>„Jakou přidanou hodnotu přináší chatbot společnosti?“</i>			
	Přidaná hodnota	Produkce Náklady	Celkový a průměrný počet uživatelů Celková a průměrná doba trvání konverzace (počet minut a sekund) Celkový a průměrný počet konverzací Celkový počet agentů zahrnutých do konverzace Celková a průměrná doba trvání konverzace s jiným agentem (počet minut a sekund) Celkový a průměrný počet konverzací s jiným agentem Počet neúspěšných konverzací Počet nevhodných odpovědí (reakcí) Počet opakovaných dotazů	Kvantitativní

Tabulka 3: Souhrnný rámec vybraného přístupu k měření efektivity chatbota (2. část)

Zdroj: vlastní zpracování podle Peras (2018)

Souhrnný rámec v tabulkách 2 a 3 byl rozšířen o poznatky získané nebo navržené autorkou diplomové práce na základě analýzy procesu měření efektivity chatbota Radima v ČSOB podle vybraného přístupu. Ty byly odlišeny hnědou barvou a metrika,

kteřá byla shledána opomenutelnou, byla označena barvou šedou. Jedná se tedy o následující rozšíření či připomínkování daného přístupu.

Perspektiva uživatelské přívětivosti:

- externí hodnocení založené na pětihvězdičkovém systému či přisouzení náležitě emotikony – vyjadřuje míru uživatelské přívětivosti chatbota, kterou tvoří průměrný počet všech hodnocení (čím vyšší počet hvězdiček či pozitivních emotikon je chatbotovi uživateli uděleno, tím je pro ně přívětivější),
- využití tzv. fokusních skupin – metoda hodnocení založená na názoru a nápadech určité skupiny lidí, kdy mohou být i větší skupiny (tzv. Sounding Board), jejíž výsledky dokážou za krátký čas ovlivnit celkovou budoucí podobu chatbota,
- celkový a průměrný počet konverzací, ve kterých byl předán osobní údaj uživatele – tato metrika poukazuje na atributy důvěryhodnost a ochrana soukromí (čím více osobních údajů uživatel chatbotovi předá, tím věrohodněji působí, a naopak chatbot musí respektovat nařízení GDPR, což znamená, že se nutně neptá na informace, které nejsou potřebné pro splnění úkolu a respektuje soukromí uživatele).

Perspektiva vyhledávání (generování) informací:

- míra jistoty chatbota – vyjadřuje s jakou jistotou chatbot, při vyhledávání nejrelevantnější odpovědi ze sady Q&A, volí daný výstup. Tato metrika je založená na vybavenosti jeho znalostní báze, resp. čím více vzorových Q&A umí (v souvislosti s významností učených slov), tím se zvyšuje jistota nalezení adekvátní odezvy a uspokojení potřeb uživatele. Hodnota se pohybuje v rozmezí intervalu od 0 (jistý) do 1 (nejistý). Tuto metriku lze poté sledovat jako průměrnou hodnotu k danému dni, která se skládá ze všech zpráv a těch, které se nacházejí nad prahovou hodnotou (tj. když chatbot nenachází shodu a odpovídá, že neví).

Technologická perspektiva:

- na základě teoretických i praktických východisek byla zpochybněna účinnost Turingova testu, který je vhodné implementovat pouze v případě obecné umělé

intelligence. Tímto testem prozatím žádný chatbot neprošel, a tudíž není potřebné ho zařazovat do portfolia metrik,

- míra jistoty chatbota – s čím vyšší jistotou je chatbot schopen vyhledat a vracet výstupy (v souvislosti s vybaveností znalostní báze), tím lidštějším dojmem jeho reakce působí.

Doporučení pro všechny perspektivy:

- analýza časových řad – dané ukazatele je vhodné sledovat v čase s jejich grafickým vyjádřením (vyhovující je sloupcový a spojnicový graf), kdy se jedná o intervalové časové řady, u kterých je potřeba odstraňovat kalendářní variace a pro zjištění sezónních vlivů je nutné analyzovat dlouholetá data s periodicitou jednotlivých sledování kratších než jeden rok (měsíc, kvartál apod.).

8.2 Diskuze vybraného přístupu

Ve vybraném přístupu se nacházejí pouze nefinanční měřitelné a kvalitativní ukazatele, které jsou koncipovány v souladu s dosažením očekávaných (chtěných) efektů chatbota. Těmito efekty jsou myšleny cíle každé z perspektiv, tzn., aby byl chatbot uživatelsky přívětivý, dokázal během konverzace uspokojit především informační potřeby uživatelů, produkoval pouze relevantní a úplné výstupy, působil co nejméně strojovým dojmem a zároveň vytvářel společnosti přidanou hodnotu. Diplomová práce se zabývala výše uvedenými metrikami, ovšem, aby podnik dokázal vnímat investici do chatbota jako vydařenou, tak potřebuje vidět i nějakou její finanční návratnost. V teoretické části byly v rámci podkapitoly 4.1.2 zmíněny ukazatele efektivnosti investic ROI, TCO, NPV, kdy daní autoři uvádějí, že je prostřednictvím nich často indikováno na nevýhodnost investic do IS/ICT či je těžké je odhadnout a mají tedy nízkou vypovídací schopnost. To ale neznamená, že by tyto ukazatele měly být opomíjeny, jenom je upozorňováno na to, že pokud v období projektu nevycházejí jejich nejpříznivější hodnoty, nemusí to nutně znamenat, že se investice opravdu nevyplatí. Příčinou může být nezahrnutí přesných hodnot všech vstupů a výstupů potřebných pro jejich výpočet, a to z důvodu, že u vyhodnocování přínosů IS/ICT hrají podstatnou roli právě již zmiňované nefinanční měřitelné a kvalitativní ukazatele. Toto znovu odkazuje na metriky navržené ve vybraném přístupu s tím, že i když jmenované ukazatele efektivnosti investic nebyly předmětem zkoumání diplomové

práce, jedná se o osvědčené ukazatele, nad jejichž použitím by mělo být uvažováno i v tomto konkrétním případě (chatbot by měl v budoucnosti uspořít plánovaný počet FTE).

Přístup od Peras (2018) nabízí již rozebraný rámec, ale nenavrhuje celý formální postup. Aby byl rámec aplikovatelný, musí už od počátku zahrnovat veškeré dílčí části vedoucí k výsledku. Prvním krokem je podle Molnár (2000) definování cílů, kterých má být zavedením technologie (chatbota) dosaženo a na základě nich by organizace měla uvažovat nad možnými (očekávanými) přínosy (pro uživatele i společnost). Spolu s tím by měl být určen i zodpovědný manažer. Podle Učeň a kol. (2001), jak bylo zmíněno v podkapitole 4.1.1, je potřeba nejdříve zaměřit svou pozornost na účelnost (*měřit správné věci*), tedy určit které perspektivy je pro daný podnik relevantní vyhodnocovat a poté se soustředit na účinnost (*měřit věci správně*), k čemuž je vhodné použít rozšířený souhrnný rámec (viz tabulka 2 a 3) a postupovat podle schématu 3. Vzhledem k různým typům (viz 3.3) a důvodům implementace (viz 3.4) chatbota je vhodné si ještě před zahájením volby perspektiv uvědomit, jaký chatbot bude posuzován a poté mu zbytek procesu přizpůsobit.

Dále bylo během analýzy procesu měření efektivnosti chatbota Radima v ČSOB podle přístupu od Peras (2018) povšimnuto, že se napříč i v rámci jednotlivých perspektiv některé metriky opakují. Například celková a průměrná doba trvání konverzace ukazuje na kategorie: spokojenost uživatelů, účinnost z hlediska předání informací, resp. dokončení úkolu a přidané hodnoty pro společnost. Proto je vhodné dané metriky vybírat podle principu popsaném ve schématu 3 a věnovat jim pozornost zvláště vzhledem k dosažení cíle každé z perspektiv a zároveň by za správnost jejich měření měly jednoznačně odpovídat určené osoby.

Na základě všech získaných teoretických a praktických poznatků bylo poté možné sestavit postup vedoucí k měření a vyhodnocování efektivnosti chatbota čili rozšířit vybraný přístup. Jeho návrh je rozdělen do dvou etap, a to etapy plánování a následné kontroly, které jsou charakterizovány nadcházejícími kroky.

Etapa plánování zavedení chatbota do podniku:

1. Definování cílů, kterých má být zavedením chatbota dosaženo – cíle by měly splňovat podmínky SMART.
2. Uvažování nad očekávanými přínosy (podle účelu užití chatbota (viz 3.4) – rozšíření a zkvalitnění služeb zákazníkům, rostoucí počet spokojených a získání nových zákazníků, úspora času zaměstnanců, podpora nehlasových interakcí, podpora prodeje či jeho přímé zvyšování apod.) a současné určení zodpovědného manažera.
3. Zohlednění typu chatbota – zda se jedná o vyhledávací či generativní model a podle toho uzpůsobit výběr vhodných perspektiv.
4. Na základě předchozích kroků zvolit vhodné perspektivy (z pěti rozebraných), s kterými jsou vázány dílčí cíle (co chce podnik v dané perspektivě zjišťovat, jaké efekty ho zajímají) a přiřadit jim priority.
5. V rámci každé z vybraných perspektiv je poté potřeba postupovat podle schématu 3 – výsledkem je portfolio metrik (s určením dimenze (jednotka měření), výchozí a cílové (chtěné) hodnoty a zdroje dat pro měření).

Etapa následné kontroly:

6. Sledování metrik – stanovené ukazatele je za pomoci příslušného softwaru potřeba pravidelně sledovat, vč. vytvoření intervalových časových řad, nejlépe s jejich grafickým vyjádřením (vhodný je sloupcový a spojnicový graf), kdy vlastníci metrik (zodpovědný manažer) musí mít od odpovědných osob k dispozici sledované hodnoty v předem dohodnutých lhůtách.
7. Vyhodnocení – spolu s šestým krokem zjišťovat současný stav daných ukazatelů, registrovat změnu a porovnávat ji s chtěnou.
8. Zhodnocení výsledné efektivnosti chatbota – vyvozování závěrů o dosahování definovaných cílů z prvního kroku (pokládání otázek, zda chatbot účinně splňuje svůj účel a pokud ne, tak co s tím lze dělat) a případné úpravy metrik.

Pro rychlejší orientaci byly tyto kroky převedeny do podoby následujícího schématu 4.

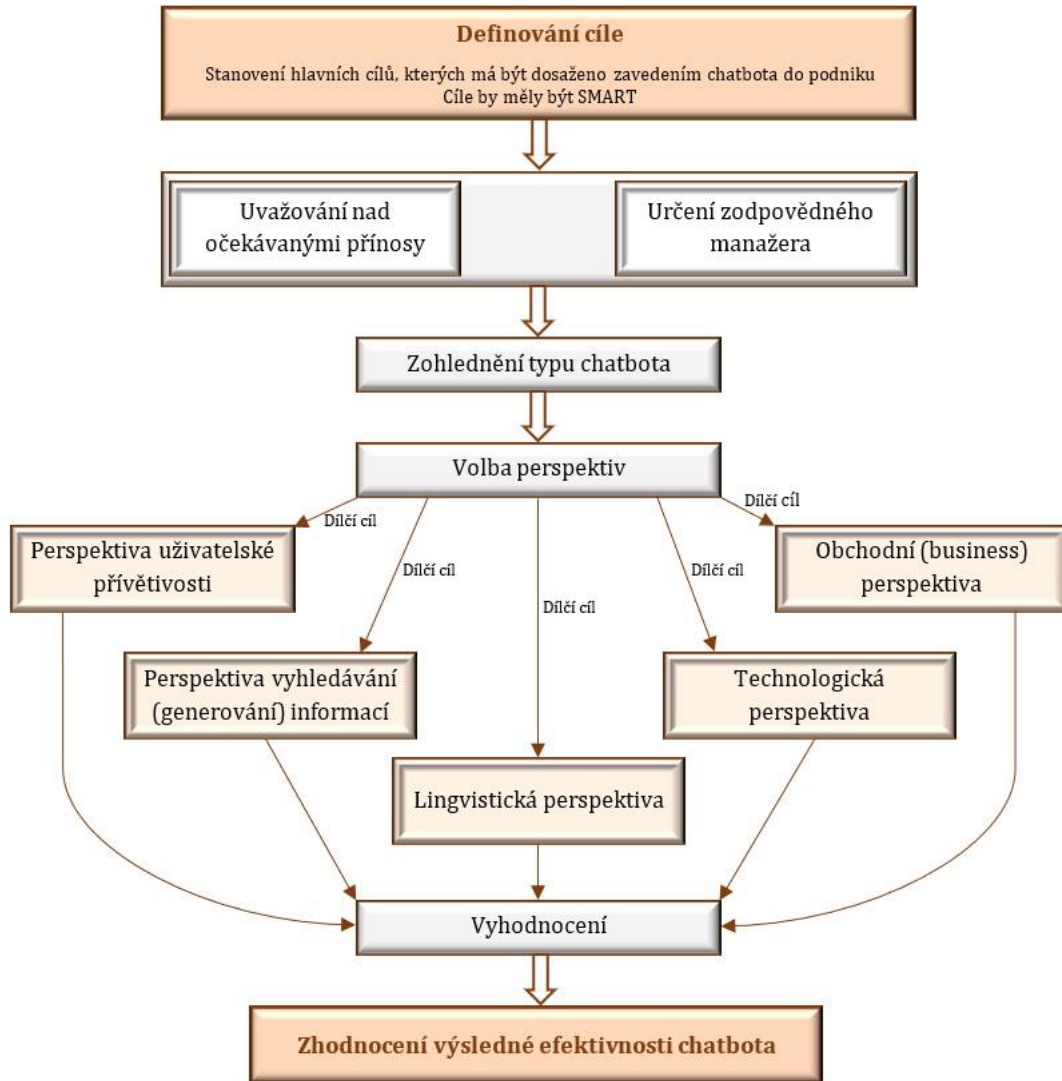


Schéma 4: Návrh na rozšíření vybraného přístupu k měření efektivity chatbota
 Zdroj: vlastní zpracování

9 Závěry a doporučení

Cílem diplomové práce bylo představit a porovnat dostupné přístupy k měření efektivnosti chatbota, vybrat jeden vhodný, rozebrat podle něho daný proces v reálné společnosti a na základě získaných teoretických a praktických poznatků navrhnout jeho rozšíření.

Pro zpracování práce bylo využito dostupných zahraničních studií čerpaných z odborných vědeckých databází, v nichž bylo těžké najít adekvátní zdroje zainteresované v řešené problematice. Té byla totiž dosud věnována velmi malá pozornost, a i z toho důvodu neexistuje veřejně známý přístup, který by se běžně používal. V tomto ohledu bylo tedy registrováno určité omezení.

Chatbot přece jenom není pouhý informační systém, ale je to technologie, která má za úkol v jisté (nejčastěji rutinní) činnosti nahradit člověka, s čímž je spojeno jeho působení, schopnosti, způsob interakce a další efekty. Proto je potřeba nad měřením jeho efektivnosti dostatečně uvažovat a zaměřit se na konkrétní oblasti vyhodnocující plnění jeho účelu (přinášení užitku), čímž je primárně myšlena komunikace s koncovými uživateli. Výběr vhodného přístupu, kdy bylo usilováno o nalezení toho nejkompexnějšího, byl podmíněn klíčovými kritérii: uživatelská přívětivost, informační schopnost a vybavenost, jazyková úroveň a vybavenost, lidskost a obchodní hledisko. Pro jejich stanovení bylo využito nejen literatury, ale i odbornosti a možnosti zajistit expertní stanovisko od specialistů z řad zaměstnanců společnosti ČSOB. Tato banka má v současnosti nasazeno přes 80 plně funkčních robotů, včetně několika chatbotů, a tak ví co od takové technologie reálně očekávat a jak ji zajišťovat. Na základě provedené vícekriteriální analýzy poté došlo k výběru přístupu od chorvatské autorky D. Peras publikovaného v roce 2018, která k měření efektivnosti chatbota přistupuje z pěti pohledů, kterými jsou: perspektiva uživatelské přívětivosti, perspektiva vyhledávání (generování) informací, lingvistická perspektiva, technologická perspektiva a obchodní (business) perspektiva. Ty dále rozděluje do kategorií, atributů a metrik, kdy metriky do nich zařazené odpovídají na otázky, jak je chatbot přívětivý, resp. použitelný, výkonný, působící a uspokojující, jakou přidanou hodnotu (užitek) přináší pro uživatele, jaká je kvalita, kvantita, relevance, jednoznačnost a gramatická přesnost výstupů, které produkuje,

zda se chová jako člověk a jakou přidanou hodnotu vytváří pro společnost. Přesněji se jedná o nefinančně měřitelné a kvalitativní ukazatele. V reakci na tuto volbu byl podle daného přístupu rozebrán proces měření efektivnosti chatbota (zaměřeného na úkol) v reálné společnosti, tedy ČSOB. Autorka vybraného přístupu uvádí, že ne všechny perspektivy je potřeba do hodnocení zahrnovat, protože vždy záleží na konkrétním typu chatbota a jeho účelu. Ke stejnému závěru došlo i během realizace dané analýzy v ČSOB, kdy byla lingvistická perspektiva shledána za nepotřebnou, a to právě z důvodu způsobu implementace zkoumaného chatbota. Za pomoci uskutečněného výzkumu bylo zjištěno, že se metriky nacházející ve vybraném přístupu používají i v reálné společnosti, a tudíž z hlediska obsahu lze označit daný přístup za použitelný. Navíc je vytvořený rámec přehledně rozdělený, což usnadňuje jeho uplatnění při měření a získávání odpovědí na výše zmíněné dílčí otázky. V souvislosti s touto analýzou bylo autorkou diplomové práce navrženo doplnění daného rámce o metriky a další skutečnosti, kterými je: způsob externího hodnocení za pomoci pětihvězdičkového systému a udělení náležité emotikony, využití tzv. fokusních skupin či tzv. Sounding Board, celkový a průměrný počet konverzací, ve kterých byl předán osobní údaj uživatele, míra jistoty chatbota, zpochybnění účinnosti Turingova testu a doporučení provádění analýzy časových řad.

Hlavním nedostatkem daného přístupu je absence jednotlivých kroků vedoucích k samotnému měření, resp. jak lze poznat co je u chatbota potřeba měřit a poté, jak s nasbíranými daty nakládat. Z tohoto důvodu bylo v souladu s teoretickou částí a získanými praktickými poznatky autorkou této práce navrženo další rozšíření v podobě celistvého postupu. Ten byl rozdělen do dvou etap, kdy, aby podnik hodnotil pro něho pouze relevantní perspektivy, musí už v etapě plánování zavedení chatbota vědět, jaký by měl být výsledný efekt a co je tedy cílem jeho nasazení, co je od něho očekáváno, přičemž je potřeba zohlednit jeho implementaci (tím je myšleno, na jaký typ konverzace je zaměřen (úkol či téma) a jak tvoří výstupy). Toto lze shrnout jako fázi, kdy se firma zabývá účelností. Ve chvíli, kdy je jasné, co se bude měřit, přichází teprve na řadu daný rámec, který udává, jakým způsobem se bude měřit a poukazuje tedy na účinnost (vložených zdrojů) chatbota vzhledem k získanému užítku (výslednému efektu). Po nasazení této technologie do produkčního prostředí dochází ke sběru dat, na což bezprostředně navazuje etapa kontroly, ve které jsou

vyhodnocovány výsledné hodnoty stanovených ukazatelů v předem dohodnutých lhůtách a porovnávány s chtěným stavem. V případě nedosahování definovaných cílů nastává rozhodování o změně, čímž by se měl zabývat strategický management. Hlavním doporučením je tedy pozorovat nasbíraná data v časových řadách a podrobit je analýze pro nalezení a pochopení zákonitostí a změn spojených s vývojem chatbota od začátku jeho zavedení až po odhadnutí jeho vývoje v budoucnosti.

S ohledem na výše zmíněné byl cíl práce naplněn v plném rozsahu a zároveň byla nalezena odpověď na stanovenou výzkumnou otázku.

V návaznosti na výsledky práce by bylo vhodné rozšířený přístup aplikovat v reálné společnosti a okomentovat jeho celistvou proveditelnost na nově zaváděném chatbotovi, což nebylo v případě diplomové práce možné uskutečnit. Zároveň by bylo příhodné se v rámci šetření zaměřit i na chatbota typu generative-based. Pro další zkoumání této problematiky by mohly posloužit i ostatní přístupy, které byly v práci představeny a porovnány. Přínosem by mohlo být i nalezení nových perspektiv a následně kategorií, atributů a metrik, jimž by měla být věnována pozornost z hlediska měření efektivnosti chatbota.

Diplomová práce přináší další pohled na danou problematiku a zároveň může posloužit, jak analyzované společnosti, tak těm, které neví, jaké oblasti a metriky je u chatbota vhodné posuzovat či se mohou inspirovat a doplnit tak své portfolio metrik. Zajišťování co největší efektivnosti chatbota vede ke zvyšování konkurenceschopnosti firem, které ho mají, což je na současném trhu společností přijímajících čtvrtou průmyslovou revoluci krucíální.

10 Seznam použité literatury

10.1 Tištěné zdroje

ČECH, P., BUREŠ, V. *Podniková informatika*. 1. vyd., Hradec Králové: Gaudeamus, 2009, 232 s. ISBN 978-80-7041-479-8.

DITTRICHOVÁ, J., SVOBODOVÁ, L., SOUKAL, I., JINDRA, V. *Základy financí*. 2. vyd., Hradec Králové: Gaudeamus, 2014, 214 s. ISBN 978-80-7435-409-0.

MINSKY, M., L. *Computation: Finite and Infinite Machines*. Prentice-Hall Series in Automatic Computation. Englewood Cliffs, N.J., Prentice-Hall, 1967, 317 s. ISBN 0-13-165563-9.

MOHELSKÁ, H., PITRA, Z. *Manažerské metody*. 1. vyd., Praha: Professional Publishing, 2012, 343 s. ISBN 978-80-7431-092-8.

MOLNÁR, Z. *Efektivnost informačních systémů*. Systémová integrace (Grada). 1. vyd., Praha: Grada, 2000, 142 s. ISBN 80-7169-410-X.

REVENDA, Z., MANDEL, M., KODERA, J., MUSÍLEK, P., DVOŘÁK, P. *Peněžní ekonomie a bankovníctví*. 5. aktualiz. vyd., Praha: Management Press, 2014, 423 s. ISBN 978-80-7261-279-6.

RUSSEL, S., NORVIG, P. *Artificial Intelligence: A modern Approach*. 3rd ed., Upper Saddle River, N.J.: Prentice Hall, 2010, 1132 s. ISBN 978-0-13-604259-4.

SHEVAT, A. *Designing Bots: Creating Conversational Experiences*. Boston: O'Reilly Media, 2017, 348 s. ISBN 978-149-1974-827.

SOKOLOVÁ, M. a kol. *Základy managementu*. 1. vyd., Hradec Králové: Gaudeamus, 2015, 265 s. ISBN 978-80-7435-553-0.

UČEŇ, P. a kol. *Metriky v informatice: Jak objektivně zjistit přínosy informačního systému*. 1. vyd., Praha: Grada, 2001, 139 s. ISBN 80-247-0080-8.

10.2 Elektronické zdroje

ABBAS, A. *Chatbot 2019 Trends and Stats with Insider Reports*. Chatbots Life [online]. 2019 [cit. 2019-12-23]. Dostupné z: <https://chatbotslife.com/chatbot-2019-trends-and-stats-with-insider-reports-fb71697deee4>.

ACCENTURE. *Chatbots in Customer Service*. Accenture Interactive: Part of Accenture Digital [online]. 2016 [cit. 2020-01-12]. Dostupné z: https://www.accenture.com/t00010101T000000_w_/br-pt/_acnmedia/PDF-45/Accenture-Chatbots-Customer-Service.pdf.

BLOOM, E. *Efficiency vs. effectiveness within IT*. CIO [online]. IDG Communications, Inc., 2016 [cit. 2019-11-05]. Dostupné z: https://www.cio.com/article/3076102/efficiency-vs-effectiveness-within-it.html?fbclid=IwAR20n2px8NVZUvL9Jw3xk4DOszHqj_5Hpi45geCBLlIRtRZTCc4jDfSqU.

BRANDTZAEG, P., B., FØLSTAD, A. *Why people use chatbots*. KOMPATSIARIS, I., CAVE, J., ed. *Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017, Proceedings* [online]. Cham: Springer International Publishing, 2017, Lecture Notes in Computer Science, vol. 10673 [cit. 2020-01-12]. ISBN 978-3-319-70283-4. DOI: 10.1007/978-3-319-70284-1. Dostupné z: <https://link.springer.com/book/10.1007/978-3-319-70284-1>.

BUSINESS INSIDER INTELLIGENCE. *80% of businesses want chatbots by 2020*. Businessinsider.com [online]. 2016 [cit. 2019-10-22]. Dostupné z: <https://www.businessinsider.com/80-of-businesses-want-chatbots-by-2020-2016-12>.

CUI, L., HUANG, S. et al. *SuperAgent: A Customer Service Chatbot for E-commerce Websites*. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics-System Demonstrations* [online]. Vancouver: Association for Computational Linguistics, 2017, s. 97-102 [cit. 2019-09-07]. DOI: 10.18653/v1/P17-4017. Dostupné z: <https://www.aclweb.org/anthology/P17-4017.pdf>.

ČERMÁK, P. *Umělá inteligence: Studijní opora pro kombinované studium*. Olomouc: Moravská vysoká škola Olomouc, o. p. s. [online]. 2018 [cit. 2019-07-13]. Dostupné z: <https://mvso.cz/wp-content/uploads/2018/02/Um%c4%9bl%c3%a1-inteligence-studijn%c3%ad-text.pdf>.

ČSOB. *ČSOB získala s týmovým robotem „Ró“ 3. místo v prestižní soutěži*. ČSOB [online]. 2019d [cit. 2020-02-02]. Dostupné z: <https://www.csob.cz/portal/-/tz191016b>.

ČSOB. *Chatbot Radim*. ČSOB [online]. 2020a [cit. 2020-02-07]. Dostupné z: <https://www.csob.cz/portal/lide/pujcky/pujcka-na-cokoliv>.

ČSOB. *Klientské centrum*. ČSOB [online]. 2019c [cit. 2019-10-11]. Dostupné z: <https://www.csob.cz/portal/csob/kariera/klientske-centrum>.

ČSOB. *O ČSOB a skupině: Informace o řízení a obchodní činnosti společnosti*. ČSOB [online]. 2019a [cit. 2019-08-24]. Dostupné z: <https://www.csob.cz/portal/csob/o-csob-a-skupine>.

ČSOB. *Půjčka na cokoliv: Peníze pro každou situaci – rychle a jednoduše*. ČSOB [online]. 2020b [cit. 2020-02-07]. Dostupné z: https://www.csob.cz/portal/lide/pujcky/pujcka-na-cokoliv?il1=radim_cl#spocitejte-si.

ČSOB. *Výroční zpráva 2018: Československá obchodní banka, a. s.* ČSOB [online]. 2019b [cit. 2019-08-25]. Dostupné z: <https://www.csob.cz/portal/documents/10710/444804/vz-csob-2018.pdf>.

DELOITTE. *ČSOB s pomocí Deloitte nasazuje do provozu další roboty* [online]. 2019 [cit. 2020-02-02]. Dostupné z: <https://www2.deloitte.com/cz/cs/pages/press/articles/cze-tz-csob-s-pomoci-deloitte-nasazuje-do-provozu-dalsi-roboty.html>.

DRIFT, SURVEY MONKEY AUDIENCE, SALESFORCE, MYCLEVER. *The 2018 State of Chatbots Report*. Drift [online]. 2018 [cit. 2019-12-23]. Dostupné z: <https://www.drift.com/wp-content/uploads/2018/01/2018-state-of-chatbots-report.pdf>.

DRIFT, SURVEY MONKEY AUDIENCE. *2019 State of Conversational Marketing [Free Report]*. Drift [online]. 2019 [cit. 2019-12-23]. Dostupné z: <https://www.drift.com/blog/state-of-conversational-marketing/#state-of-convo-marketing>.

EXPERT CHOICE. *The Analytic Hierarchy Process: Structured Decisions: Decision Making for Better Decisions*. Expert Choice [online]. 2020 [cit. 2020-01-19]. Dostupné z: <https://www.expertchoice.com/ahp-software>.

FACEBOOK IQ. *3 Ways Messaging Is Transforming the Path to Purchase* [online]. 2018 [cit. 2019-09-07]. Dostupné z: https://www.facebook.com/business/news/insights/3-ways-messaging-is-transforming-the-path-to-purchase?ref=fbiq_series#What-it-means-for-marketers.

FEEDYOU. *Portfolio* [online]. 2019 [cit. 2019-10-26]. Dostupné z: <https://feedyou.ai/success-stories/>.

GRICE, H., P. *Logic and Conversation*. Cole, P. & J. L. Morgan, ed. *Syntax and Semantics Vol. 3* [online]. New York: Academic Press, 1975, 41–58 [cit. 2019-10-10]. ISBN 0127854231. Dostupné z: <https://lawandlogic.files.wordpress.com/2018/07/grice1975logic-and-conversation.pdf>.

HOLZMAN, O. *Vojtěch Dlouhý z Feedyou: Chatbot není umělá inteligence. Dokáže ale nahradit zaměstnance a šetřit náklady*. CzechCrunch [online]. 2019 [cit. 2020-02-09]. Dostupné z: <https://www.czechcrunch.cz/2019/03/vojtech-dlouhy-z-feedyou-chatbot-neni-umela-inteligence-dokaze-ale-nahradit-zamestnance-a-setrit-naklady/>.

ISO. *ISO 9241-11:2018(en) Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts*. Online Browsing Platform (OBP) [online]. 2018 [2020-01-10]. Dostupné z: <https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en>.

ISO. *Popular Standards: ISO 9000 Family – Quality Management* [online]. International Organization for Standardization, 2019 [cit. 2019-12-14]. Dostupné z: <https://www.iso.org/iso-9001-quality-management.html>.

KBC. *Who we are*. KBC Group [online]. 2019 [cit. 2019-08-24]. Dostupné z: <https://www.kbc.com/en/who-we-are>.

KHOROZOV, A. *Trends Driving the Chatbot Growth*. Chatbots Magazine [online]. 2017 [cit. 2020-03-22]. Dostupné z: <https://chatbotsmagazine.com/trends-driving-the-chatbot-growth-77b78145bac>.

KOJOUHAROV, S. *Ultimate Guide to Leveraging NLP & Machine Learning for your Chatbot*. Chatbots Life [online]. 2016 [cit. 2019-09-10]. Dostupné z: <https://chatbotslife.com/ultimate-guide-to-leveraging-nlp-machine-learning-for-you-chatbot-531ff2dd870c>.

KULIGOWSKA, K. *Commercial Chatbot: Performance Evaluation, Usability Metrics and Quality Standards of Embodied Conversational Agents*. Professionals Center for Business Research [online]. 2015, vol. 2, no. 2 [cit. 2020-01-20]. ISSN 2409-9783. DOI: 10.18483/PCBR.22. Dostupné z: https://www.researchgate.net/publication/282474589_Commercial_Chatbot_Performance_Evaluation_Usability_Metrics_and_Quality_Standards_of_Embodied_Conversational_Agents.

LOKMAN, A., S., ZAIN, J., M. *Chatbot Enhanced Algorithms: A Case Study on Implementation in Bahasa Malaysia Human Language*. ZAVORAL, F., YAGHOB, J., PICHAPPAN, P., EL-QAWASMEH, E., ed. *Networked Digital Technologies* [online]. Berlin: Springer International Publishing, 2010, Communications in Computer and Information Science, vol. 87 [cit. 2019-07-29]. ISBN 978-3-642-14291-8. DOI: 10.1007/978-3-642-14292-5_5. Dostupné z: https://link.springer.com/chapter/10.1007%2F978-3-642-14292-5_5.

MAROENGSI, W. et al. *A Survey on Evaluation Methods for Chatbots*. Proceedings of 2019 7th International Conference on Information and Education Technology (ICIET 2019) [online]. 2019, s. 111–119 [cit. 2019-10-10]. DOI: 10.1145/3323771.3323824. Dostupné z: <https://dl.acm.org/doi/10.1145/3323771.3323824>.

MAŘÍK, V. a kol. *Národní iniciativa: Průmysl 4.0*. Ministerstvo průmyslu a obchodu [online]. 2015 [cit. 2019-10-05]. Dostupné z: https://automatizace.hw.cz/files/prumysl_40_23stran_publicace_vladycr_0.pdf.

MICROSOFT. *Co je služba QnA Maker?* [online]. 2019 [cit. 2020-01-15]. Dostupné z: <https://docs.microsoft.com/cs-cz/azure/cognitive-services/qnamaker/overview/overview>.

MICROSOFT. *Cognitive Services Pricing - QnA Maker* [online]. 2020 [cit. 2020-01-15]. Dostupné z: <https://azure.microsoft.com/en-us/pricing/details/cognitive-services/qna-maker/>.

NICASTRO, D. *8 Ways to Measure Chatbot Program Success*. CMS WiRE, Channel: Customer Experience [online]. 2018 [cit. 2019-10-23]. Dostupné z: <https://www.cmswire.com/customer-experience/8-ways-to-measure-chatbot-program-success/>.

OVERBY, S., GREINER, L., PAUL, L., G. *What is an SLA? Best practices for service-level agreements*. CIO [online]. IDG Communications, Inc., 2017 [cit. 2019-12-15]. Dostupné z: <https://www.cio.com/article/2438284/outsourcing-sla-definitions-and-solutions.html>.

PERAS, D. *Chatbot Evaluation Metrics: Review Paper*. 36th International Scientific Conference on Economic and Social Development – "Building Resilient Society" [online]. Varazdin: Varazdin Development and Entrepreneurship Agency (VADEA), 2018, s. 89–97 [cit. 2019-10-10]. ISSN 1849-7535. Dostupné z: https://www.esd-conference.com/upload/book_of_proceedings/Book_of_Proceedings_esdZagreb2018_Online.pdf.

QUALTRICS. *What is Net Promoter Score (NPS)? Definition & Examples* [online]. 2019 [cit. 2019-11-05]. Dostupné z: <https://www.qualtrics.com/experience-management/customer/net-promoter-score/>.

QUI, M., LI, F.-L., WANG, S., GAO, X., CHEN, Y., ZHAO, W., CHEN, H., HUANG, J., CHU, W. *AliMe Chat: A Sequence to Sequence and Rerank based Chatbot Engine*. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics [online]. Vancouver: Association for Computational Linguistics, vol. 2: Short Papers, 2017, s. 498–503 [cit. 2019-10-01]. DOI: 10.18653/v1/P17-2079. Dostupné z: <https://www.aclweb.org/anthology/P17-2079/>.

RADZIWILL, N., MORGAN, B. *Evaluating Quality of Chatbots and Intelligent Conversational Agents* [online]. 2017 [cit. 2020-01-10]. CoRR abs/1704.04579. Dostupné z: <https://arxiv.org/abs/1704.04579>.

RAMOS, R. *Screw the Turing test — chatbots don't need to act human*. VentureBeat [online]. 2017 [cit. 2020-01-10]. Dostupné z: <https://venturebeat.com/2017/02/03/screw-the-turing-test-chatbots-dont-need-to-act-human/>.

RIEKE, T., D. *The relationship between motives for using a Chatbot and satisfaction with Chatbot characteristics in the Portuguese Millennial population: an exploratory study*. Porto, 2018. Dissertation. University of Porto. Faculty of Economics. Supervised by Helena Martins. Dostupné z: <https://repositorio-aberto.up.pt/bitstream/10216/116509/2/296743.pdf>.

SANJEEVI, M. *Chapter 11: ChatBots to Question & Answer systems*. Medium [online]. 2018 [cit. 2019-09-10]. Dostupné z: <https://medium.com/deep-math-machine-learning-ai/chapter-11-chatbots-to-question-answer-systems-e06c648ac22a>.

SHAH, N. *Which Is Best for You: Rule-Based Bots or AI Bots?* Chatbots Magazine [online]. 2017 [cit. 2019-12-28]. Dostupné z: <https://chatbotmagazine.com/which-is-best-for-you-rule-based-bots-or-ai-bots-298b9106c81d>.

SHAWAR, B., A., ATWELL, E. *Chatbots: Are they Really Useful?* LDV Forum 22 [online]. 2007, s. 29-49 [cit. 2020-01-12]. Dostupné z: https://pdfs.semanticscholar.org/8d82/84bfba7ebcb4e2575d864ec7c16ea6a168f0.pdf?_ga=2.213540061.837378187.1578848909-862667441.1578595632.

SIKOROVÁ, M. *Seznamte se, váš nový bankovní úředník je robot*. EkonTech.cz [online]. 2019 [cit. 2020-02-02]. Dostupné z: <https://www.ekontech.cz/clanek/seznamte-se-vas-novy-bankovni-urednik-je-robot>.

SLOUKA, D. *AI vs ML: co je umělá inteligence a co strojové učení?* IDG Communications, Inc. [online]. 2019 [2019-12-28]. Dostupné z: <https://aiworld.cz/umela-inteligence/ai-vs-ml-co-je-umela-inteligence-a-co-strojove-uceni-288>.

ŠVÍK, M. *ROI, TCO a NPV: Svatá trojice*. CIO [online]. IDG Communications, Inc., 2009 [cit. 2019-12-14]. Dostupné z: <https://businessworld.cz/it-strategie/roi-tco-a-npv-svata-trojice-5303>.

Veřejný rejstřík a Sběrka listin: Výpis z obchodního rejstříku. Justice.cz [online]. Praha: Ministerstvo spravedlnosti České republiky, 2019 [cit. 2019-08-25]. Dostupné z: <https://or.justice.cz/ias/ui/rejstrik-firma.vysledky?subjektId=56666&typ=PLATNY>.

WAKEFIELD, J. *The hobbyists competing to make AI human*. BBC [online]. 2019 [cit. 2020-01-13]. Dostupné z: <https://www.bbc.com/news/technology-49578503>.

WALTER, B., ALBERS, A., SCHELLEIS, A., KURRLE, A. *Efficient use of Sounding Board Method at project milestones and its potentials for virtualization*. Procedia CIRP [online]. 2017, vol. 60, s. 284-289 [cit. 2020-03-12]. DOI: 10.1016/j.procir.2017.02.004. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S2212827117300689>.

XU, A., LIU, Z., GUO, Y., SINHA, V., AKKIRAJU, R. *A new chatbot for customer service on social media*. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems [online]. Denver: ACM, 2017, s. 3506–3510 [2019-10-01]. DOI: 10.1145/3025453.3025496. Dostupné z: <https://dl.acm.org/doi/10.1145/3025453.3025496>.

ZÁHORA, J. *Učebnice statistiky* [online]. 12. verze, Univerzita Karlova v Praze, Lékařská fakulta v Hradci Králové, 2015 [cit. 2019-12-13]. ISBN 978-80-88176-00-8. Dostupné z: <https://publi.cz/books/201/Cover.html>.

11 Přílohy

- 1) Dotazník sloužící k expertnímu stanovení vah kritérií pro výběr vhodného přístupu k měření efektivnosti chatbota
- 2) Saatyho matice s jednotlivými a konečnými váhami kritérií

Uživatelská přívětivost × **Informační schopnost a vybavenost**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Uživatelská přívětivost × **Jazyková úroveň a vybavenost**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Uživatelská přívětivost × **Lidskost**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Uživatelská přívětivost × **Obchodní hledisko**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Informační schopnost a vybavenost × **Jazyková úroveň a vybavenost**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Informační schopnost a vybavenost × **Lidskost**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Informační schopnost a vybavenost × **Obchodní hledisko**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Jazyková úroveň a vybavenost × **Lidskost**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Jazyková úroveň a vybavenost × **Obchodní hledisko**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Lidskost × **Obchodní hledisko**

9	8	7	6	5	4	3	2	1	2	3	4	5	6	7	8	9
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Děkuji za Vámi věnovaný čas a spolupráci.

Bc. Andrea Kotoučková

Kritéria:

- K1 Uživatelská přívětivost
 K2 Informační schopnost a vybavenost
 K3 Jazyková úroveň a vybavenost
 K4 Lidskost
 K5 Obchodní hledisko

Experti:

- E1 Expert č. 1
 E2 Expert č. 2
 E3 Expert č. 3
 E4 Expert č. 4
 E5 Expert č. 5
 E6 Expert č. 6

Další značení:

G_i – geometrický průměr i -tého kritéria

v_i – normovaná váha i -tého kritéria

Jednotlivé Saatyho matice:

	K1	K2	K3	K4	K5	G_i	v_i
K1	1	1/3	1	1	1/7	0,544	0,077
K2	3	1	4	6	1/5	1,705	0,242
K3	1	1/4	1	1	1/4	0,574	0,082
K4	1	1/6	1	1	1/5	0,506	0,072
K5	7	5	4	5	1	3,707	0,527
Celkem						7,036	1,000

Tabulka 4: Saatyho matice – E1

Zdroj: vlastní zpracování

	K1	K2	K3	K4	K5	G_i	v_i
K1	1	1/4	7	8	4	2,237	0,305
K2	4	1	6	8	2	3,288	0,448
K3	1/7	1/6	1	3	1/2	0,514	0,070
K4	1/8	1/8	1/3	1	1/5	0,253	0,034
K5	1/4	1/2	2	5	1	1,046	0,143
Celkem						7,338	1,000

Tabulka 5: Saatyho matice – E2

Zdroj: vlastní zpracování

	K1	K2	K3	K4	K5	G _i	v _i
K1	1	1/3	3	7	3	1,838	0,271
K2	3	1	3	7	3	2,853	0,420
K3	1/3	1/3	1	7	1/3	0,763	0,112
K4	1/7	1/7	1/7	1	1/5	0,226	0,033
K5	1/3	1/3	3	5	1	1,108	0,163
Celkem						6,788	1,000

Tabulka 6: Saatyho matice - E3

Zdroj: vlastní zpracování

	K1	K2	K3	K4	K5	G _i	v _i
K1	1	1/4	5	6	1/3	1,201	0,164
K2	4	1	5	8	1/2	2,402	0,329
K3	1/5	1/5	1	1	1/6	0,367	0,050
K4	1/6	1/8	1	1	1/7	0,312	0,043
K5	3	2	6	7	1	3,022	0,414
Celkem						7,304	1,000

Tabulka 7: Saatyho matice - E4

Zdroj: vlastní zpracování

	K1	K2	K3	K4	K5	G _i	v _i
K1	1	1/4	1/3	3	1/4	0,574	0,090
K2	4	1	1	6	1	1,888	0,295
K3	3	1	1	5	1/3	1,380	0,215
K4	1/3	1/6	1/5	1	1/5	0,295	0,046
K5	4	1	3	5	1	2,268	0,354
Celkem						6,405	1,000

Tabulka 8: Saatyho matice - E5

Zdroj: vlastní zpracování

	K1	K2	K3	K4	K5	G _i	v _i
K1	1	1	6	5	4	2,605	0,378
K2	1	1	8	8	1	2,297	0,333
K3	1/6	1/8	1	3	1/3	0,461	0,067
K4	1/5	1/8	1/3	1	1/4	0,291	0,042
K5	1/4	1	3	4	1	1,246	0,181
Celkem						6,900	1,000

Tabulka 9: Saatyho matice - E6

Zdroj: vlastní zpracování

	R1	R2	R3	R4	R5	R6	Výsledné (jednotné) váhy	Výsledné pořadí
K1	0,077	0,305	0,271	0,164	0,090	0,378	0,214	3.
K2	0,242	0,448	0,420	0,329	0,295	0,333	0,345	1.
K3	0,082	0,070	0,112	0,050	0,215	0,067	0,099	4.
K4	0,072	0,034	0,033	0,043	0,046	0,042	0,045	5.
K5	0,527	0,143	0,163	0,414	0,354	0,181	0,297	2.
							1,000	

Tabulka 10: Stanovení výsledných (jednotných) vah kritérií a jejich příslušného pořadí

Zdroj: vlastní zpracování

Seznam zkratek

AHP	Analytic Hierarchy Process
AI	Artificial Intelligence
B2C	Business to Consumer
CRM	Customer Relationship Management
ČSOB	Československá obchodní banka, a. s.
EFMA	European Financial Management Association
FAQ	Frequently Asked Questions
FO	Fyzická osoba
FTE	Full Time Equivalent
GDPR	General Data Protection Regulation
ICT	Information and Communication Technologies
IS	Information Systems
ISO	Organization for Standardization
IT	Information Technology
KBC	Kredietbank ABB Insurance CERA Bank
KC	Klientské centrum
KPI	Key Performance Indicators
MIT	Massachusetts Institute of Technology
NBC	National Broadcasting Company
NLP	Natural Language Processing
NPS	Net Promoter Score
NPV	Net Present Value
NV	Naamloze Vennootschap
Q&A	Questions and Answers
QR	Quick Response
ROI	Return on Investments
SLA	Service Level Agreement
SMART	Specific, Measurable, Acceptable/Achievable, Relevant/Realistic, Time Specific/Trackable
SME	Small and Medium Enterprise
TCO	Total Cost of Ownership
URL	Uniform Resource Locator

Zadání práce

UNIVERZITA HRADEC KRÁLOVÉ
Fakulta informatiky a managementu
Akademický rok: 2019/2020

Studijní program: Systémové inženýrství a informatika
Forma studia: Prezenční
Obor/kombinace: Informační management (im2-p)

Podklad pro zadání DIPLOMOVÉ práce studenta

Jméno a příjmení: **Bc. Andrea Kotoučková**
Osobní číslo: **I1800138**
Adresa: Pod Zámečkem 1762/11, Hradec Králové – Nový Hradec Králové, 50006 Hradec Králové 6, Česká republika
Téma práce: Měření efektivity chatbota
Téma práce anglicky: Measurement the effectiveness of chatbot
Vedoucí práce: prof. Ing. Hana Mohelská, Ph.D.
Katedra managementu

Zásady pro vypracování:

Cíl práce: Představení a zhodnocení možností měření efektivity chatbota a jejich aplikace ve vybraném podniku.

Osnova:

1. Úvod
2. Cíl práce a metodika
3. Literární rešerše
4. Měření efektivity konkrétního chatbota ve vybraném podniku
5. Shrnutí výsledků
6. Závěry a doporučení
7. Seznam použité literatury
8. Přílohy

Seznam doporučené literatury:

- MAROENGSI, W. et al. *A Survey on Evaluation Methods for Chatbots*. In: Proceedings of 2019 7th International Conference on Information and Education Technology (ICIET 2019) [online]. 2019, s. 111 ? 119 [cit. 2019-10-10]. Dostupné z: <https://doi.org/10.1145/3323771.3323824>.
- PERAS, D. *Chatbot Evaluation Metrics: Review Paper*. In: 36th International Scientific Conference on Economic and Social Development ? „Building Resilient Society“ [online]. Varazdin: Varazdin Development and Entrepreneurship Agency (VADEA), 2018, s. 89 ? 97 [cit. 2019-10-10]. Dostupné z: <https://search.proquest.com/docview/2176212638/fulltextPDF/B6CAFE7628848A3PQ/1?accountid=14623>.
- RADZIWIŁ, N., M., BENTON, M., C. *Evaluating Quality of Chatbots and Intelligent Conversational Agents*. Software Quality Professional, 2017 [cit. 2019-10-10]. Dostupné z: <https://arxiv.org/ftp/arxiv/papers/1704/1704.04579.pdf>.
- RUSSELL, S., NORVIG, P. *Artificial Intelligence: A modern Approach*. 3rd Edition, Upper Saddle River: Prentice Hall, 2010, 1152 s. ISBN 0-13-604259-7 978-0-13-604259-4.
- SHAWAR, B., A., ATWELL, E. *Different measurements metrics to evaluate a chatbot system*. In: NAACL-HLT-Dialog '07 Proceedings of the Workshop on Bridging the Gap: Academic and Industrial Research in Dialog Technologies [online]. Stroudsburg: Association for Computational Linguistics, 2007, s. 89 ? 96 [cit. 2019-10-10]. Dostupné z: <https://doi.org/10.3115/1556328.1556341>.

SINGH, A., RAMASUBRAMANIAN, K., SHIVAM, S. *Processes in the Banking and Insurance Industries*. In: *Building an Enterprise Chatbot: Work with Protected Enterprise Data Using Open Source Frameworks* [online]. 1. vyd., Berkeley: Apress, 2019, s. 1 ? 18 [cit. 2019-10-10]. ISBN 978-1-4842-5034-1. Dostupné z: https://doi.org/10.1007/978-1-4842-5034-1_1.

Podpis studenta:

Datum:

Podpis vedoucího práce:

Datum: