

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií



Bakalářská práce

Hlasoví asistenti

Tomáš Kundrt

© 2022 ČZU v Praze

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Provozně ekonomická fakulta

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Tomáš Kundrt

Informatika

Název práce

Hlasoví asistenti

Název anglicky

Voice assistants

Cíle práce

Bakalářská práce je tematicky zaměřena na problematiku hlasových asistentů. Hlavním cílem práce je analýza a komparace řešení hlasových asistentů s důrazem na využití v chytré domácnosti s pilotní implementací. Dílčí cíle práce jsou:

- vypracování přehledu trhu s hlasovými asistenty,
- charakteristika technologií pro hlasové asistenty,
- definování nedostatků současných hlasových asistentů.

Metodika

Metodika řešené problematiky bakalářské práce je založena na studiu a analýze odborných informačních zdrojů. Vlastní práce spočívá v objektivní analýze a komparaci současných řešení a technologií pro hlasové asistenty s pilotní implementací. Na základě syntézy teoretických poznatků a výsledků praktické části budou formulovány závěry bakalářské práce.

Doporučený rozsah práce

40 – 50 stran textu.

Klíčová slova

Hlasový asistent, hlasové ovládání, chytrá domácnost, umělá inteligence, Internet věcí, rozpoznání hlasu, soukromí

Doporučené zdroje informací

ČÍKA, Petr. Internet věcí pro inteligentní domácnost: Internet of things for smart home : zkrácená verze habilitační práce. Brno: Vysoké učení technické v Brně, nakladatelství VUTIUM, 2017. ISBN 978-80-214-5559-7

DUNIN-UNDERWOOD, Anna. Alexa, can you keep a secret? Applicability of the third-party doctrine to information collected in the home by virtual assistants. 29. Information & Communications Technology Law. 2020, s. 101-119. ISSN 1360-0834. Dostupné z: doi:10.1080/13600834.2020.1676956

GREENGARD, Samuel. The internet of things. Cambridge, Massachusetts: MIT Press, [2015]. ISBN 978-0-262-52773-6.

HOY, Matthew B. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. 37. Medical Reference Services Quarterly. 2018, s. 81-88. ISSN 0276-3869. Dostupné z: doi:10.1080/02763869.2018.1404391

SHARIF, Khairunisa a Bastian TENBERGEN. Smart Home Voice Assistants: A Literature Survey of User Privacy and Security Vulnerabilities. Complex Systems Informatics and Modeling Quarterly. 2020, s. 15-30. ISSN 22559922. Dostupné z: doi:10.7250/csimq.2020-24.02

Předběžný termín obhajoby

2021/22 LS – PEF

Vedoucí práce

Ing. Pavel Šímek, Ph.D.

Garantující pracoviště

Katedra informačních technologií

Elektronicky schváleno dne 10. 8. 2021

doc. Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 5. 10. 2021

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 10. 03. 2022

Čestné prohlášení

Prohlašuji, že svou bakalářskou práci "Hlasoví asistenti" jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené bakalářské práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 15. 3. 2022

Poděkování

Rád bych touto cestou poděkoval panu doc. Ing. Pavlu Šimkovi, Ph.D. za příkladné vedení mé práce, za jeho ochotu, za cenné rady a připomínky při konzultacích a vypracování bakalářské práce.

Hlasoví asistenti

Abstrakt

Bakalářská práce se zabývá analýzou a komparací řešení hlasových asistentů s následným návrhem a pilotní implementací autora vlastního hlasového asistenta. Seznamuje čtenáře s všeobecným přehledem internetových věcí specifikovaných pro téma práce, s bezpečností použití, srovnáním růstu trhu s hlasovými asistenty, a jejich budoucí využití. Práce popisuje vlastnosti a srovnání vybraných hlasových asistentů pro použití v chytré domácnosti. V práci je kladen důraz na problematiku hlasových asistentů, jak pro domácí použití, tak pro použití v počítačových a mobilních zařízeních.

Klíčová slova: Hlasový asistent, hlasové ovládání, chytrá domácnost, umělá inteligence, internet věcí, rozpoznání hlasu, soukromí, domácí asistenti, analýza řešení, pilotní implementace, bakalářská práce

Voice assistants

Abstract

The bachelor thesis deals with the analysis and comparison of voice assistant solutions with the subsequent design and pilot implementation of the author's own voice assistant. It introduces the reader to a general overview of the IoT specified for the thesis topic, the safety of use, a comparison of the growth of the voice assistant market, and their future use. The thesis describes the features and comparison of selected voice assistants for use in the smart home. The thesis focuses on the issue of voice assistants, both for home use and for use in computing and mobile devices.

Keywords: Voice assistant, voice control, smart home, artificial intelligence, internet of things, voice recognition, privacy, home assistants, solution analysis, pilot implementation, bachelor thesis

Obsah

1 Úvod	13
2 Cíl práce a metodika	14
2.1 Cíl práce.....	14
2.2 Metodika.....	14
3 Teoretická východiska	15
3.1 Internet věcí	15
3.1.1 Historie IoT.....	15
3.1.2 Aplikace.....	16
3.1.3 Spotřebitelská využití a chytrá domácnost	17
3.2 Hlasoví asistenti	18
3.2.1 Úvod do hlasových asistentů	18
3.2.2 Historie hlasového ovládání	19
3.2.3 Co jsou vlastně hlasoví asistenti	20
3.2.4 Co dokážou hlasoví asistenti	21
3.2.5 Jazyk, kompatibilita a umělá inteligence	23
3.3 Ochrana soukromí a možnosti zabezpečení.....	25
3.3.1 Ochrana osobních údajů	25
3.3.2 Zabezpečení	26
3.3.3 Možnosti zabezpečení	27
3.4 Domácí asistenti.....	28
3.4.1 Amazon Echo.....	28
3.4.2 Apple HomePod.....	29
3.4.3 Google Home	30
3.5 Ostatní hlasoví asistenti.....	32
3.5.1 Microsoft Cortana	32
3.5.2 Samsung Bixby	33
3.5.3 Seznam hlasový asistent	34
3.6 Stav trhu hlasových asistentů.....	35
3.6.1 Uživatelé a firmy.....	35
3.6.2 Přehled trhu.....	36
3.6.3 Klíčové trendy trhu	37
3.7 Potenciální budoucí použití	39
3.7.1 Překladače.....	39
3.7.2 Knihovny	39
3.7.3 Průvodci.....	40
4 Vlastní řešení.....	40

4.1	Analýza a komparace řešení	40
4.1.1	Úvod do řešení hlasových asistentů	40
4.1.2	Všeobecná studie hlasových asistentů.....	42
4.1.3	Komparace mezi hlasovými asistenty	43
4.1.4	Omezení a budoucí možnosti využití	44
4.1.5	Závěr analýzy a komparace řešení	46
4.2	Návrh aplikace	46
4.2.1	Běžové prostředí	47
4.2.2	1. Fáze vývoje	47
4.2.3	2. Fáze vývoje	48
4.2.4	3. Fáze vývoje	49
4.3	Implementace aplikace	50
4.3.1	Import knihoven a nastavení enginu aplikace	51
4.3.2	Vytvoření hlasového asistenta	52
4.3.3	Nastavení hlavních funkcí	54
4.3.4	Vyřešení zjištěných nedostatků a možných hrozeb	56
5	Výsledky a diskuse	59
5.1	Výsledky	59
5.2	Diskuse	59
6	Závěr	60
7	Seznam použitých zdrojů	61

Seznam obrázků

Obrázek 1 [34] Amazon echo 4. generace. Dostupné online z:

<<https://www.alza.cz/amazon-echo-dot-4-generace-glacier-white-s-hodinami-d6242639.html>>

Obrázek 2 [37] Apple HomePod. Dostupné online z:

<<https://hlasovi-asistenti.heureka.cz/apple-homepod/>>

Obrázek 3 [39] Google Home. Dostupné online z:

<<https://www.alza.cz/google-home-d4254706.html>>

Obrázek 4 [50] Předpověď růstu trhu. Dostupné online z:

<<https://www.mordorintelligence.com/industry-reports/voice-assistant-application-market>>

Obrázek 5 [54] Systém dialogů.

R. Sivapriyan, N. Sakshi and T. Vishnu Priya, "Comparative Analysis of Smart Voice Assistants," 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2021, pp. 1-6,

Dostupné z doi: 10.1109/CSITSS54238.2021.9683722.

Obrázek 6 Import knihoven [autor]

Obrázek 7 Nastavení enginu aplikace [autor]

Obrázek 8 Vytvoření hlasového asistenta část 1 [autor]

Obrázek 9 Vytvoření hlasového asistenta část 2 [autor]

Obrázek 10 Ukázka nastavení hlavních funkcí část 1 [autor]

Obrázek 11 Ukázka nastavení hlavních funkcí část 2 [autor]

Obrázek 12 Ukázka nastavení hlavních funkcí část 3 [autor]

Obrázek 13 Vyřešení nedostatku – Poznámky [autor]

Obrázek 14 Vyřešení nedostatku – Textové zprávy a hovory [autor]

Obrázek 15 Vyřešení nedostatku – Zábava [autor]

Obrázek 16 Možnosti pro zabezpečení [autor]

Seznam tabulek

Tabulka 1 Všeobecná analýza hlasových asistentů

Tabulka 2 Komparace řešení

Tabulka 3 Omezení a pohled na budoucnost

Seznam použitých zkratk

IoT – Internet of Things

D2D – Device to Device communication

M2M – Machine to Machine communication

Auto-ID – Automatic identification

MIT – Massachusetts Institute of Technology

RFID – Radio Frequency Identification

NFC – Near Field Communication

QR – Quick Response

ITU – International Telecommunication Union

Wi-Fi – Wireless Fidelity

IBM – International Business Machines Corporation

USA – United States of America

ML – Machine Learning

IFTTT – If This Then That – internetová služba

Voice-ID – Voice identification

CAPTCHA-Completely Automated Public Turing test to tell Computers and Humans Apart

HSBC – The Hongkong and Shanghai Banking Corporation Limited

AI – Artificial Intelligence

CAGR – Compound Annual Growth Rate

COVID-19 - Coronavirus disease 2019

WHO – World Health Organization

CDC – Centers for Disease Control and Prevention

NIH – National Institutes of Health

SNCF – Société nationale des chemins de fer français – národní železniční dopravce ve Francii

INRIA – National Institute for Research in Digital Science and Technology

CPOE – Computerized physician order entry

EHR – Electronic health record

DNN – Deep Neural Networks

HTML – HyperText Markup Language

XML – Extensible Markup Language

GUI – Graphical User Interface

CLI – Command Line Interface

API – Application Programming Interface

1 Úvod

Přirozenou snahou člověka je si život a každou činnost v něm zjednodušovat. Tento fakt lze pozorovat v každém odvětví lidského zájmu od počátku věků. Čím více člověk s problémem přichází do styku, tím častěji přemýšlí o tom, jak by si danou činnost mohl zjednodušit či zkrátit čas pro ni potřebný – například ovládání přístrojů hlasem.

Hlasové technologie a hlasoví asistenti umožňují nejpřirozenější formu komunikace – mluvení. V minulých letech se jednalo především o rozvoj chatbotů, kteří byli schopni psanou formou komunikovat a interagovat s člověkem. Nyní jsou v hlasové formě a stávají se z nich praví osobní asistenti. Pod pojmem hlasový asistent se skrývá aplikace, která rozumí ve větší či menší míře přirozenému jazyku člověka a dokáže na nej stejně přirozeně odpovídat či reagovat, a to tak, jak by bylo očekáváno od člověka, který by nám dělal asistenta. Osobu, která by řešila většinu fádních každodenních věcí za nás a na nás by zbyla pouze tvůrčí činnost. Apple Siri, Amazon Alexa, Microsoft Cortana a Google Assistant jsou nejoblíbenějšími hlasovými asistenty a jsou vestavěné do smartphonů nebo vyhrazených domácích reproduktorů. Uživatelé mohou svým asistentům klást otázky, ovládat zařízení domácí automatizace a přehrávání médií pomocí hlasu a spravovat další základní úkoly, jako je e-mail, seznamy úkolů a kalendáře, pomocí slovních příkazů.

Hlasoví asistenti v chytrých reproduktorech, tzv domácí asistenti jsou v současné době hlavními hnacími silami hlasových technologií, v USA již více než třetina všech domácností vlastní hlasového asistenta. Stávají se tak technologií, které je potřeba začít věnovat pozornost. Asistenti stále nejsou adaptováni na český jazyk, což velmi ovlivňuje možnost jejich využití v České republice. Je proto otázkou, zda má pro firmy smysl se jimi zabývat už teď, nebo mají čekat až na jazykové přizpůsobení, u kterého není jisté kdy přijde.

Hlavním cílem této práce je analýza a komparace řešení mezi hlasovými asistenty převážně využívané v chytré domácnosti s pilotní implementací vlastního hlasového asistenta vytvořeného na základě zjištěných nedostatků. Mezi dílčí cíle pak patří zpracování přehledu trhu s hlasovými asistenty a jeho budoucí vývoj, charakteristika využívaných technologií pro hlasové asistenty a zjištění současných nedostatků těchto asistentů.

2 Cíl práce a metodika

2.1 Cíl práce

Bakalářská práce je tematicky zaměřena na problematiku hlasových asistentů. Hlavním cílem práce je analýza a komparace řešení hlasových asistentů s důrazem na využití v chytré domácnosti s pilotní implementací. Dílčí cíle práce jsou:

- vypracování přehledu trhu s hlasovými asistenty,
- charakteristika technologií pro hlasové asistenty,
- definování nedostatků současných hlasových asistentů.

2.2 Metodika

Metodika řešené problematiky bakalářské práce je založena na studiu a analýze odborných informačních zdrojů a dostupných zdrojů online. Teoretická část bakalářské práce spočívá v řešení problematiky hlasových asistentů, jejich potenciální narušení soukromí a následné možnosti zabezpečení, srovnání přehledu trhu a obeznámení s hlavními zástupci domácích hlasových asistentů.

Praktická část bakalářské práce bude zpracována dle získaných poznatků z teoretické části práce. Tato část bude zpracována jako analýza a komparace řešení s následnou pilotní implementací. Pro analýzu bude použita obecná exaktní metoda, přesněji srovnávací metoda, pro pilotní implementaci bude použit programovací jazyk Python v prostředí PyCharm. V pilotní implementaci bude vyvíjen hlasový asistent s použitými metodami ostatních asistentů a následně budou vyřešeny zjištěné nedostatky a potenciální možné hrozby při používání podobných zařízení. Pro vývoj aplikace bude použita kaskádová metodika vývoje, která je běžná pro implementaci.

Na základě syntézy teoretických poznatků a výsledků praktické části budou formulovány závěry bakalářské práce.

3 Teoretická východiska

3.1 Internet věcí

Pojem Internet věcí hraje stále větší roli jak v průmyslu, tak v ostatních sektorech. Daný koncept má obrovský vliv na to, jak se díváme na naši práci, ale i na náš každodenní život. IoT zařízení jsou stále více rozšiřována a používána jak v domácím, tak i v pracovním prostředí v podobě různých senzorů a kamer, ale i při používání svého chytrého telefonu. V roce 2016 bylo přidáno téměř půl miliardy (429 milionů) mobilních zařízení, čímž se počet zařízení na světě vyšplhal na osm miliard. Spolu s počtem mobilních zařízení stoupá i počet přenesených dat mobilní sítí. Podle odhadu společnosti Cisco se počet dat přenesených sítí za 1 měsíc zvýší během čtyř let o 38 exabytů. Od roku 2016 by to byl nárůst o 700 %. [1] Experti odhadli, že Internet věcí v roce 2020 zahrnoval přibližně 30 miliard zařízení. Hodnota trhu se odhaduje na 80 miliard dolarů. [2]

IoT je koncept hovořící o tom, že jakékoli zařízení, které může být připojeno k Internetu, je součástí IoT. IoT zařízení jsou pomocí Internetu nebo jiné služby ovládané a využívány pro získávání a zpracovávání informací, spínání zařízení nebo i komunikaci s člověkem. Podle způsobu komunikace účastníků konverzace se dělí IoT na komunikaci dvou zařízení – D2D a komunikaci dvou strojů – M2M. Koncept IoT zahrnuje inteligentní kávovary, mobilní telefony, sluchátka, žárovky, ale i proudové motory, drony, vrtačky či digitální vodoměry, elektroměry apod. [3]

3.1.1 Historie IoT

Historie IoT se započala vznikem komunikace pomocí elektronických zařízení. První zařízení pro zmíněný druh komunikace bylo vytvořeno v roce 1832, kdy Baron Shilling v Rusku zkonstruoval elektromagnetický telegraf. Nicméně první zařízení, které bylo zařazeno do skupiny IoT, bylo vytvořeno v roce 1990. Byl to topinkovač, který mohl být zapnut a vypnut pomocí Internetu. [4] Ovšem zařízení, které bylo vůbec poprvé připojeno k internetové síti, byl modifikovaný stroj na výrobu Coca-Coly na univerzitě Carnegie Mellon v Pittsburghu.

Ačkoliv zařízení, které jsou řazeny do skupiny IoT, byly vytvářeny již dříve, pojem "Internet of Things" byl poprvé použit Kevinem Ashtonem, ředitelem Auto-ID Center, který

je jeden z předních světových průkopníků IoT. Uvedený technologický průkopník ho použil ve své prezentaci pro firmu Procter & Gamble v roce 1999. [5]

Koncepce Internetu věcí se stala populární v již zmíněném roce 1999, prostřednictvím Auto-ID Center v MIT a souvisejících zveřejnění analýz trhu. Radiofrekvenční identifikaci (RFID) viděl Kevin Ashton (jeden ze zakladatelů původního Auto-ID Centra) jako předpoklad pro Internet věcí. Ashton preferoval frázi „Internet pro věci“. Pokud by všechny objekty a lidé v každodenním životě byly vybaveny identifikátory, počítače by je mohly spravovat a inventarizovat.

Označování věcí může být řešeno např. pomocí technologií:

- RFID
- NFC
- čárový kód
- QR kód
- digitální vodoznak

V roce 2005 nastal velký rozmach IoT, vlivem ITU zveřejněním zprávy, v níž se uvádí, že kdykoli a z jakéhokoli místa na světě se bude možné připojit k jakémukoli zařízení v jiné části světa. Uvedené připojení vytvoří novou dynamickou síť sítí - Internet věcí.

V roce 2006 byly na trh uvedeny Nabaztag, mechanické zajíci připojené k internetu pomocí technologie Wi-Fi, které dokázaly se svým uživatelům komunikovat o nových pohybech na trhu akcií, titulcích zpráv, budících, a jiné. [6]

3.1.2 Aplikace

Aplikace pro zařízení připojená na internet jsou rozsáhlé. Došlo k mnohačetným pokusům o kategorizaci, z nichž většina souhlasí s rozdělením mezi spotřební, podnikatelské a infrastrukturní aplikace. George Osborne, ex-kancléř britské pokladny, prohlásil,

že Internet věcí je další stadium informační revoluce a zmínil propojitelnost všeho od městské dopravy přes nemocniční zařízení až po domácí využití.

Možnost propojit zařízení s limitovaným procesním výkonem, pamětí a energetickou spotřebou znamená, že IoT nachází využití v téměř každém oboru. Takové systémy

by mohly být ve vedení sběru informací v situacích sahajících od přírodních ekosystémů až po budovy a továrny, čímž nacházejí uplatnění v oblastech snímání prostředí a urbanistického plánování. [7]

Inteligentní obchodní systémy by například mohly monitorovat nakupovací zvyky konkrétních uživatelů sledováním jejich konkrétních mobilních telefonů. Těmto uživatelům by poté mohly být poskytnuty speciální nabídky na jejich oblíbené produkty nebo by jim bylo dokonce možné hlásit umístění toho, co potřebují, o co si automaticky zažádala jejich lednička. Další příklady snímání a ovládání jsou v aplikacích zahrnujících teplo, vodu, elektřinu a řízení spotřeby energie nebo asistence s naváděním přepravních systémů. Jiné aplikace, které IoT může poskytnout, je umožnit rozšíření zabezpečení a automatizace domácnosti. Koncept „internet žijících věcí“ byl navrhnout jako popis sítí s biologickými senzory, které by mohly využít analýzy založené na cloudech, aby umožnily uživatelům studium DNA nebo jiných molekul.

3.1.3 Spotřebitelská využití a chytrá domácnost

Rostoucí podíl získávají IoT zařízení pro spotřebitelské využití. Příklady spotřebitelských aplikací zahrnují propojené auto, zábavu, domácí automatizaci (známé také jako chytrá zařízení do domácnosti), nositelnou technologii, zařízení denní potřeby, propojený zdravotní stav, spotřebiče jako pračky a sušičky, robotické vysavače, čističky vzduchu, trouby nebo chladničky a mrazničky, které používají Wi-Fi ke vzdálenému sledování.

Přístroje IoT jsou součástí širší koncepce domácí automatizace, známé také jako domotika. Velké inteligentní domácí systémy využívají hlavní router nebo switch, který uživatelům poskytuje centrální ovládání jejich veškerých zařízení.

Jednou z klíčových aplikací inteligentní domácnosti je pomoc zdravotně znevýhodněným a starým lidem. Domácí systémy používají podpůrnou technologii přizpůsobenou specifickému postižení člověka. Hlasové ovládání může pomoci uživatelům s omezením zraku a pohyblivosti, výstražné systémy mohou být připojeny přímo k implantátům, které užívají sluchově postižení lidé apod. Využit lze senzory, které monitorují zdravotní krizové situace, jako jsou pády nebo záchvaty. Technologie chytré domácnosti používané tímto způsobem mohou poskytnout uživatelům větší svobodu a vyšší kvalitu života.

3.2 Hlasoví asistenti

3.2.1 Úvod do hlasových asistentů

Už v minulém století si lidé představovali vedle létajících aut i domácí spotřebiče ovládané hlasem, jako nevyhnutelnou budoucnost. Ačkoliv jsme se stále běžného rozšíření létajících aut nedočkali, s chytrou domácností jsme pokročili mnohem více. Hlasem ovládaní chytrí asistenti jsou již poměrně běžné především na západ od nás. Masovému rozšíření asistentů pro chytrou domácnost u nás stále brání do jisté míry jazyková bariéra, která však postupně také opadá. Chytrá domácnost již v dnešní době nemusí být vůbec nákladnou záležitostí, jak tomu bylo dříve. [8]

Ještě před několika desítkami let se myšlenka smysluplné konverzace s počítačem zdála futuristická, ale technologie, která umožňuje hlasové rozhraní používat a být široce dostupná, zde již je. Několik produktů na úrovni spotřebitele vyvinutých v posledních několika letech přineslo do každodenního používání levné hlasové asistenty a neustále přibývají další funkce a platformy. Uživatelé mohou dělat vše od kladení jednoduchých informačních otázek po přehrávání hudby a vytáčení telefonů nebo zapnutí a vypnutí světel pomocí hlasového ovládání. Tento sloupec prozkoumá základní fungování a společné vlastnosti dnešních hlasových asistentů. Rovněž se bude diskutovat o některých otázkách ochrany soukromí a zabezpečení, které jsou vlastní hlasovým asistentům, a o možném budoucím použití těchto zařízení. Jelikož jsou hlasoví asistenti stále více využíváni, knihovníci se budou chtít seznámit s dosud nedefinovanou technologií, která má potenciál být mechanismem doručení knihovnických materiálů a služeb.

Hlasové ovládání je úzce spjato s rozvojem technologií a rozvojem rozpoznávání řeči. Technologie musí nejprve rozpoznávat a rozumět hlasům a teprve potom je můžeme ovládat svými hlasy.

Rozpoznávání řeči analyzuje zvuky, které slyší, a filtruje, co daná osoba říká. Poté je předá v digitální podobě, aby bylo možné je pochopit a interpretovat. Jak se naučit a zlepšit rozpoznávání řeči spojené se strojovým učením. Používá se velké množství nahrávek lidského hlasu, na kterých se algoritmus učí rozumět hlasu. Tyto nahrávky jsou vyžadovány nejen pro jednotlivé jazyky, ale také dialekty a přízvuky, protože každý z nich významně ovlivňuje porozumění asistentovi. [9]

Další funkce, jako je provedení příkazu „napsat zprávu“ nebo „přehrát hudbu ze Spotify“, lze přiřadit pouze k této funkci.

Hlavní výhody hlasové technologie jsou:

- Rychlost – rychlejší mluvení než psaní,
- Přesnost – gramaticky správné psaní (i když někdy nerozpoznává slova správně),
- Produktivita – hlasová technologie zvyšuje produktivitu,
- Bezpečnost – To je případ například u řízení aut. [10]

Mezi další výhody může patřit fakt, že tuto technologii mohou využívat i lidé se zdravotním postižením a ovládat tak počítače a další zařízení, což by jinak nebylo možné. Využití hlasové technologie běžným uživatelem dalo vzniknout mobilním aplikacím využívajícím hlasové prvky. Například iTranslate umožňuje hlasový překlad věty z jednoho jazyka do druhého. Věty k překladu jsou zaznamenány, přepsány do textu, přeloženy a zároveň mohou být okamžitě převedeny do druhého jazyka. Vše se děje během několika sekund, což umožňuje rychlé dohody i bez znalosti druhého jazyka. Kromě toho, že jazyk, který aplikace má velmi širokou schopnost překladu (pod 100) a 12, má také obrovský způsob použití na světě. Podle tvůrců stáhl více než 100 milionů lidí téměř každé zemi. [11]

3.2.2 Historie hlasového ovládání

První rozpoznávání hlasu se datuje do půlky 20. století, kdy byl v roce 1952 vytvořen systém Audrey, který byl schopen rozpoznat čísla. Do té doby zařízení byla schopná pouze hlas zaznamenávat.

Následně v roce 1962 byl vytvořen systém IBM vytvořený systém Shoe box, který zvládl rozeznat 16 slov a 9 číslic v angličtině. V roce 1971 byl USA založen Institut pro rozpoznávání řeči. představuje první velkou změnu v hlasovém ovládání, která vyvrcholila systémem Harpey, systém je schopen rozpoznat přes 1000 slov. [12]

Důležitý pokrok byl Hidden Markov model, který vyměřoval pravděpodobnost neznámých zvuků jako slov. To umožnilo rozpoznat slovo v řádech tisíců. Jako příklad použití tohoto modelu přišla IBM s Project Tangor, který dokáže rozpoznat až 20 000 slov a některé celé věty. Tento projekt byl veden Čechem Frederickem Jelínkem, který se zasadil o základy toho, jak se vyvíjí programy pro převod řeči do digitální podoby. [13]

Největší pokrok však přišel až v roce 1997 u Dragon's Naturally speaking softwaru, který umožňoval, že nebylo potřeba se zastavit za každý proneseným slovem a rozuměl až 100 slovům za minutu. To, s mírnými úpravami, byl vrchol hlasového rozpoznávání do doby, než se zapojila aplikace Google Voice Search v roce 2008. Strojové učení proniklo

i do této oblasti a Google díky tomu bylo schopné pomoci aplikaci, která z velkého objemu trénovacích dat obstojně rozuměla lidskému hlasu.

Další větší zlomek přišel v roce 2011, kdy společnost Apple spustila Siri, prvního hlasového asistenta, který zvládne provést jednoduché příkazy. Spustil tak vlnu tvorby asistentů dalších firem. V roce 2013 představil Microsoft hlasového asistenta Cortana. V roce 2014 Amazon navázal asistentem Alexa. V roce 2016 jej následoval Google s asistentem Google a v roce 2017 Samsung 13 s asistentem Bixby. Hlavní zprávy v roce 2017 však byly zprávy, že společnost Google získala dostatečnou úroveň přesnosti porozumění anglickým hlasovým dotazům na 95 %. O tom se mluví jako o procentu, kterému běžně rozumí i člověk. [14]

3.2.3 Co jsou vlastně hlasoví asistenti

Jednoduše řečeno, hlasoví asistenti jsou realizací interakce rozhovoru s počítači. Apple Siri, Microsoft Cortana, Amazon Alexa a Google Assistant jsou všichni softwaroví agenti, kteří běží na účelových reproduktorových zařízeních nebo smartphonech. Software neustále poslouchá klíčové slovo, aby ho probudil. Jakmile toto klíčové slovo uslyší, zaznamená hlas uživatele a odešle jej na specializovaný server, který jej zpracuje a interpretuje jako příkaz. V závislosti na příkazu server poskytne hlasovému asistentovi příslušné informace, aby jej mohl uživatel přečíst zpět, přehrát médium požadované uživatelem nebo dokončit úkoly s různými připojenými službami a zařízeními. Počet služeb, které podporují hlasové příkazy, rychle roste a výrobci zařízení tzv. Internetu věcí také integrují hlasové ovládání do svých produktů. [15]

Asistent společnosti Siri od společnosti Apple je nejstarší, byl vydán jako samostatná aplikace v roce 2010 a do systému iOS byl přidán v roce 2011. Microsoft následoval krátce poté s Cortanou v roce 2013. Amazon uvedl Alexu s domácím reproduktorem připojeným k Echo v roce 2014 a v roce 2016 byl oznámen asistent Google spolu s domácím reproduktorem a je také zabudován do aplikace Google pro smartphony založené na Androidu. Každý asistent má své vlastní jedinečné vlastnosti, ale základní funkce jsou stejné. Jedna z následujících kapitol se věnuje podrobnějšímu popisu těchto asistentů.

Hlasoví asistenti se liší od dřívějších technologií aktivovaných hlasem v tom, že dokážou reagovat na mnohem větší počet příkazů a otázek. Je to proto, že jsou vždy připojeni k internetu; každá interakce je odeslána zpět do centrálního výpočetního systému, který

analyzuje hlasové příkazy uživatele a poskytuje asistentovi správnou odpověď. Dřívější hlasem aktivovaná zařízení se spoléhala na menší sadu „vestavěných“ příkazů a odpovědí. Nedávný pokrok ve zpracování přirozeného jazyka, známý také jako výpočetní lingvistika, umožnil hlasovým asistentům rychle vytvářet smysluplné odpovědi. Hirschberg a Manning připisují tato nedávná vylepšení zpracování přirozeného jazyka čtyřem věcem:

- obrovské zvýšení výpočetního výkonu,
- dostupnost velmi velkého množství jazykových dat,
- vývoj vysoce úspěšných ML metod
- mnohem bohatší porozumění struktuře lidského jazyka a jeho nasazení v sociálních kontextech. [16]

Vzhledem k tomu, že osobní počítače jsou stále levnější a výkonnější a lidé vytvářejí stále více online textů k analýze, vědci použil tento text k trénování hlasových asistentů, aby naslouchali a reagovali na naše požadavky přirozenějšími a smysluplnějšími způsoby. Hlasoví asistenti mohou analyzovat požadavky formulované mnoha různými způsoby a interpretovat, co uživatel s největší pravděpodobností bude chtít. Například, když chcete požádat asistenta Google, aby si pamatoval, kde někdo zaparkoval své auto, může uživatel říci kteroukoli z mnoha frází: „Pamatujte si, kde jsem zaparkoval,“ „Zaparkoval jsem zde,“ „Nechal jsem auto na 6. ulici, nebo „auto je na jihu“, všichni dosáhnou podobného výsledku. Google si bude pamatovat, kde uživatel zaparkoval auto, a když se ho později zeptá, bude schopen odpovídajícím způsobem reagovat. Uživatel může klást otázky podobně přirozeným způsobem; dotaz „kde jsem zaparkoval“, „kde jsem nechal auto“ nebo „pamatujete si, kde jsem zaparkoval“ - to vše vyvolá očekávanou odpověď. Zpracování přirozeného jazyka se vyhne frustraci uživatelů dřívějších systémů rozpoznávání hlasu, které pro správné fungování vyžadovaly specifické fráze a vzorce.

3.2.4 Co dokážou hlasoví asistenti

Ačkoli každý aktuálně dostupný hlasový asistent disponuje jedinečné funkce, sdílejí některé podobnosti a jsou schopni provádět následující základní úkoly:

- odesílat a číst textové zprávy, telefonovat, odesílat a číst e-mailové zprávy;
- odpovídat na základní informační dotazy („Kolik je hodin? Jaká je předpověď počasí?“);
- nastavit časovače, alarmy a položky kalendáře;

- nastavovat připomenutí, vytvářet seznamy a provádět základní matematické výpočty;
- ovládat přehrávání médií z připojených služeb, jako jsou Amazon, Google Play, iTunes, Pandora, Netflix a Spotify;
- ovládat zařízení podporující internet věci, jako jsou termostaty, světla, alarmy a zámky;
- vyprávět vtipy a příběhy.

Kromě těchto úkolů mohou hlasoví asistenti přidat další funkce, často nazývané „dovednosti“, které rozšiřují jejich schopnosti propojením s jinými programy pomocí hlasových příkazů. Alexa od Amazonu má dovednosti v přehrávání médií, objednávání obvyklých nápojů a jídel z místní restaurace a vyvolávání Uber nebo Lyft pomocí údajů z připojeného účtu. Asistent Google má podobné dovednosti, ale zaostává za Amazonem v naprostém počtu dostupných dovedností, a to především díky pozdějšímu vydání. [17] Tyto dovednosti jsou vytvářeny vývojáři třetích stran, podobně jako vývoj aplikací pro chytré telefony. Google Assistant se také integruje s několika nástroji, které uživatelům umožňují vytvářet vlastní dovednosti. Pomocí webových služeb, jako jsou Tasker a IFTTT, mohou uživatelé vytvářet dovednosti, které jim umožní automatizovat příspěvky na sociálních médiích, zapínat a vypínat zařízení a stovky dalších možností. Například říct asistentovi „Dobré ráno“ by mohlo spustit řadu akcí určených k urychlení ranní rutiny uživatele: zapnutí kávovaru, čtení zpráv a událostí v kalendáři, otevření garážových vrat a nastartování vozu a poté zamknutí dveří, aktivace alarmu a nastavení termostatu po odchodu lidí do práce. Všechna zařízení nezbytná k provádění těchto úkolů na dálku jsou dnes k dispozici; hlasový asistent jednoduše poskytuje most, který umožňuje uživatelům vydávat příkazy verbálně, nikoli prostřednictvím aplikace.

Hlasoví asistenti jsou k dispozici na široké škále hardwarových platform. Amazon a Google uvádějí na trh vyhrazená domácí reproduktorová zařízení pro své hlasové asistenty. Amazon vyrábí několik variant svého produktu Echo, od malé Echo Dot až po Echo Show, která má audio a video funkce. Hardware reproduktorů Google se nazývá Home a dodává se také v modelech mini a full-size. Apple právě vstupuje na trh domácích reproduktorů s oznámením svého HomePod zařízení s podporou Siri, které má být k dispozici v prosinci 2017. Společnost Microsoft se zaměřila na zabudování Cortany do počítačů a telefonů

s Windows 10 a nedávno spolupracovala se společností Harmon Kardon na vývoji domácího reproduktoru s podporou Cortany.

Asistenti jsou k dispozici také na většině platform chytřích telefonů; Asistent Google je integrován do telefonů Android a lze jej nainstalovat jako samostatnou aplikaci na iPhone, i když některé funkce jsou deaktivovány. [18] Alexa od Amazonu má verze aplikací pro Android a iOS a Microsoft a Amazon spolupracují na tom, aby přivedly Cortanu na zařízení Amazon a Alexa do počítačů. [19] Apple nechává svého asistenta mimo zařízení, která nejsou iOS, ale Siri je k dispozici na všech zařízeních Apple, včetně iPhone, MacBooků, iPadů a Apple Watch. Jak se trh s hlasovým asistentem stabilizuje, je pravděpodobné, že bude existovat další integrace a tyto sady funkcí u hlavních hlasových asistentů budou podobné. V tuto chvíli je Amazon dominantním hráčem v této oblasti, protože nejprve uvedl na trh domácí produkt s velkou knihovnou médií, která je k dispozici ihned po vybalení. Google buduje kapacitu a přidání domácího reproduktoru a integrace do dalších produktů Google posílí jejich podíl na trhu. Apple se také může stát více uchazečem s vydáním HomePodu koncem roku 2017 a přidáním dalších připojených domácích produktů značky Apple. Microsoft pravděpodobně nezíská větší poutavost, protože jejich podíl na trhu smartphonů je zanedbatelný a chybí jim poutavý domácí produkt. [15]

3.2.5 Jazyk, kompatibilita a umělá inteligence

Jakým jazykem budete se svým chytrým asistentem mluvit je pro nás, jako Čechy, velice důležité. Je to také jazyková bariéra, které prozatím brzdí větší rozšíření chytrých hlasových asistentů u nás. Všechny zařízení pro chytrou domácnost z naší recenze zvládají na výborné úrovni angličtinu, s češtinou to pak je už mnohem horší. Volbu češtiny, jako výchozího jazyka nabízí zatím pouze Google Asistent. Čeština je však v případě Asistenta Google docela omezená, při jejím využívání budete limitováni některými funkcemi, které Google Home v češtině nezvládne, rovněž celá umělá inteligence nemá češtinu tak “naučenou”, jako angličtinu a je to na celém hlasovém projevu velice znát. Základní pokyny a funkce však “český” Google Asistent s přehledem zvládne. Siri od firmy Apple bohužel češtinu nepodporuje. Alexa od firmy Amazon češtinu aktuálně také nepodporuje. Asistent Google češtinu zvládá, ale celý systém je při využití češtiny střidmější než anglická verze.

Umělá inteligence chytrých asistentů je založena na strojovém učení. Program si tak musí zpracovat nepřehledné množství psaného či mluveného projevu v daném jazyce, aby fungoval co nejlépe. Plnohodnotným rozšířením do dalších jazyků tak aktuálně brání hlavně nedostatek vstupů či dat v daném jazyce.

Jak je váš asistent pro chytrou domácnost opravdu chytrý je také velice důležitým kritériem při výběru toho správného. Čím více zařízení různých výrobců můžete do ekosystému své chytré domácnosti připojit je klíčově důležitým kritériem. Jak na tom tedy Asistent Google, Siri a Alexa od Amazonu jsou?

Většina z prodávaného příslušenství (chytré zásuvky, žárovky, kamery, zámky či meteostanice) jsou kompatibilní se všemi ze třech hlavních systémů pro chytrou domácnost. V popisku produktu by mělo být vždy napsáno, zda výrobek podporuje zrovna vašeho hlasového asistenta. Rozdíly nejsou ani u služeb s předplatným (Spotify, Netflix, HBO GO...), podpora je vždy stejná u všech tří zařízení.

Drobné rozdíly však naleznete v kompatibilitě s dalšími produkty značky. Například Google podporuje u svého chytrého asistenta protokol Cast, pokud tak máte TV s Androidem nebo využíváte zařízení Chromecast, můžete si nechat hlasovým povelům převést libovolný audiovizuální soubor na obrazovku, kterou si vyberete. V tomto ohledu nejvíce zaostává Amazon Alexa, protože společnosti Google i Apple mají spoustu dalších produktů, které lze do chytré domácnosti integrovat.

Díky umělé inteligenci dokáže hlasový asistent pružně reagovat na různé otázky a být vždy pohotový. Čím lepší pak je jeho AI tím lépe dokáže odpovídat a tím přesnější odpovědi také poskytuje. Je tedy zřejmé, že úroveň umělé inteligence je pro hlasového asistenta klíčová.

Otázka tedy zní, jaká umělá inteligence hlasového asistenta pro chytrou domácnost je ta nejlepší? Odpověď je poměrně jednoduchá, ta od Googlu. Asistent Google se již několik let pravidelně umisťuje na prvním místě v nezávislém srovnání schopností hlasových asistentů (test je prováděn v angličtině) s poměrně velkým náskokem, následně jsou v hodnocení umístěny asistenty od Applu a Amazonu, pouze s malým rozdílem. Ačkoliv je pravda, že Apple a Amazon se poslední roky na Google dotahují, bude to trvat asi ještě dlouho, než budou na podobné úrovni, kde je Google nyní.

Nebylo by však vhodné říci, že Siri a Alexa nejsou dobrými hlasovými asistenty, pouze je faktem tvrzení, že Google je o kousek dál. Během každodenního používání není možné si všimnout malých rozdílů mezi jednotlivými asistenty . [21]

3.3 Ochrana soukromí a možnosti zabezpečení

Používání hlasových asistentů doma také přináší obavy o soukromí. V domácnosti už existuje zařízení, které funguje tak, že je poslouchá. Teoreticky může naslouchat v každé situaci a získat velmi citlivá data o tom, co slyší. Lidé si tato rizika uvědomují, a to je jeden z faktorů, který brání hlasovým asistentům v plném využití jejich potenciálu.

3.3.1 Ochrana osobních údajů

Hlasové ovládání je stále nová technologie a firmy se na zabezpečení těchto dat teprve připravují. Únik informací a přístup k datům je obecně častým problémem těchto technologických gigantů jejichž technologie je používána každý den a kteří často vědí o lidech více než lidé sami o sobě.

Nedávné incidenty téměř všech asistentů také ukazují že se jedná o novou technologii a zabezpečení dat se stále zkoumá. Obrovská vlna obav o soukromí ohledně používání asistentky vyvolala zprávy o asistentovi Alexa. Bylo poukázáno na to, že tým Amazonu poslouchá nahrávky rozhovorů s asistentem, aby mohli v budoucnu lépe reagovat. Zmínil také e-mail od mluvčího Amazonu, že zaměstnanci mají velmi přísná pravidla a nemohou přímo identifikovat uživatele, když se zaregistrují, ale stále vidí jména sériová čísla a čísla účtů. Ukazuje se, ale také že zaměstnanci, kteří nahrávky přepisují mají přístup k poloze uživatele a v některých případech i k vlastní adrese zákazníka. [22] K vyřešení těchto problémů poskytl Amazon v září 2019 funkci pro odstranění interakcí s asistentem.

Otázka hledání nahrávek rozhovorů s asistenty se objevila i u dalších společností jako Apple nebo Google. Apple měl poté problémy s uchováváním záznamů a jejich poslechem, což provedlo drastickou změnu, kdy se automaticky neukládají žádné nahrávky, pokud někdo nechce nahrávky uložit, aby pomohl vyvinout asistenta. Ve společnosti Google byly zveřejněny dokonce zvukové nahrávky 1000 lidí z Belgie.

V roce 2017 se lidé nebáli a potenciální rizika přehlíželi, protože necítili, že oni jsou cílem, a důvěřovali společnosti, že opatření pro zabezpečení jsou dostatečná. V roce 2019, si však lidé už problémy se soukromím více uvědomovali. Z dat Microsoftu z roku 2019 se 52 % uživatelů obává toho, že jejich osobní data nejsou řádně zabezpečena a 41 % uživatelů se obecně obává toho, že je zařízení aktivně poslouchá a/ nebo nahrává. Tyto obavy mohou být jedním z důvodů, proč se lidé zdráhají použít ke zveřejnění lékařských záznamů a dalších citlivých informací. [23]

Problém ochrany soukromí při používání asistenta se netýká pouze domácích asistentů. Při používání asistenta na svém mobilním telefonu se ujistěte že vás mezi lidmi slyší. I tento aspekt je problémem a lidé jej nechtějí používat na veřejnosti.

3.3.2 Zabezpečení

Jak již bylo řečeno, domácí hlasoví asistenti představují velký vpád do života lidí a do jejich soukromí. Mohou mít přístup k velmi citlivým údajům, takže je důležité si uvědomit potenciální nebezpečí, která mohou představovat. Existuje řada možných útoků proti asistentům a stále se objevují nové, protože jsme v raných fázích této technologie a všechny její slabiny musí být teprve řádně objeveny.

I když hlasoví asistenti mají zajímavé a užitečné funkce, představují také několik jedinečných problémů. Jedním z hlavních problémů těchto hlasem aktivovaných zařízení je zabezpečení. Kdokoli, kdo má přístup k zařízení aktivovanému hlasem, mu může klást otázky, shromažďovat informace o účtech a službách přidružených k zařízení a požádat jej o provádění úkolů. To představuje velké bezpečnostní riziko, protože tato zařízení budou velmi číst obsah kalendáře, e-maily a další osobní informace. V jednom hlášeném případě muž zjistil, že iPad v jeho obývacím pokoji odemkne přední dveře každému, kdo stojí venku a požádá Siri, aby je pustil dovnitř. [24] Google nedávno upgradoval svůj software Assistant na hlasový tisk, který jednoznačně identifikuje každého uživatele hlasem a brání zařízení ve čtení osobních údajů. Apple také učí Siri rozpoznávat hlas uživatele, ale tato funkce ještě nebyla vydaná v době psaní tohoto článku. Alexa společnosti Amazon je stejně náchylná k těmto bezpečnostním problémům a Amazon pracuje na zavedení podobného systému hlasového tisku. Alexa má další problém zabudování do rozhraní obchodu Amazon. Ve výchozím nastavení si kdokoli s hlasovým přístupem k zařízení může objednávat položky pomocí účtu Amazon vlastníka. Existují možnosti, jak nastavit hlasový kód pro potvrzení nákupů, a veškeré zboží bude odesláno na evidovanou adresu vlastníka, ale stále existuje potenciál pro uživatele se zlými úmysly nakupovat zboží na účtu vlastníka. Členové domácnosti mohli také provádět neoprávněné nákupy, jako šestiletá holčička, která si prostřednictvím Alexy objednala domeček pro panenky a čtyři libry cukrovinek. [25]

Hlasoví asistenti jsou také zranitelní několika dalšími útoky. Vědci nedávno prokázali, že hlasoví asistenti budou reagovat na neslyšitelné příkazy vydávané na ultrazvukových frekvencích. [26] To by útočníkovi umožnilo přiblížit se k oběti, přehrát ultrazvukový příkaz a zařízení oběti by reagovalo. Existuje také možnost, že by tento typ

útohu mohl být vložen ve vysílacích médiích. Reklama na produkt v televizi, která obsahuje vložené ultrazvukové příkazy pro přidání položky do nákupního košíku, vás může pobídnout k jejímu zakoupení. Tato myšlenka se zdá být přitažlivá, ale řada novinových zpráv o dříve zmíněné dívce, která si objednala domeček pro panenky, způsobila problémy mnoha lidem. Hlasatel během hlášení příběhu promluvil slovy: „Alexo, objednej mi domeček pro panenky“ a stovky Alex v domovech diváků odpověděly.

3.3.3 Možnosti zabezpečení

Jednou z možností může být zvuková CAPTCHA, která vyžaduje, aby byl přepsán text, který asistent řekne před provedením příkazu. Tento záznam by asistent těžko rozeznal, zatímco pro člověka by to neměl být problém. To však výrazně omezuje využití hlasových asistentů obecně, kdy by jakékoli použití asistenta muselo být potvrzováno přes CAPTCHA. Může však být řešením pro některé obtížně přístupné příkazy. Dalším problémem je, že vývoj rozpoznávání řeči jde rychleji než vývoj audio CAPTCHA, takže to pravděpodobně nebude fungovat. [27]

Druhým způsobem ochrany je využití funkce rozpoznání hlasu „majitele“ hlasového asistenta. Asistent nyní dokáže rozpoznat hlas uživatele, který jej používá nejčastěji. Otázkou však je, zda to bude zásadní překážka, neboť hlasy lze naklonovat, a i tato bezpečnostní opatření lze obejít.

Další je funkce rozpoznávání hlasu firmy Nuance, která umožňuje ověřit hlas uživatele na základě předchozích hlasových nahrávek. Hlas je také jedinečná lidská vlastnost, jako je obličej nebo otisk prstu, a lze jej proto použít k ověření. Analyzovali více než 10 faktorů pro hlasové nahrávky uživatelů a byli schopni identifikovat duplicitní a pozměněné nahrávky. [28]

Poskytuje dva typy Voice ID, aktivní a pasivní. Pro aktivní identifikaci je potřeba říct frázi „můj hlas je moje heslo“. Ten pak systém porovná s nahraným záznamem, pod kterým je uživatel identifikován. Tuto funkci lze také propojit s webem nebo aplikací, aby byly zajištěny bezpečné transakce. Forma pasivní identifikace umožňuje ověřit osobu bez použití frází, pouze naslouchá na pozadí telefonického rozhovoru zákazníka s informační linkou. Tímto způsobem rozpozná hlas do 10 sekund a neruší konverzaci. Také je tento způsob těžší prolomit, protože již nestačí napadat jednoduché věty.

Nastavení autentizace není složité, pro proaktivní ověření stačí nahrát zmíněnou frázi vícekrát a zabezpečení by mělo být na místě. Velké společnosti jako Vodafone již tuto metodu ověřování používají, stejně jako velké globální banky jako Barclays a HSBC. [29]

3.4 Domácí asistenti

Domácí asistent je zařízení, určené pro automatizaci domácnosti, ovládání zařízení na základě hlasových příkazů a v některých případech i na pomoc v domácnosti. Tato zařízení jsou velkým bodem zájmu technicky zaměřených společností. Na trhu je možné najít desítky elektronických pomocníků se zaměřením na pomoc v domácnosti a hlasové ovládání zařízení propojených pomocí internetu nebo lokální sítě. Domácí asistenti používají pro své funkce, jako například přehrávání hudby, vyhledávání informací nebo řízení jiných připojených zařízení v domácnosti, umělou inteligenci. [30]

3.4.1 Amazon Echo

Amazon Echo je hlasový domácí asistent od firmy Amazon, který byl uveden na trh v roce 2015. Byl to první domácí asistent disponující vlastním hardwarovým rozhraním, díky čemuž se stal prvním hlasově aktivovaným bezdrátovým reproduktorem. Obsahuje sedm mikrofونů reagujících na příkazy zadávané uživatelem. Jméno osobního inteligentního asistenta implementovaného v Amazon Echo je Alexa. Dokáže nastavit budík, vytvářet seznamy, streamovat podcasty, přehrávat hudbu a dodávat informace o dopravě a počasí v reálném čase. Původním důvodem vynalezení Amazon Echo bylo, avšak zjednodušení nakupování na internetových stránkách Amazonu. Prostřednictvím daného zařízení je dnes možné si nechat domů poslat téměř cokoli, a to jen za pomoci hlasových příkazů. zmíněný hlasový asistent je také velmi přizpůsobivý i v oblasti připojení služeb třetí strany. Dokáže zavolat Uber, přehrát oblíbený seznam skladeb ze služby Spotify nebo i objednat pizzu.

Chytré reproduktory s asistentkou Alexou jsou velice oblíbené v USA [31] nebo v případě Evropy v Německu. Jejich nevýhodou však je mírně horší umělá inteligence při porovnání s konkurencí a nulová podpora češtiny. Tyto reproduktory také velice těžší z propojení s Amazonem, který u nás bohužel nepůsobí, a to jim ubírá velkou část z jejich výhod. Další menší nevýhodou je obecně horší úroveň asistentky Alexi při porovnání s

konkurencí. Alexa si vede mírně hůř než Siri a z trojice porovnávaných hlasových asistentů je na tom nejhůře (přestože jen o malý kousek). [32]

- Příznivá cena zařízení s Alexou
- Široká podpora napříč chytrými spotřebiči
- Neustále se lepší umělá inteligence
- Benefity z propojení s Amazonem, možnost nakupovat hlasem
- Amazon u nás nepůsobí
- Z porovnávané trojice má nejhorší výsledky, co se umělé inteligence týče



Obrázek 1 Amazon Echo 4. generace [33]

3.4.2 Apple HomePod

V červenci 2017 oznámila společnost Apple vydání svého inteligentního reproduktoru, podporujícího technologii Siri. Cílem zařízení Apple HomePod bylo nabídnout uživatelům zbývajících zařízení vyvíjených společností Apple co nejlepší zážitek především při poslechu hudby. Proto obsahuje pole sedmi reproduktorů a subwoofer s vlastním motorem, které spolu přinášejí ten nejlepší zvuk mezi zařízení podobného typu. Zážitek je umocněn šest-mikrofonním polem, které umožňuje zařízení rozeznat hlasové příkazy i během hlasitého přehrávání hudby. Mezi jeho četné funkce je zahrnuto podávání informace

o počasí, připomínky, překladač nebo ovládání inteligentních domácích zařízení, které však musí být podporovány službou HomeKit, což je nevýhodou v porovnání s ostatními zařízeními, vzhledem k jejich ceně. Daná nevýhoda se prokázala v krátkém čase po zahájení prodeje, kdy se počet prodaných kusů zařízení Apple HomePod pohyboval pouze v 3 % celkového počtu prodaných domácích asistentů. [34]

Siri je již velice dlouho součástí iPhoneů, ale tohoto chytrého asistenta lze nalézt i jako střed ekosystému Apple HomeKit. Siri má v testech o něco lepší výsledky než Amazon Alexa, ale poměrně rozdíl je opravdu malý. Na chytrého Asistenta od Google Siri stále ztrácí. Velkým benefitem je snadná integrace do celého ekosystému produktů Apple. [35]

- Snadná integrace do ekosystému Apple
- Umělá inteligence na poměrně vysoké úrovni
- Snadná integrace s chytrými spotřebiči
- Nutnost využívat služby Apple (Apple Music, atd...)
- Poměrně drahá zařízení
- Je nutné využít ekosystém produktů Google



Obrázek 2 Apple Home Pod [36]

3.4.3 Google Home

V současnosti největším konkurentem zařízení Amazon Echo je vynález společnosti Google, Google Home-chytrý reproduktor, který pro své fungování využívá inteligentního asistenta Google Assistant. Google Home je rozměrově menší ve srovnání se zařízením od společnosti Amazon.com. Dokáže ovládat službu Google Play, Spotify, vytvářet úlohy

a připomínky v Google kalendáři, informovat o novinkách ve světě, pracovat se službou Netflix nebo platformou YouTube pomocí Chromecastu. Základní myšlenkou daného zařízení však je ovládání elektronických zařízení v celé domácnosti. Pomocí internetového připojení dokáže ovládat jakékoliv na to uzpůsobené zařízení, například inteligentní žárovky, senzory, garážová vrata, kávovar a mnoho dalších. [37]

Silnou stránkou zařízení Google Home je ohromná schopnost Google asistenta odpovídat na otázky. Ať už je to nejrychlejší cesta do práce nebo školy, předpověď počasí v jakékoli části Země nebo aktuální cena ropy na trhu, Google asistent zná odpověď.

Google Assistant je aktuálně na vzestupu v podstatě na všech trzích. Výhodou tohoto chytrého asistenta je pokročilá AI, která v nezávislých testech poměrně přesvědčivě vítězí nad Amazonem i Applem. Umělá inteligence se také rychle zlepšuje a může těžit z rozsáhlého technického zázemí a dat, která Google na rozvoj umělé inteligence a Google Asistenta uvolňuje. Chytré reproduktory pro chytrou domácnost od Googlu mohou také využívat celý ekosystém produktů od Googlu a výhody z propojení s operačním systémem Android.

- Vysoká úroveň umělé inteligence
- Možnost těžit z ekosystému zařízení a služeb od Googlu
- Rychle se zlepšující AI
- Příznivá cena chytrých reproduktorů
- Asistent Google je součástí i některých vybraných telefonů s Androidem
- špatná kompatibilita s produkty od Apple



Obrázek 3 Google Home [38]

3.5 Ostatní hlasoví asistenti

3.5.1 Microsoft Cortana

Microsoft Cortana je virtuální asistent, který je součástí operačních systémů Windows Phone 8.1 a Windows 10 od společnosti Microsoft. Na základě hlasových povelů umí pracovat s plánovacím kalendářem, udržovat seznamy kontaktů, otevírat aplikace, posílat e-maily i SMS, synchronizovat data mezi mobilním telefonem a počítačem a další podobné funkce. [39]

Cortana byla poprvé představena Joem Belfiorem na konferenci Microsoft BUILD 2014 v San Franciscu. Cortana je dále dostupná také na herních konzolích Xbox One. Název Cortana je odvozen od fiktivní postavy ze série her Halo. Hlas Cortany patří, stejně jako v sérii Halo, Jen Taylorové.

Cortana zatím rozumí jen angličtině, čínštině (v čínských regionech se jmenuje Xiao Na) a od roku 2014 také italštině, francouzštině, němčině a španělštině. [40] Tyto verze byly nejprve dostupné jen ve verzi ALPHA, to znamená, že jej testovali sami uživatelé a Microsoft (formou zpětné vazby) sbíral informace o používání atd. V těchto jazykových verzích Cortana zdaleka neumí všechno to, co umí její anglicky mluvící verze. Cortana čerpá data převážně z vyhledávače Bing, ale také z Wikipedie nebo z aplikací od Microsoftu

(například: Počasí nebo Finance). Dalším zdrojem informací jsou pro ni aplikace v počítači nebo mobilu, které její uživatel používá.

3.5.2 Samsung Bixby

Bixby je virtuální asistent vyvinutý společností Samsung Electronics. Samsung 20. března 2017 oznámil, že se chystá vydat hlasového digitálního asistenta jménem "Bixby". Bixby byl představen společně s Samsung Galaxy S8 a S8+ v průběhu Samsung Galaxy Unpacked 2017 event, která se konala 29. března 2017. Některé součásti Bixby mohou být také staženy pomocí apk balíčku i na starší Galaxy zařízení s operačním systémem Android Nougat. Bixby představuje hlavní předělanou aplikaci S Voice, kterou Samsung představil v roce 2012 společně s Galaxy S II. V květnu roku 2017, Samsung oznámil, že Bixby bude zařazen do série Family Hub 2.0, což je první produkt, ke kterému není potřeba mobilní telefon a byl do něj přidán virtuální asistent Bixby. [41]

Bixby se skládá ze čtyř aplikací. Jsou to: Bixby Voice, Bixby Vision, Bixby Home a Bixby Reminder.

- Bixby Voice může uživatel spustit tím, že někomu zatelefonuje nebo stiskne a přidrží tlačítko umístěné pod tlačítky hlasitosti, nazvané Bixby Button. Chvilí před představením telefonu bylo tlačítko Bixby Button přeprogramované a uživatel si jej mohl nastavit na otevření dalších aplikací nebo asistentů, například konkurenčnímu Google Assistant. Pár dní před spuštěním prodeje telefonu Galaxy S8 byla tato funkce odstraněna firmwarovou aktualizací.
- Bixby Vision je umístěn do systémové aplikace fotoaparát a může identifikovat objekty v reálném čase, vyhledávat pro ně různé služby, nabídnout výrobek k nákupu, je-li nákup k dispozici. Bixby je také schopen překládat text, přečíst QR kódy a poznat památky. [42]
- Bixby Home (s příchodem OneUI 2.0 už jen jako Samsung Daily) lze nalézt přejetím zleva doprava na domovské obrazovce. Lze vertikálně rolovat seznamem informací, které Bixby připravil. Může komunikovat například s počasím nebo fitness aktivitami (z předinstalované aplikace Samsung Health).

Bixby podporuje více jazyků, včetně angličtiny, korejštiny, němčiny, italštiny, čínštiny atd. Tyto jazyky podporují kontextové a vizuální vyhledávání. V listopadu 2019 se Bixby

dostalo dalšího rozšíření s plánovanou implementací do dalších druhů chytrých zařízení od Samsungu.

Samsung neustále rozšiřuje svou nabídku jazyků, a nyní již podporuje více než 10 jazyků včetně 6. dubna 2017 představené němčiny. Bixby v korejštině byl spuštěn 1. května 2017. Bixby je dostupný celosvětově avšak je limitován úzkou nabídkou jazyků. Čínská verze Bixbyho je k dispozici pouze na zařízeních, které se oficiálně prodávají v Číně.

3.5.3 Seznam hlasový asistent

Seznam.cz začal pracovat na vlastním digitálním asistentovi a období zahraničních služeb Apple Siri, Google Now nebo Amazon Alexa. Jako maskot slouží pes Krasty, služba se však nejmenuje po něm, ale nese označení Seznam hlasový asistent. V současné době je k dispozici veřejná testovací verze, a to pouze pro chytré telefony s Androidem 10 (nikoliv s nejnovější jedenáctou edicí).

Seznam s hlasovým asistentem přichází do doby, kdy se zejména Google, Apple a Amazon snaží uživatele naučit ovládat elektroniku a služby na ní navázané hlasovými povely. Podobně Microsoft delší dobu budoval asistentku Cortana, postupně ji ale spíše utlumuje. Velká trojka každopádně vedle samotných softwarových/cloudových pozadí těchto služeb postupně vypouští nový hardware (HomePod, Echo a spol.), aby adopci podpořil.

Hlasového asistenta Seznamu si může vyzkoušet každý. V Google Play Store je třeba si stáhnout aplikaci Seznam.cz, jít do nastavení a v sekci „Vychytávky“ službu aktivovat. Aplikace si vyžádá přístup k mikrofonu a také je třeba v nastavení povolit aplikaci Seznam.cz chod nad dalšími programy. Následně se totiž po stranách displeje začne zobrazovat ikona psa Krastyho, který čeká na hlasové povely. Krastyho lze po okrajích přesouvat podobně jako chatové bubliny.

Hlasový asistent se pak aktivuje povelom „Sezname“. Místo Krastyho se objeví kruh aplikací, se kterými asistent aktuálně spolupracuje. Jde o Seznam.cz, Mapy.cz, Jízdní řády, Email.cz, Novinky.cz, Sport.cz, TV Seznam, Pohádky, Super.cz, TV program a Spotify.

Množina povelů, se kterými hlasový asistent dokáže pracovat, je prozatím omezená. Krasty například umí otevřít aplikaci Mapy.cz, prozatím ale nezvládne pomocí hlasu vyhledat adresu či nastavení navigace. Tyto možnosti přibudou až v budoucnu. Dále lze přehrávat hudbu ze Spotify („Sezname, přehraj Children of Bodom“), ukazovat zprávy,

přehrávat videa ze služeb Seznamu („Sezname, přehraj Lajnu“), dodávat základní informace nebo nastavovat věci jako budík. Povel „Sezname, zavři se“ lze asistenta ukončit.

Česká internetová jednička asistenta vypustila po sérii interních a uzavřených testů. S pomocí uživatelů v rámci veřejného testování chce zrychlit učení celého systému a službu průběžně zlepšovat. [43]

3.6 Stav trhu hlasových asistentů

Trh s hlasovými asistenty se právě teď potýká s těžkou konkurencí, zejména mezi Amazonem a Googlem. Každý z nich těží z jiných částí – Amazon a Alexa mohou mít v jakémkoli zařízení v domácnosti a zastávat silnou dominantní pozici v prostoru chytrých reproduktorů, což je aktuální hnací motor v oblasti hlasu. Silně upřednostňuje společnosti, které vytvářejí aplikace, které přidávají další funkce, a snaží se usnadnit celý proces vývoje, do kterého investuje značné prostředky. Google je v oblasti chytrých reproduktorů v defenzivě a musí chránit svůj podíl na trhu, odborníci se shodují na bodu, který stojí za zmínku, že Google v prostoru chytrých reproduktorů prohrál. Ale chytré reproduktory nejsou všechno, Google nedávno vydal prohlášení, že Google Assistant je na více než 1 miliardě zařízení. [44] Většinu toho zvládají mobilní telefony se systémem Android, ale Susan Westwater poznamenává, že lidé si často ani neuvědomují, že mají na svém mobilním zařízení hlasového asistenta.

Google se však nezaměřuje pouze na mobilní telefony, nedávno získal společnost Fitbit, která podniká především v oblasti chytrých hodinek. To vám umožní zařadit asistenta do této kategorie a být blíže uživateli. Akvizice navíc vytváří další překážku pro vstup konkurentů, jako je Amazon, do segmentu. Alexis Hue, rizikový kapitalista a generální ředitel společnosti Voice and You, proto věří, že pro Amazon bude obtížné dostat se z domu, protože nemají mobilní telefony ani jiná zařízení určená pro venkovní použití. [45]

3.6.1 Uživatelé a firmy

Co se týče amerického trhu, tak lidé si stále zvykají na to, že existuje hlas a asistenti i po několika letech jsou stále používány na jednoduché příkazy, což také potvrzují data

z teoretické části a stejně tak i data z vyscrapovaných stránek Amazonu. Jednoduché příkazy a jednoduché aplikace jsou tím nejpoužívanějším.

Použití hlasu je pak jiná věc. Stále se hledají nové způsoby jeho použití. Hlasová technologie se neustále vyvíjí a stejně tak se vyvíjí i myšlenka používat hlas. Odborníci se shodují, že již začínáme rozumět hlasu a jeho možnému využití, ale v oblasti hlasových technologií nedošlo k velkému rozmachu. Nyní, podle Google Trends, můžeme sledovat vývoj zájmu o Alexa a Google Assistant za poslední 3 roky, přičemž největší zájem byl během vánočního období, což představuje vrcholy k dosažení obchodních cílů, o společnostech se říká, že jsou v tomto ohledu pozadu.

Společnosti by měly přemýšlet o hlasové technologii, ale vše musí být v souladu s celkovou strategií společnosti. Pokud se společnosti rozhodnou pouze vytvářet hlasové aplikace bez další podpory, nebudou moci uspět. S rostoucím počtem hlasových aplikací není snadné uspět bez aplikací, které komunikují na jiných kanálech. Pouze třetina aplikací má hodnocení a ani ne 10 % aplikací má více než 5 hodnocení. Navíc 50 % uživatelů aplikaci aktivně nevyhledává, takže i když má aplikace vysokou pozici ve své kategorii, má k ní přístup pouze polovina z celkového počtu uživatelů.

3.6.2 Přehled trhu

V prognózovaném období 2021–2026 se předpokládá růst trhu s aplikacemi hlasových asistentů o 27,3 % CAGR. Hlasový asistent je především digitální asistent postavený na používání technologií AI, strojového učení a rozpoznávání hlasu. Aplikace hlasového asistenta jsou schopny sledovat hlasové příkazy a pomáhat při rutinních pracích, jako je zadávání online objednávek, plánování schůzek, zapínání připojených světel a funguje jako bezproblémový zprostředkovatel pro psaní textových zpráv nebo volání.

Podle „Cisco Annual Internet Report“ bude počet zařízení připojených k internetovým sítím do roku 2023 více než trojnásobkem celosvětové populace. Podle této zprávy bude na jednoho obyvatele do roku 2023 připadat přibližně 3,6 síťových zařízení, oproti 2,4 zařízení na obyvatele v roce 2018. Rovněž dále předpokládá, že do roku 2023 bude 29,3 miliardy síťových zařízení, což představuje nárůst z 18,4 miliardy v roce 2018. Každý rok se na trhu vyvíjejí a přijímají různá nová zařízení v různých formách s vylepšenými schopnostmi a inteligencí. Vzhledem k tomu, že počet připojených zařízení roste, bude to vést k přijetí řešení aplikací hlasového asistenta. [46]

Některé aplikace hlasového asistenta, které získaly uznání ve všech sektorech, jsou Siri společnosti Apple, Cortana společnosti Microsoft a Asistent Google. Malé a střední podniky se navíc podílejí na řízení technologie inteligentních hlasových asistentů, čímž se stávají klíčovými účastníky trhu s aplikacemi hlasových asistentů.

Hlasové systémy, které uživatelům umožňují vyslovit dotaz, jsou mezi podniky stále běžnější. Hlasové technologie jsou však známé jako ničitelé šířky pásma a mohou snadno přetížít výkon sítě, zejména pokud používají cloudovou infrastrukturu. Očekává se, že to poškodí přijetí hlasových technologií pro podniky, které s takovými problémy bojují. Edge computing je jednou ze slibných technologií, které mohou podporovat zařízení podporující hlas. Technologie Edge Computing umísťuje vysoce výkonná úložiště a síťové prostředky co nejbližší koncovým uživatelům a zařízením.

Hráči neustále pracují na vývoji aplikací hlasového asistenta s podporou edge computingu. Například v prosinci 2019 oznámil Sensory, vývojář AI pro edge computing, spuštění speciální verze platformy pro rozpoznávání řeči TrulyNatural pro domácí spotřebiče. Výrobci elektronických zařízení využijí nově vyvinutou technologii k přidání funkcí hlasového ovládání do svých zařízení.

Podobně další společnosti, jako je Microsoft, v roce 2019 rozšiřují roli Cortany, aby získaly konkurenční výhodu. Společnost chce, aby byla Cortana autentickým hlasovým asistentem, což je funkce, která vždy pracuje na pozadí a pomáhá spotřebiteli být produktivní. Společnost plánuje zavést Cortanu do Outlooku a získat nové možnosti, které lidem pomohou zůstat v pohybu, i když jsou v práci. [47]

V době pandemie COVID-19 nabízely populární aplikace hlasových asistentů poměrně přímočaré pohledy na rostoucí případy a bezpečnostní pokyny WHO, CDC a NIH.

V dubnu 2020 spustila Univerzita v Paříži, francouzská železniční společnost SNCF a francouzský výzkumný ústav INRIA hlasového asistenta, který může pomoci volajícím s potenciálními příznaky COVID-19. Řešení je také instruuje k záchranným službám nebo jejich lékařům pomocí AI a dalších špičkových technologií.

3.6.3 Klíčové trendy trhu

Očekává se, že odvětví zdravotní péče v prognózovaném období výrazně poroste. Poskytovatelé zdravotní péče analyzují výhody aplikací hlasového asistenta pro zvýšení efektivity a zlepšení zkušeností pacientů. Z tohoto důvodu několik značek ve zdravotnictví

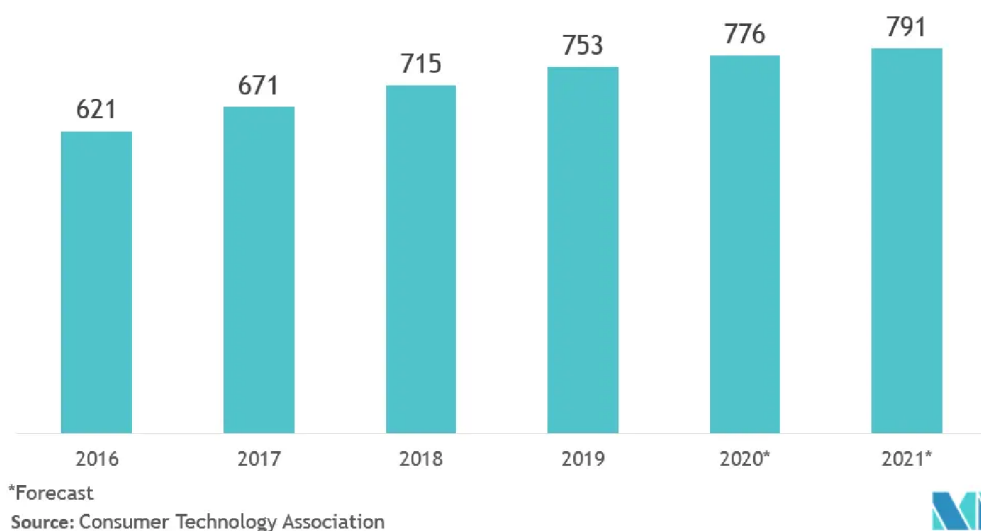
chytře využívá technologii podporovanou hlasem, což dává konkrétní důvody věřit v její kompetenci.

Aplikace hlasového asistenta ve zdravotnictví mohou pomoci při zvládnání chronických onemocnění pacientů. Například Sugarpod, řešení pro řízení diabetu 2. typu, které má vlastní úkoly pro pacienty s diabetem, vychází z jejich preferencí. Sugarpod, když je integrován s hlasovými asistenty, umožňuje pacientům řídit své léčebné postupy a efektivně sledovat pokrok.

Poskytovatelé technologií se také zaměřují na vývoj aplikací hlasových asistentů pro zdravotnický průmysl. Například společnost Nuance Communications, která je softwarovou společností, vyvinula hlasového asistenta speciálně pro zdravotnický prostor s názvem Dragon Medical Virtual Assistant. Řešení přináší propracované konverzační dialogy a dovednosti ve zdravotnictví, které automatizují vysoce hodnotné klinické úkoly od vyhledávání v grafu po CPOE. [48]

V červnu 2019 Saykara oznámila vydání Kara 2.0, zdravotnického asistenta poháněného umělou inteligencí, který lékařům zjednodušuje dokumentační proces. Kara 2.0 je inovativní hlasová aplikace poháněná umělou inteligencí pro zdravotnictví, která umožňuje pacientům a lékařům komunikovat jako obvykle. V době rozhovoru mezi pacientem a lékařem aplikace hlasového asistenta naslouchá, přepisuje text, zpracovává text do strukturovaných dat a vyplňuje každý formulář v EHR pacienta. Tyto nově se objevující technologické trendy jsou hybnou silou růstu tohoto segmentu na trhu.

Unit Shipments of Connected Consumer Devices (in Millions), United States, 2016-2021*



Obrázek číslo 4 Předpověď růstu trhu (jednotky v milionech dolarů) [49]

3.7 Potenciální budoucí použití

Hlasoví asistenti mají potenciál radikálně změnit způsob interakce uživatelů s počítači. Pro mnoho uživatelů je schopnost číst a psát překážkou v přístupu k informacím. Hlasoví asistenti mohou těmto uživatelům překlenout informační mezeru. Nedávný výzkum ukázal, že hlasoví asistenti mohou prospět osobám trpícím demencí tím, že poskytnou všudypřítomný hlas, který dokáže znovu a znovu odpovídat na stejné otázky, aniž by ztratil trpělivost a v případě potřeby nabídl povzbuzení. [50] Pro ostatní může být obtížné přečíst si pokyny, které jim poskytne lékař. Budování těchto schopností do aktuálně dostupného spotřebitele technologie by byly mnohem nákladově efektivnější než účelové zařízení a mnoho uživatelů by již bylo pohodlné s ovládáním těchto zařízení. Hlasoví asistenti mohli uživatelům také číst knihy a jiné dokumenty v dlouhém formátu. I když stále znějí trochu jako roboti, hlasové kvality hlasových asistentů se rychle zlepšují. Jakmile se zdokonalí natolik, aby nevycházely, každá kniha mohla být audioknihou.

3.7.1 Překladače

Hlasoví asistenti mají také potenciál převratu v překladu. Google nedávno oznámil novou sadu sluchátek, která se spárují s hlasovým asistentem pro hlasový překlad v reálném čase. Uživatelé spustí aplikaci tím, že požádají asistenta, aby jim pomohl mluvit jazykem. Telefon uživatele zachycuje zvuk, který druhá strana předává, přenáší jej na překladatelské servery Google a přeloženou verzi přehrává ve sluchátku uživatele. Když uživatel chce svou řeč přeložen, stiskne tlačítko na sluchátku, které odešle zvuk na server, kde je přeložen a předán posluchači pomocí reproduktoru na sluchátku uživatele. Přestože je překlad společnosti Google notoricky náchylný k chybám, zejména u lékařského nebo hovorového jazyka, výsledky téměř okamžité by měly být užitečné pro jednoduché konverzace. [51]

3.7.2 Knihovny

Hlasoví asistenti mohou být užiteční také pro propagaci a správu knihoven. K dispozici jsou již nástroje, které knihovnám umožňují vytvářet dovednosti pro hlasové

asistenty, kteří seznamují události v knihovně v místních komunitních kalendářích. Kódování dalších funkcí, které by čtenářům umožnilo slyšet hodiny a oznámení, obnovovat jejich položky, slyšet, jaké nové položky jsou k dispozici, nebo naplánovat konzultaci s knihovníkem, by bylo relativně jednoduché. Více složité úkoly, jako je prohledávání databází a žádosti o meziknihovní výpůjčky, jsou pravděpodobně mimo rozsah nástroje, který může poskytovat zpětnou vazbu pouze hlasem. [52]

3.7.3 Průvodci

Hlasové asistenty lze také snadno naprogramovat tak, aby fungovali jako virtuální průvodci v menších galerijních nebo výstavních prostorách. Čtenáři mohli požádat asistenta, aby jim řekl o výstavě, a asistent si může přečíst připravené poznámky. Knihovny se zaměřením na technologie možná budou chtít zvážit zapůjčení těchto zařízení a poskytnutí základního školení, aby čtenáři mohli experimentovat s těmito zařízeními také ve svých domovech.

4 Vlastní řešení

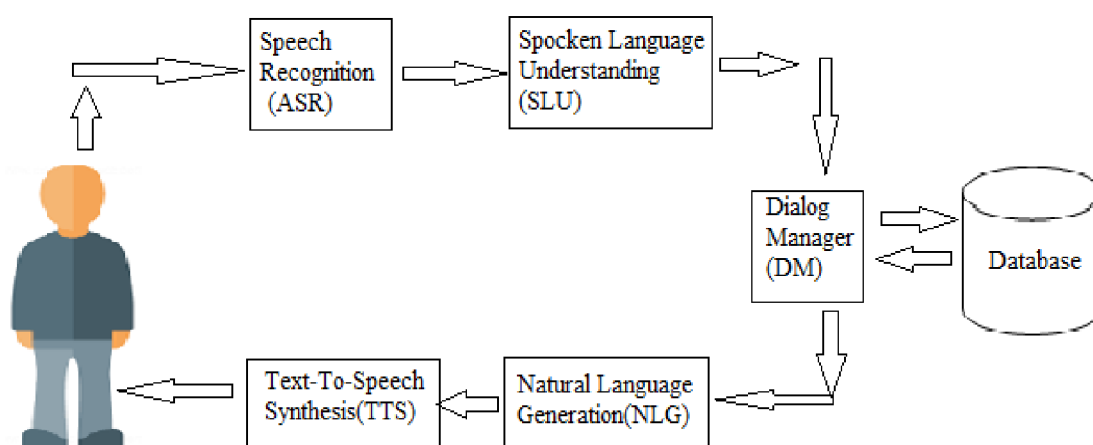
Jako způsob vlastního řešení bylo zvoleno naprogramování softwaru aplikace hlasového asistenta v programovacím jazyce Python. Nástroj na programování v tomto jazyce byla použita aplikace PyCharm, opatřena z webové stránky jetbrains.com.

4.1 Analýza a komparace řešení

4.1.1 Úvod do řešení hlasových asistentů

Navrhování a implementace přirozených a intuitivních způsobů interakce je primárním výzkumným polem v doméně interakce člověk-počítač. Systémy, které mohou komunikovat s uživatelem v jejich přirozeném jazyce, jsou v současnosti intenzivně zkoumány. Virtuální osobní asistenti jsou s příchodem zařízení a technologií jako Google Home, Siri atd. stále populárnější. Uživatelé dokončují úkoly efektivněji pomocí systémů mluveného dialogu prostřednictvím mluvené interakce. Také systémy mluveného dialogu jsou tak začleňovány do různých zařízení, jako jsou chytré televizory, chytré telefony, v automobilovém navigačním systému. Dialogové systémy nebo konverzační systémy mohou

také podporovat širokou škálu aplikací nebo zařízení v obchodních podnicích, vzdělávání, státní správě, zdravotnictví a zábavě. V posledních letech došlo k výraznému nárůstu nasazení a používání hlasem ovládaných osobních digitálních asistentů. Během posledních desetiletí byla nezbytná integrace zařízení a technologií (tj. kombinace správy dialogů a syntézy textu na řeč, porozumění přirozenému jazyku, automatického rozpoznávání řeči) středem zájmu rozsáhlého akademického a průmyslového výzkumu, který nakonec vyústil v komerční produkty, jako je Apple Siri, Google Home, Amazon Alexa a Microsoft Cortana. Tyto produkty nejen ukazují pokroky či nové technologie, ale dokážou dovést koncept umělého osobního asistenta, tedy systému, který je schopen, alespoň do určité míry rozumět a reagovat na mluvené vstupy.



Obrázek 5 Systém dialogů [54]

Neustálý pokrok a aktivity probíhaly jak z technologického hlediska, tak i s ohledem na podporované aplikační domény. Umělé služby, které mohou začít, pokračovat a zvládnout složité interakce s lidskými partnery v jejich přirozeném jazyce. Mnoho společností používá systémy mluveného dialogu k návrhu svého zařízení pro dialog nebo spouštění takových programů, jako je Cortana od společnosti Microsoft, Siri společnosti Apple, Amazon Alexa a Google Assistant. Společnosti používaly různé přístupy k navrhování a zlepšování svých dialogových systémů. Google například vylepšil Google Assistant pomocí metody DNN, která zvýrazňuje hlavní komponenty dialogových systémů a nové architektury hlubokého učení používané pro tyto komponenty efektivně. Google vyvinul virtuálního osobního asistenta, Google Assistant pro mobilní zařízení a zařízení pro chytrou domácnost. Na rozdíl od Google Now se Asistent Google může zapojit

do obousměrných konverzací nebo v nich pracovat. Microsoft použil Microsoft Azure Machine Learning Studio s dalšími komponentami Azure ke zlepšení a zefektivnění dialogového systému Cortana. Amazon poskytuje pokročilé funkce hlubokého učení a pro převod řeči na text využívá služby automatického rozpoznávání řeči a rozpoznávání záměru textu a také využívá porozumění přirozenému jazyku pro vytváření vysoce poutavých uživatelských zkušeností, jako jsou konverzační interakce. pro vývojáře. Siri je virtuální asistentka s hlasově ovládaným rozhraním přirozeného jazyka, která využívá kontextové povědomí a sekvenční vyvozování, aby pomohla uživatelům iOS provádět osobní úkoly nebo operace. Siri je program umělé inteligence sestávající ze strojového učení, zpracování přirozeného jazyka a algoritmu vyhledávání na webu. Siri a tato technologie je integrována do zařízení iPhone, iPod touch a iPad a do nového systému domácí automatizace Home Kit společnosti Apple. Cortana je operační systém, který je orientovaným hlasovým asistentem dostupným na zařízeních s Windows 10 a také na konzolích Xbox One. Cortana je úzce spojena s produktivitou a jejími aktivitami.

4.1.2 Všeobecná studie hlasových asistentů

Název zařízení	Siri	Google assistant	Cortana	Alexa
Operační systém	iOS 5, macOS sierra, tvOs, watchOS	Android, iOS, KaiOS (jio Phone)	Windows, iOS, Android, Xbox OS	Fire OS, iOS, Android
Platforma	iPhone, iPad, Mac, Apple TV, Apple Watch, HomePod	Android, Google Home, Android TV, Smart Speakers, Headphones, Small Speakers, Headphones, Small Displays,	Windows 10, Windows 10 mobile, Windows Phone 8.1, Harmon Kardon Invoke, Microsoft Band 2, Microsoft Band, Android, Xbox One, Skype, iOS, Cyanogens	Amazon Echo, Fire TV, Fire HD, iOS, Android

		Google Allo, iOS	OS, Windows Mixed Reality, Amazon Alexa	
Dostupné jazyky	angličtina, arabština, čínština, dánština, holandština, finština, francouzština, němčina, hebrejština, italština, japonština, korejština, malajština, norština, portugalština, ruština, španělština, švédština, thaiština, turečtina	angličtina, hindština, holandština, francouzština, němčina, italština, japonština, korejština, portugalština, španělština	angličtina, čínština, francouzština, němčina, italština, japonština, portugalština, španělština	angličtina, němčina, japonština
Typ zařízení	Inteligentní osobní asistent	Inteligentní osobní asistent	Inteligentní osobní asistent	Inteligentní osobní asistent, cloudová hlasová služba

Tabulka 1 Všeobecná analýza hlasových asistentů

4.1.3 Komparace mezi hlasovými asistenty

Asistent	Hlasová zařízení	Přednosti
Amazon Alexa	Amazon Echo Dot (druhá generace), Amazon Echo, Amazon Echo plus, Amazon Echo spot, Sonos one, Amazon echo show	Vylepšený zvuk Dolby, 2,5palcový displej, hudba pro více místností, zesilovače třídy D, vlastní ovladače, videohovory
Microsoft Cortana	Harmon Kardon Invoke	360stupňový zvuk Zvuk, možnost propojení s aplikací Skype

		Integrace s ostatními zařízeními Hlasová technologie vzdáleného pole
Apple Siri	Apple HomePod	Automatické ladění hudby, 360stupňový zvuk, sedm výškových reproduktorů, Airplay 2 pro poslech ve více místnostech
Google assistant	Google Home Max, Google home, Google Home Mini, Sony LF-S50G, JBK	Streamování Chromecastu, individuální rozpoznávání hlasu, google translate, Duální 4,5palcové basové reproduktory, vlastní výškové reproduktory, chytrá optimalizace zvukové místnosti, optimalizace vzdálené místnosti, hlasové ovládání vzdáleného pole, bezdrátové připojení s více reproduktory párování, 360stupňový zvuk, ovládání gesty, odolný proti stříkající vodě, zobrazení hodin, výdrž baterie 10 hodin, výstup 50 W

Tabulka 2 Komparace řešení

4.1.4 Omezení a budoucí možnosti využití

Virtuální asistent	Omezení	Budoucí možnosti
Apple Siri	Siri nefunguje se slabými mobilními službami. Pro interakci cloudového serveru vyžaduje silné mobilní připojení. Nízká mobilní síť generuje omezení ve funkčnosti Siri. Rozumí, pouze	V dnešní době se více využívají digitální technologie. Pokud bude Siri schopna získat a využívat digitální technologie, má Siri velkou budoucnost.

	<p>pokud mluvíte přesně. Jazykové problémy mohou znamenat potíže rozpoznání hlasu při zadávání požadavků a instrukcí. Siri neovládá prvky zábavy.</p>	
Microsoft Cortana	<p>V jednu chvíli může Cortanu používat pouze jedna osoba. Pouze „Hey Cortana“ se používá ve formě klíčového slova. Když Cortana začíná odemknutím obrazovky, provádí pouze omezené úkoly, ne všechny možné, protože zachovává soukromí. Více kompatibilní s Windows, nemá příliš kompatibilitu ohledně platformy používané vývojáři aplikací a péči o uživatele. Přístup k systému iOS a Android je složitý. Cortana neovládá psaní textových zpráv, volání a psaní poznámek</p>	<p>Cortana by měla přijít s lepšími hlasovými schopnostmi, vyšší rychlostí vyhledávání a dostupností ve více světových jazycích. Cortana by měla automaticky sdělit, kdy se počítač pokazil nebo byl infikován virem. Cortana by měla být schopna se spojit se sociálními účty, aby získala informace o přátelích.</p>
Amazon Alexa	<p>Alexa nerozumí vícenásobné výslovnosti. Přidáním dat s více akcenty angličtiny lze tento problém vyřešit, i když ukládání tak velkého množství dat je také problém. Mnoho hlasových profilů není pro vývojáře k dispozici.</p>	<p>Alexa by měla k řešení svých problémů a problémů s úložištěm použít koncept Big Data.</p>
Google Assistant	<p>Google umožňuje provozovat více účtů současně, a proto je méně zabezpečen. Asistent Google provádí několik úkolů jediným hlasovým příkazem.</p>	<p>Funkce AI v asistentovi Google mu umožní uchovávat data týkající se chování, preferencí a voleb uživatele a pomáhat uživatelům s každodenními úkoly, ať už jde o přípravu do práce nebo</p>

	Google assistant neovládá psaní poznámek.	správu domácích prací v případě chytrých domácností.
--	---	--

Tabulka 3 Omezení a pohled na budoucnost

4.1.5 Závěr analýzy a komparace řešení

Google a Siri si vedly nejlépe v plánování a organizování úkolů. Alexa si nejlépe vedla v online nakupování. Cortana je svou povahou adaptivní. Asistent Google používá Google jako vyhledávač, který je lídrem všech vyhledávačů. Asistent Google je tedy nejlepší hlasový asistent z porovnávaných asistentů. Nicméně, jeho vlastnosti a dovednosti mají také své určité softwarové mezery a nedostatky, a proto byl autorem navrhnut a pilotně implementován vlastní hlasový asistent, kde budou například vyřešeny tyto nedostatky:

- Vyhledávání pouze jedním vyhledávačem
- Psaní textových zpráv
- Volání
- Psaní poznámek
- Zábava

4.2 Návrh aplikace

Hlavním účelem hlasové aplikace je vytvořit rozhraní mezi počítačem řízenou aplikací a uživatelem, který komunikuje hlasem. Je potřeba určit jednotlivé parametry pro návrh interaktivního hlasového rozhraní, které zaručí, aby aplikace byla uživatelsky příjemná, přehledná, byla přínosná pro uživatele a snadno se ovládala a uměla základní prvky pro chod chytré domácnosti.

Je potřeba vybrat kvalitní rozpoznávací systém, který bude zvolen pro rozpoznávání promluvy, tj. přesnost rozpoznávání slov by měla být co nejvyšší, aby se dialog neprotahoval zbytečnými chybami v nerozpoznaných promluvách uživatelů. Dále je nutné rozlišovat pro jakou kategorii lidí a komunikační oblast témat (typy úloh) je návrh tvořen. Pokud se jedná o zkušené uživatele hlasových dialogových systémů, není třeba vkládat časté nápovědy. V opačném případě se nápovědy vloží tak, aby se s interaktivním rozhraním a hlasovým

dialogovým systémem naučil nezkušený, či méně zkušený uživatel a nedocházelo ke špatným promluvám, které rozpoznávací systém není schopen rozpoznat. Tím je zabráněno nežádoucímu prodlužování dialogu. U tvorby interaktivního rozhraní také záleží na oblasti (typu úlohy), kde bude využívána. V současné době není možné vytvořit aplikaci, která by pomocí hlasového digitálního systému mohla ovládat více oblastí najednou.

S parametrem kategorie lidí souvisí další parametr pro návrh a tím je design aplikace. V případě tvorby rozhraní pro kategorii lidí ve věku mladých dětí a mladistvých je nutné, aby aplikace uživatele do určité míry zaujala (barevný vzhled, animované obrázky apod.) a byla uživatelsky příjemná. Výsledkem při návrhu interaktivního hlasového rozhraní by měla být spokojenost potenciálního uživatele s pohodlím a snadností obsluhy aplikace.

4.2.1 Běhové prostředí

Pro běhové prostředí aplikace byl vybrán programovací jazyk Python. Tento programovací jazyk je high-level, objektově orientovaný a interpretovaný jazyk, který umožňuje dynamické typování. Díky tomu, že je tento programovací jazyk open source, disponuje velikou škálou distribucí, podporou různých platforem a velkým množstvím dodatečných balíčků k různému použití. V tomto případě se bude jednat o balíčky vhodné k použití implementace aplikace hlasového asistenta. Dále také disponuje jednoduchou a snadno čitelnou syntaxí. Vzhledem k těmto vlastnostem, umožňuje rychlý vývoj aplikací.

Samotný vývoj aplikace byl rozdělen do několika nezávislých částí – fází, které byly po jejich dokončení vždy alespoň částečně otestovány.

4.2.2 1. Fáze vývoje

V první fázi vývoje byla autorem provedena menší analýza návrhu uživatelského rozhraní pro hlasové asistenty.

- Uživatelský výzkum – O faktu, že uživatel dokončí akci s hlasovým asistentem, mohou rozhodovat drobné detaily a znalosti jeho potřeb, které jsou nejlepší prevencí, jak uživatele nedoradit od dalšího používání asistenčních programů. Je důležité se zaměřit na vyjadřovací způsoby cílové skupiny a používaný způsob komunikace následně přetvořit v komunikační manuál, který bude využit při tvorbě dialogů.

- Nasimulování nejpravděpodobnějších situací – Používání hlasových asistentů má svá specifika. Dají se předpokládat situace, kdy uživatel používá ovládání hlasem častěji, např. při řízení dopravního prostředku, kdy uživatel pouhou jednou větou „play some music“ spustí hudbu, místo komplikovaného manuálního – a v případě řízení dopravního prostředku nebezpečného vyhledávání.
- Předpřipravit jednoduché dialogy – Při psaní dialogů je důležitým faktorem čas, čím déle bude operace trvat, tím horší bude dojem z celé akce. Zdlouhavé procesy mohou úplně odradit od dalšího užívání.
- Počítat s lidskou chybou – Důležité je počítat s nestandardními situacemi kdy se uživatel bude vyjadřovat jinak než spisovně, proto je důležité zahrnout slangová slova do návrhu.

4.2.3 2. Fáze vývoje

V druhé fázi vývoje byly přidány do programovací aplikace PyCharm níže uvedené rozšiřující balíčky. Odstavec obsahuje popis jednotlivých dodatečných modulů a následně kód, který je třeba vypsát do terminálu aplikace, aby byl poté následně možný import všech potřebných modulů.

- Speech Recognition – balíček umožní aplikaci rozpoznat hlas a řeč a reagovat na tyto podněty.
`pip install speechRecognition`
- Pyttsx3 – pyttsx3 obsahuje knihovnu pro převod textu na řeč, která funguje v offline režimu na rozdíl od alternativních knihoven a je kompatibilní jak s nejnovější, tak i se starší verzí Pythonu.
`pip install pyttsx3`
- Pywhatkit – instalací tohoto balíčku bude zajištěno lepší ovládání při vyhledávání v internetovém prohlížeči.
`pip install pywhatkit`
- Wikipedia – Wikipedia je knihovna Pythonu, která usnadňuje přístup k datům z Wikipedie a jejich analýzu.
`pip install wikipedia`

- Pyjokes – na požádání řekne uživateli vtip, tento balíček neslouží přímo pro užívání hlasového asistenta jako pomocníka v různém směru, ale spíše pro odlehčení a zábavu.

```
pip install pyjokes
```

- Twilio – Twilio je rozhraní pro API, které je používáno pro komunikaci, jako jsou telefonní hovory, zprávy a dvou faktorové ověřování.

```
pip install twilio
```

- Wolframalpha – Wolfram Alpha je API, které dokáže vypočítat odpovědi na expertní úrovni pomocí algoritmů, znalostní báze a technologie AI.

```
pip install wolframalpha
```

- Winshell – Modul winshell je lehkým obalem funkcí prostředí Windows. Zahrnuje pohodlné funkce pro přístup ke speciálním složkám, pro použití funkce kopírování, přejmenování a mazání souborů shellu a určitou podporu pro strukturované úložiště.

```
pip install winshell
```

Některé použité balíčky již jsou součástí nainstalované aplikace Python, a proto nebylo nutné je instalovat. Jedná se především o tyto doplňky:

- Subprocess – Tento modul je používán pro získání podrobností o podprocesu systému, které se používají v různých příkazech, např. Shutdown nebo Sleep.
- Tkinter – Tkinter je de-facto standardní balíček GUI, který je objektově orientovaný.
- Web browser – Modul webového prohlížeče je pohodlný ovladač webového prohlížeče. Poskytuje rozhraní na vysoké úrovni, které umožňuje uživatelům zobrazovat dokumenty založené na webu. webový prohlížeč lze také použít jako nástroj CLI.
- Datetime – Modul Datetime v Pythonu poskytuje třídy pro práci s datem a časem. Tyto třídy poskytují řadu funkcí pro práci s daty, časy a časovými intervaly.

4.2.4 3. Fáze vývoje

Třetí fáze vývoje byla zaměřena na importování nainstalovaných rozšiřujících balíčků a na naprogramování základních dovedností hlasového asistenta, kterými jsou:

- Rozpoznání řeči
- Vyhledávání informací

- Ovládání základních funkcí na spuštěném zařízení
- Používání nainstalovaných aplikací
- Ovládat hovory a spravovat zprávy
- Používat navigaci a určování polohy
- Ovládání časových nástrojů např. budík

4.3 Implementace aplikace

Strategie implementace aplikace hlasového asistenta byla realizována v prostředí PyCharm Community Edition 221.4165.171 s Python verzí 3.10.2. V současné době se jedná o jeden z nejpoužívanějších programovacích jazyků. Protože autorova vlastní zkušenost s programovacím jazykem v začátku vytváření bakalářské práce nebyla velká, bylo postupováno od studií daného jazyku, jeho struktury, zápisu, přes vytváření menších lehčích aplikací, které by bylo možné využít pro vložení do samotné aplikace až po samotné programování navržené aplikace.

4.3.1 Import knihoven a nastavení engine aplikace

```
main.py x
1 import ctypes
2 import datetime
3 import json
4 import os
5 import shutil
6 import smtplib
7 import subprocess
8 import time
9 import webbrowser
10 from urllib.request import urlopen
11 import progress
12 import pyjokes
13 import pyttsx3
14 import requests
15 import speech_recognition as sr
16 import wikipedia
17 import winshell
18 import wolframalpha
19 from twilio.rest import Client
```

Obrázek 6 Import knihoven

Knihovny jsou importovány jak z nainstalovaných, tak z předinstalovaných modulů. Jednoduše se mohou některé importované soubory přidat nebo odstranit. Z některých modulů byly nainstalovány ještě tzv. podmoduly. Z `urllib.request` byl naimportován `urlopen`, který slouží pro zpracování a načítání URL adres pomocí různých protokolů. Modul `Client`, získaný z `twilio.rest` zajišťuje lepší komunikaci při hovorech a posílání zpráv.

```
21 engine = pyttsx3.init('sapi5')
22 voices = engine.getProperty('voices')
23 engine.setProperty('voice', voices[1].id)
```

Obrázek 7 Nastavení engine aplikace

Engine pro aplikaci byl nastaven na `Pyttsx3`, který se používá pro převod textu na řeč v Pythonu a `sapi5` je rozhraní platformy Microsoft řečové aplikace, které je

používáno pro funkci převodu textu na řeč. Hlasové ID může být změněno na „0“ pro mužský hlas, v tomto případě byl použit ženský hlas pro veškerý převod textu na řeč, tedy ID „1“.

4.3.2 Vytvoření hlasového asistenta

```
25 def speak(audio):
26     engine.say(audio)
27     engine.runAndWait()
28
29 def wishMe():
30     hour = int(datetime.datetime.now().hour)
31     if hour >= 0 and hour < 12:
32         speak("Good Morning Sir !")
33
34     elif hour >= 12 and hour < 18:
35         speak("Good Afternoon Sir !")
36
37     else:
38         speak("Good Evening Sir !")
39
40     assname = "Assistant 1 point o"
41     speak("I am your Assistant")
42     speak(assname)
43
44 def username():
45     speak("What should i call you sir")
46     uname = takeCommand()
47     speak("Welcome Mister")
48     speak(uname)
49     columns = shutil.get_terminal_size().columns
50
51     print("#####".center(columns))
52     print("Welcome Mr.", uname.center(columns))
53     print("#####".center(columns))
54
55     speak("How can i Help you, Sir")
```

Obrázek 8 Vytvoření hlasového asistenta část 1

```

57 def takeCommand():
58     r = sr.Recognizer()
59
60     with sr.Microphone() as source:
61
62         print("Listening...")
63         r.pause_threshold = 1
64         audio = r.listen(source)
65
66     try:
67         print("Recognizing...")
68         query = r.recognize_google(audio, language='en-in')
69         print(f"User said: {query}\n")
70
71     except Exception as e:
72         print(e)
73         print("Unable to Recognize your voice.")
74         return "None"
75
76     return query
77
78 def sendEmail(to, content):
79     server = smtplib.SMTP('smtp.gmail.com', 587)
80     server.ehlo()
81     server.starttls()
82     server.login('your email id', 'your email password')
83     server.sendmail('your email id', to, content)
84     server.close()
85

```

Obrázek 9 Vytvoření hlasového asistenta část 2

Asistent je vytvořen podle návrhu aplikace, v němž byly zohledněny všechny aspekty správné komunikace mezi uživatelem a počítačem tak, aby komunikace probíhala naprosto bezproblémově.

V Pythonu je funkce definována pomocí klíčového slova `def`. Pomocí tohoto slova byly nastaven základní funkce asistenta, jakými jsou:

- `wishMe` – funkce je nastavena na poznání aktuálního času a podle toho zvolí správný pozdrav
- `username` – uživatel si nastaví své „uživatelské jméno“, kterým bude asistentem osloven

- `takeCommand` – spustí získání příkazu a zahájí jeho vyřizování a zaslání zpětné vazby uživateli
- `sendEmail` – tato funkce byla záměrně vybrána na ukázkou, jelikož aby fungovala, je zapotřebí nastavit v G-mailu nízkou úroveň zabezpečení, při vyšší úrovni to Google pokládá za možnost nabourání do schránky.

4.3.3 Nastavení hlavních funkcí

Nastavení hlavních funkcí bylo vytvořeno z poznatků získaných v teoretické části práce. Jelikož hlasoví asistenti disponují velkou řadou funkcí a možností, pro pilotní implementaci byly pro ukázkou vybrány jedny z běžných funkcí.

```

97     if 'wikipedia' in query:
98         speak('Searching Wikipedia...')
99         query = query.replace("wikipedia", "")
100        results = wikipedia.summary(query, sentences=3)
101        speak("According to Wikipedia")
102        print(results)
103        speak(results)
104
105    elif 'open youtube' in query:
106        speak("Here you go to Youtube\n")
107        webbrowser.open("youtube.com")
108
109    elif 'open google' in query:
110        speak("Here you go to Google\n")
111        webbrowser.open("google.com")
112
113    elif 'play music' in query or "play song" in query:
114        speak("Here you go with music")
115        music_dir = "C:\\Users\\Tomáš Kundrt\\Music"
116        songs = os.listdir(music_dir)
117        print(songs)
118        random = os.startfile(os.path.join(music_dir, songs[1]))
119
120    elif 'the time' in query:
121        strTime = datetime.datetime.now().strftime("% H:% M:% S")
122        speak(f"Sir, the time is {strTime}")
123
124    elif 'open chrome' in query:
125        codePath = r"C:\\Program Files (x86)\\Google\\Chrome\\Application\\chrome.exe"
126        os.startfile(codePath)

```

Obrázek 10 Ukázka nastavení hlavních funkcí část 1

```

175
176 elif "calculate" in query:
177
178     app_id = "Wolframalpha api id"
179     client = wolframalpha.Client(app_id)
180     indx = query.lower().split().index('calculate')
181     query = query.split()[indx + 1:]
182     res = client.query(' '.join(query))
183     answer = next(res.results).text
184     print("The answer is " + answer)
185     speak("The answer is " + answer)
186
187 elif 'search' in query or 'play' in query:
188
189     query = query.replace("search", "")
190     query = query.replace("play", "")
191     webbrowser.open(query)

```

Obrázek 11 Ukázka nastavení hlavních funkcí část 2

```

304 elif "weather" in query:
305
306     api_key = "Api key"
307     base_url = "http://api.openweathermap.org / data / 2.5 / weather?"
308     speak(" City name ")
309     print("City name : ")
310     city_name = takeCommand()
311     complete_url = base_url + "&appid=" + api_key + "&q=" + city_name
312     response = requests.get(complete_url)
313     x = response.json()
314
315     if x["cod"] != "404":
316         y = x["main"]
317         current_temperature = y["temp"]
318         current_pressure = y["pressure"]
319         current_humidiy = y["humidity"]
320         z = x["weather"]
321         weather_description = z[0]["description"]
322         print(" Temperature (in kelvin unit) = " + str(
323             current_temperature) + "\n atmospheric pressure (in hPa unit) = " + str(
324             current_pressure) + "\n humidity (in percentage) = " + str(
325             current_humidiy) + "\n description = " + str(weather_description))
326
327     else:
328         speak(" City Not Found ")
329

```

Obrázek 12 Ukázka nastavení hlavních funkcí část 3

Všechny příkazy, které uživatel zadal asistentovi ke zpracování jsou uloženy v „query“ a konvertovány do tzv. lower case pro snazší porozumění příkazu.

Funkce `weather`, která je jedna z nejkompexnějších ukázkových funkcí, využívá vyhledávač Google pro otevření webové stránky s počasím, aby získala API dané webové stránky a mohla tak ukázat požadovaná data o počasí ve vybrané lokalitě

4.3.4 Vyřešení zjištěných nedostatků a možných hrozeb

Při analýze a komparaci řešení nejpoužívanějších hlasových asistentů byly zjištěny nedostatky, například:

- Psaní poznámek
- Textové zprávy a hovory
- Zábava

```
263 elif "write a note" in query:
264     speak("What should i write, sir")
265     note = takeCommand()
266     file = open('assistant.txt', 'w')
267     speak("Sir, Should i include date and time")
268     snfm = takeCommand()
269     if 'yes' in snfm or 'sure' in snfm:
270         strTime = datetime.datetime.now().strftime("% H:% M:% S")
271         file.write(strTime)
272         file.write(" :- ")
273         file.write(note)
274     else:
275         file.write(note)
276
277 elif "show note" in query:
278     speak("Showing Notes")
279     file = open("assistant.txt", "r")
280     print(file.read())
281     speak(file.read(6))
```

Obrázek 13 Vyřešení nedostatku – poznámky

Jako jeden z nedostatků bylo zjištěno psaní poznámek, a to hned u dvou komparovaných asistentů – Google assistant a Cortana. Pro vyřešení tohoto nedostatku byla naprogramována výše uvedená funkce `write` a `note`, která vytvoří v zařízení textový dokument.


```

329 elif "send message" in query:
330     account_sid = 'Account Sid key'
331     auth_token = 'Auth token'
332     client = Client(account_sid, auth_token)
333     message = client.messages \
334         .create(
335         body=takeCommand(),
336         from_='Sender Number',
337         to='Receiver Number'
338     )
339
340     print(message.sid)
341
342 elif "make a call" in query:
343     account_sid = 'Account Sid key'
344     auth_token = 'Auth token'
345     client = Client(account_sid, auth_token)
346     call = client.calls.create(
347         to='Receiver Number',
348         from_='Sender Number',
349         url='http://demo.twilio.com/docs/voice.xml'
350     )
351
352     print(call.sid)
353

```

Obrázek 14 Vyřešení nedostatku – Textové zprávy a hovory

Dalším zjištěným nedostatkem bylo zaslání textových zpráva provádění hovorů. Tato funkce funguje pomocí dodatečného modulu Twilio. Aby bylo možné tuto funkci využívat, je potřeba si vytvořit účet na službě Twilio.

```

171
172 elif 'tell a joke' in query:
173     speak(pyjokes.get_joke())

```

Obrázek 15 Vyřešení nedostatku – Zábava

Tento nedostatek slouží spíše jako rozšiřující funkce než funkce potřebná pro usnadnění práce uživatele, nicméně byl tento nedostatek zaznamenán u asistenta Siri a jak již bylo zmíněno, jedná se pouze o menší doplněk.

V poslední řadě se autor věnoval bezpečnostním funkcím, které mohou ochránit naše soukromí a zamezit tak odposlouchávání běžné konverzace, které není určena přímo asistentovi. Někteří asistenti totiž po interakci s člověkem a splnění jeho požadavků zůstávají stále aktivní a mohou poslouchat a ukládat konverzace mezi ostatními členy domácnosti. Aby se zamezilo těmto hrozbám, byly v aplikaci nastaveny funkce pro vypnutí, odhlášení ze systému, či přímo funkce pro zastavení poslouchání asistentem.

```
221
222 elif 'lock window' in query:
223     speak("locking the device")
224     ctypes.windll.user32.LockWorkStation()
225
226 elif 'shutdown system' in query:
227     speak("Hold On a Sec ! Your system is on its way to shut down")
228     subprocess.call('shutdown / p /f')
229
230 elif "don't listen" in query or "stop listening" in query:
231     speak("for how much time you want to stop assistant from listening commands")
232     a = int(takeCommand())
233     time.sleep(a)
234     print(a)
235
236 elif 'exit' in query:
237     speak("Thanks for giving me your time")
238     exit()
```

Obrázek 16 Možnosti pro zabezpečení

5 Výsledky a diskuse

5.1 Výsledky

Vyvinutá aplikace pro používání v hlasově ovládaných zařízeních je výsledkem této práce. Aplikace je vyvinuta tak, aby splňovala veškerá kritéria, která uživatel očekává od hlasově ovládaného zařízení. Aplikace je jednoduchá a velice přívětivá, aby i uživatelé, kteří se doposud neseťkali s hlasově ovládacími prvky či jinými zařízeními usnadňující práci jako jsou například chatboti.

Aplikace je vyvinuta nejen pro používání hlasových asistentů v domácnostech, ale i pro používání v mobilních a jim podobných zařízeních. Aplikaci je samozřejmě možné dále rozšiřovat o další funkce, je ale potřeba nejprve zjistit, jestli požadované funkce je možné implementovat v programovacím jazyce Python a běhovém prostředí PyCharm. S úpravami lze tuto aplikaci použít pro handicapované uživatele a usnadnit jim tak každodenní život.

5.2 Diskuse

V současné době je k dispozici velké množství hlasových asistentů, ty nejrozšířenější a nejznámější byly analyzovány v praktické části práce. Při vývoji aplikace byly použity již existující doplňkové moduly a balíčky obsažené v běhovém prostředí. Je to velice pohodlná a efektivní cesta, jak dosáhnout lehčího a rychlejšího výsledku vývoje. Software obsažený v elektronických zařízeních pro ovládání hlasem jsou převážně zatím určeny pouze pro jakousi interakci mezi uživatelem a zařízením a slouží k usnadnění každodenních činností. Ne všichni asistenti však umí to samé, existují asistenti speciálně pro chytrou domácnost, kde převažuje ovládání prvků jako jsou světla, vytápění či ovládání chytrých přístrojů v domácnosti.

6 Závěr

Složitost a přesnost technologie rozpoznávání hlasu a softwaru hlasových asistentů v posledních několika letech exponenciálně vzrostla. V současnosti dostupné produkty hlasových asistentů od společností Apple, Amazon, Google a Microsoft umožňují uživatelům klást otázky a zadávat příkazy počítačům v přirozeném jazyce. Tato technologie může mít v budoucnu mnoho využití, od automatizace domácnosti přes překlady až po společnost a podporu starších lidí. V současnosti dostupné produkty hlasových asistentů však mají také několik problémů. Než se hlasoví asistenti začnou používat k čemukoli, co vyžaduje ochranu soukromí a bezpečnost, bude třeba zlepšit kontrolu soukromí a zabezpečení. důvěrnost.

V první části práce byl popsán princip hlasových asistentů, jejich dovednosti a možnosti využití nejen v chytré domácnosti. Jedním z dílčích cílů práce bylo vypracování přehledu trhu s hlasovými asistenty, kde se autor věnoval všeobecnému přehledu trhu a jeho místem mezi uživateli a firmami. Růst trhu bude velice zajímavé sledovat, jelikož i malé firmy začínají vytvářet své hlasové asistenty a chat boty, a tak se může stát, že některá z těchto firem může v budoucnu konkurovat velkým gigantům na trhu jako je Google či Apple, nebo je dokonce nahradit. Poznatky z teoretické části byly aplikovány ve vlastním řešení praktické části.

Hlavním cílem vlastního řešení práce byla analýza a komparaci řešení s pilotní implementací aplikace hlasového asistenta, kde byly autorem vyřešeny nedostatky zjištěné analýzou řešení. Komparace jednotlivých asistentů proběhla mezi nejznámějšími zástupci domácích hlasových asistentů. Pilotní implementace aplikace probíhala ve 2 fázích: návrh aplikace, kde se autor zaměřil na běhové prostředí a na zkušenosti uživatelů hlasových asistentů s jejich používáním pro vyhnutí se chyb při komunikaci mezi oběma subjekty, a samotná implementace, kde autor vyřešil nedostatky jednotlivých asistentů zjištěné srovnávací analýzou.

Hlasoví asistenti se neustále vyvíjejí a téměř každý měsíc vychází zpráva o nových funkcích, které vylepšují možnosti asistenta. Je to stále nová technologie a ani zahraniční odborníci nemají vše otestováno na 100 %. A sekce jako optimalizace hlasového vyhledávání stále čekají na další informace. Je docela překvapivé, že na úrovni adopce asistenta nezůstávají všechny použitelné části neprozkoumané.

7 Seznam použitých zdrojů

[1] Společnost Cisco: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper [online] 2017. Dostupné z: <<https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>>

[2] Home Smart IoT Home: Domesticating the Internet of Things [online]. Dostupné z: <<https://www.toptal.com/designers/interactive/smart-home-domestic-internet-of-things>>

[3] MORGAN, Jacob, časopis Forbes: A Simple Explanation Of 'The Internet Of Things' [online]. Dostupné z: <<https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#2e0b61b81d09>>

[4] ASHTON, Kevin, RFID Journal: That 'Internet of Things' Thing [online] Dostupné z: [online]. Dostupné z: <<http://www.rfidjournal.com/articles/view?4986>>

[5] International Telecommunications Union: ITU Internet Reports 2005: The Internet of Things [online]. Dostupné z: <http://www.itu.int/osg/spu/publications/internetofthings/InternetofThings_summary.pdf>

[6] Wireless Networking and Communications Group: Device-to-Device (D2D) Communication: Fundamentals with Applications to LTE. [online] Dostupné z: <<https://wncg.org/research/briefs/device-device-d2d-communication-fundamentals-applications-lte>>

[7] "The Enterprise Internet of Things Market". Business Insider. 25 February 2015. Retrieved 26 June 2015. [online] Dostupné z: <<https://www.businessinsider.com/the-enterprise-internet-of-things-market-2014-12>>

[8] "How IoT's are Changing the Fundamentals of "Retailing"". Trak.in – Indian Business of Tech, Mobile & Startups. 30 August 2016. Retrieved 2 June 2017. [online] Dostupné z: <<https://trak.in/tags/business/2016/08/30/internet-of-things-iot-changing-fundamentals-of-retailing/>>

[9] Recenze hlasových asistentů pro chytrou domácnost. [online] Dostupné z: <<https://compari.cz/hlasovi-asistenti-chytra-domacnost/>>

[10] Globalme. Top Alexa Skills and Voice Activated Apps to Look Out For. [online] Dostupné z: <<https://www.globalme.net/blog/top-8-voice-apps>>

[11] Meeker, M. (2016b). Mary Meeker Says Voice Search Is Going to Be Huge | Inc.com. [online] Dostupné z: <<https://www.inc.com/tess-townsend/mary-meeker-says-voice-search-is-going-to-be-huge.html>>

[12] iTranslate (2019). About iTranslate., [online] Dostupné z: <<https://www.itranslate.com/about>>

[13] The History and Growth of Voice AI and Search [online] Dostupné z: <<https://medium.com/cornertechandmarketing/the-history-and-growth-of-voice-ai-and-search-a0b7b4d37986> >

[14] "Nuance Announces Major New Releases of Dragon for Windows and Mac OS X". 2016. Retrieved 2016-08-22. [online] Dostupné z: <<https://www.nuance.com/index.html>>

[15] The History and Growth of Voice AI and Search [online] Dostupné z: <<https://medium.com/cornertechandmarketing/the-history-and-growth-of-voice-ai-and-search-a0b7b4d37986>>

[16] HOY, Matthew B. Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. 37. Medical Reference Services Quarterly. 2018, s. 81-88. ISSN 0276-3869. Dostupné z: doi:10.1080/02763869.2018.1404391

[17] Hirschberg, Julia, and Christopher D. Manning. "Advances in Natural Language Processing." *Science* 349 (6245): 261–266. Dostupné z: doi:10.1126/science.aaa8685.

[18] Moore, Clayton. "The Most Useful Skills for Google Home." *Digital Trends* (May 3, 2017). [online] Dostupné z: <<https://www.digitaltrends.com/home/google-home-most-useful-skills/>>

[19] Lee, Nicole. "Google Assistant on the iPhone Is Better than Siri, but Not Much." *Engadget* (May 17, 2017). [online] Dostupné z: <<https://www.engadget.com/2017/05/17/google-assistant-iphonhands-on/>>

[20] Hachman, Mark. "The Microsoft-Amazon Deal Leaves Cortana Speakers with One Advantage: Skype." *PCWorld*. [online] Dostupné z: <<https://www.pcworld.com/article/3221284/windows/the-microsoft-amazon-deal-leaves-cortana-speakers-with-one-advantageskype.html>>

[21] Spence, Ewan. "Windows Phone Is Dead, Long Live Microsoft's Smartphone Dream." *Forbes* (July 12, 2017). [online] Dostupné z: <<https://www.forbes.com/sites/ewanspence/2017/07/12/microsoftwindows-phone-windows10-mobile-strategy/>>

[22] Recenze hlasových asistentů pro chytrou domácnost. [online] Dostupné z: <<https://compari.cz/hlasovi-asistenti-chytra-domacnost/>>

[23] Day, M., Turner, G., & Drozdiak, N. (2019). Amazon Workers Are Listening to What You Tell Alexa. *Bloomberg.Com*. [online] Dostupné z: <<https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio>>

[24] Cho, E. (2019). Hey Google, Can I Ask You Something in Private? Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems – CHI '19, 1–9. [online] Dostupné z: <<https://doi.org/10.1145/3290605.3300488>>

[25] Tilley, Aaron. “How A Few Words to Apple’s Siri Unlocked a Man’s Front Door.” Forbes. [online] Dostupné z: <[https://www.forbes.com/sites/aarontilley/2016/09/21/applehomekit-siri-security/.](https://www.forbes.com/sites/aarontilley/2016/09/21/applehomekit-siri-security/)>

[26] Liptak, Andrew. “Amazon’s Alexa Started Ordering People Dollhouses after Hearing Its Name on TV.” The Verge , 2017 [online] Dostupné z: <<https://www.theverge.com/2017/1/7/14200210/amazon-alexa-tech-news-anchor-order-dollhouse.>>

[27] Zhang, Guoming, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu. “DolphinAttack: Inaudible Voice Commands.” In Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, New York, NY, 103–117. Dostupné z: doi:10.1145/3133956.3134052.

[28] Carlini, N., Pratyush M., Tavish V., Micah S., Shields C., Wagner D. & Wenchao Z. (2016). Hidden voice commands. USENIX Conference on Security [online] Dostupné z: <<https://www.usenix.org/conference/usenixsecurity16/technical-sessions/presentation/carlini>>

[29] Nuance (2018). Multimodal voice & behavioral biometric authentication technology. Dostupné z: <https://www.nuance.com/omni-channel-customer-engagement/security/identification-and-verification.html>

[30] HSBC (2019). HSBC UK REACHES DOUBLE LANDMARK IN FIGHT AGAINST FRAUD. Dostupné z: <https://www.about.hsbc.co.uk>

[31] PESTANES Philippe, GAUTIER Benjamin, Wavestone. 2017.: The rise of intelligent voice assistants: new gadget for your living room or window of opportunity to reshuffle the cards in the web economy? [online] 2017, 8

strán. Dostupné z: <<https://www.wavestone.com/app/uploads/2017/09/Assistants-vocaux-ang-02-.pdf>>

[32] RONSON, Jacqueline, Inverse innovation: Amazon Echo's Alexa's Now an Integral, Worrisome Part of the American Family

[online] 2016. Dostupné z: <<https://www.inverse.com/article/15555-amazon-echo-s-alexa-is-now-an-integral-worrisome-part-of-the-american-family>>

[33] Recenze hlasových asistentů pro chytrou domácnost. [online] Dostupné z: <<https://compari.cz/hlasovi-asistenti-chytra-domacnost/>>

[35] Společnost Apple: Apple HomePod [online] Dostupné z: <www.apple.com/homepod>

[36] Recenze hlasových asistentů pro chytrou domácnost. [online] Dostupné z: <<https://compari.cz/hlasovi-asistenti-chytra-domacnost/>>

[38] MOREN, Dan, Tom's Guide: The Best Smart Home Devices That Work with Google Home [online] 2017. Dostupné z: <<https://www.tomsguide.com/us/pictures-story/1075-google-home-compatible-devices.html#s1>>

[40] DOČEKAL, Daniel. Jak do Androidu dostat Cortanu od Microsoftu? A co to umožní? [online]. Dostupné z: 365tipu.wordpress.com

[41] ČÍŽEK, Jakub. Cortana míří do Francie a dalších zemí. Exotické jazyky mají smůlu. [online] Dostupné z: <<https://www.zive.cz/bleskovky/cortana-miri-do-francie-a-dalsich-zemi-exoticke-jazyky-maji-smulu/sc-4-a-176401/default.aspx>>

[42] Bixby : A New Way to Interact with Your Phone. [online] Dostupné z: <<https://news.samsung.com/global/bixby-a-new-way-to-interact-with-your-phone>>

[43] Three things we know Samsung's new assistant can do on the Galaxy S8. The Verge. [online] Dostupné z: <<https://www.theverge.com/2017/3/29/15097744/samsung-bixby-galaxy-s8-assistant-vs-siri-alexa-android>>

[44] Seznam přehraj lajnu, seznam.cz spouští hlasového asistenta ve stylu Siri [online] Dostupné z: <<https://www.lupa.cz/clanky/sezname-prehraj-lajnu-seznam-cz-spousti-hlasoveho-asistenta-ve-stylu-siri/>>

[45] Bohn, D. (2019). Google Assistant will soon be on a billion devices, and feature phones are next. [online] Dostupné z: <<https://www.theverge.com/2019/1/7/18169939/google-assistant-billion-devices-feature-phones-ces-2019>>

[46] Osterloh, R. (2019). Helping more people with wearables: Google to acquire Fitbit., [online] Dostupné z: <<https://blog.google/products/hardware/agreement-with-fitbit/>>

[47] Cisco Annual Internet Report (2018-2023) White Paper [online] Dostupné z: <<https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html>>

[48] Voice Assistant Application Market – Growth, Trends, Covid 19 Impact, and Forecasts (2021-2026) [online] Dostupné z: <<https://www.mordorintelligence.com/industry-reports/voice-assistant-application-market>>

[49] Společnost Nuance. [online] Dostupné z: <<https://www.nuance.com/index.html>>

[51] Wolters, Maria Klara, Fiona Kelly, and Jonathan Kilgour. “Designing a Spoken Dialogue Interface to an Intelligent Cognitive Assistant for People with Dementia.” *Health Informatics Journal* 22 no. 4 (12/1, 2016): 854–866. Dostupné z: doi:10.1177/1460458215593329.

[52] Palladino, Valentina. “Google Pixel Buds Are Wireless Earbuds That Translate

Conversations in Real Time.” *Ars Technica* (10/4, 2017). [online] Dostupné z: <<https://arstechnica.com/gadgets/2017/10/google-pixel-buds-are-wireless-earbuds-that-translate-conversations-inreal-time/>>

[53] Patil, Sumant, and Patrick Davies. “Use of Google Translate in Medical Communication: Evaluation of Accuracy.” *BMJ* 349 (12/15, 2014): g7392. Dostupné z: doi:10.1136/bmj.g7392.