



BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF FOREIGN LANGUAGES

ÚSTAV JAZYKŮ

ARTIFICIAL INTELLIGENCE IN INTERNET COMMUNICATION

UMĚLÁ INTELIGENCE V INTERNETOVÉ KOMUNIKACI

BACHELOR'S THESIS

BAKALÁŘSKÁ PRÁCE

AUTHOR

AUTOR PRÁCE

Bc. Evgeniy Merzliakov

SUPERVISOR

VEDOUČÍ PRÁCE

Mgr. Jana Jašková, Ph.D.

BRNO 2020



VYSOKÉ UČENÍ FAKULTA ELEKTROTECHNIKY
TECHNICKÉ A KOMUNIKAČNÍCH
V BRNĚ TECHNOLOGIÍ

Bakalářská práce

bakalářský studijní obor Angličtina v elektrotechnice a informatice

Ústav jazyků

Student: Bc. Evgeniy Merzliakov

ID: 145848

Ročník: 3

Akademický rok: 2019/20

NÁZEV TÉMATU:

Umělá inteligence v internetové komunikaci

POKYNY PRO VYPRACOVÁNÍ:

Cílem práce je popsat zařízení, která jsou schopná samostatně komunikovat s lidmi prostřednictvím internetu.

DOPORUČENÁ LITERATURA:

Russell, S. J., & Norvig, P. (2014). Artificial intelligence: a modern approach (3rd ed). Harlow: Pearson Education.

Poole, D. L., & Mackworth, A. K. (2010). Artificial intelligence: foundations of computational agents. Cambridge: Cambridge University Press.

Song, J. (Ed.). (2018). The impact of artificial intelligence on communication networks and services (monothematic edition). ITU Journal: ICT Discoveries, 1(1). Available from https://www.itu.int/dms_pub/itu-t/obj/tut/T-TUT-ITUJOURNAL-2018-P1-PDF-E.pdf

Termín zadání: 6.2.2020

Termín odevzdání: 12.6.2020

Vedoucí práce: Mgr. Jana Jašková, Ph.D.

doc. PhDr. Milena Krhutová, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do dílech autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

Fakulta elektrotechniky a komunikačních technologií, Vysoké učení technické v Brně / Technická 3058/10 / 616 00 / Bmo

Abstrakt:

Rozvoj takových technologií jako cloud computing, big data a deep learning způsobuje to, že se umělá inteligence pomalu stává součástí běžného života. Tato bakalářská práce popisuje základní systémy s umělou inteligencí, které už mohou být používány i v dnešní době, například Smart Home, autonomní vozidla a virtuální asistenti. Zároveň je zde popsána krátká historie podobných systémů a jejich vliv na komunikaci a Internet. Dále jsou zde popsány základní principy umělých neuronových sítí. V dané práci jsou porovnány tři různé chatboty (Mitsuku, Uberbot, Cleverbot). Dvě metody jsou pak použity pro porovnání: zjednodušený Turingův Test a lingvistická analýza. Z výsledků vyplývá, že Cleverbot je lepší než Mitsuku a Uberbot.

Klíčová slova: Virtuální asistent; Smart Home; Mitsuku; Uberbot; Cleverbot; Chatboty; Umělá inteligence.

Abstract:

With the development of such technologies as cloud computing, big data and deep learning, the artificial intelligence (or AI) is becoming a part of everyday life, and its capabilities are rapidly evolving. This bachelor thesis defines the basic AI systems, which can be found and used today, such as self-driving vehicles, Smart Home systems, intelligent virtual assistants. Brief history, pros and cons as well as the impact of such systems on communication and Internet are also described there. Furthermore, basic principle of artificial neural networks is also explained in this thesis. Three different chatbots (Mitsuku, Uberbot, Cleverbot) are tested and compared. Two different methods are used for this purpose: simplified Turing Test and linguistic analysis. Results show that Cleverbot is slightly better than Mitsuku and Uberbot.

Keywords: Intelligent virtual assistant; Smart Home; Chatbots; Mitsuku; Uberbot; Cleverbot; Artificial intelligence.

MERZLIAKOV, Evgeniy. *Umělá inteligence v internetové komunikaci* [online]. Brno, 2020 [cit. 2020-06-12]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/127141>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav jazyků. Vedoucí práce Jana Jašková.

Prohlášení

Prohlašuji, že svou bakalářskou práci na téma *Umělá inteligence v internetové komunikaci* jsem vypracoval samostatně pod vedením vedoucího bakalářskou práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č.121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne: 12.06.2020

.....

podpis autora

Table of Contents

List of Figures	1
List of Tables.....	2
Introduction	3
1. Artificial Intelligence in Everyday Life	4
1.1. Self-driving Vehicles	4
1.2. Internet of Things.....	7
1.3. Smart Home	9
2. AI and the Internet Communication.....	11
2.1. Market Research and Advertisement	12
2.2. Intelligent Virtual Assistants.....	13
2.3. Chatbots	17
3. Comparison of chatbots.....	18
3.1. Turing Test.....	19
3.2. Linguistic analysis of chatbots	23
3.2.1. Evaluation of Mitsuku chatbot.....	23
3.2.1.1. Lexical level	24
3.2.1.2. Sentence level	26
3.2.1.3. Text level.....	28
3.2.2. Evaluation of Uberbot.....	30
3.2.2.1. Lexical level	30
3.2.2.2. Sentence level.....	32
3.2.2.3. Text level.....	34
3.2.3. Evaluation of Cleverbot.....	35
3.2.3.1. Lexical level	35
3.2.3.2. Sentence level.....	37
3.2.3.3. Text level.....	40
3.3. Results.....	41
Conclusion.....	42
Rozšířený český abstrakt	43
References	46

List of Figures

<i>Figure 1. The six Levels of Driving Automation (Synopsys, 2019)</i>	6
<i>Figure 2. Sensor and Control System of Self-driving Car (Augur, 2015)</i>	7
<i>Figure 3. Operation Principle of the IoT System (Rouse, 2019)</i>	8
<i>Figure 4. Smart Home System (Consumer Reports, 2014)</i>	10
<i>Figure 5. Neural Network Operation (Prince, 2017)</i>	12
<i>Figure 6. Percentage of Correctly Answered Questions(Munster and Thompson,2019)</i> ..	16
<i>Figure 7. Results of the Query (Munster and Thompson, 2019)</i>	16
<i>Figure 8. First phase of the Turing test (Minini, 2020)</i>	19
<i>Figure 9. Second phase of the Turing test (Minini, 2020)</i>	20
<i>Figure 10. Simplified Turing Test (Minini, 2020)</i>	20

List of Tables

<i>Tab.1. Simplified Turing Test results for Mitsuku</i>	21
<i>Tab.2. Simplified Turing Test results for Uberbot.</i>	22
<i>Tab.3. Simplified Turing Test results for Cleverbot</i>	23
<i>Tab.4. Examined linguistic features.</i>	23
<i>Tab.5. Results of linguistic analysis.</i>	41

Introduction

With the development of such technologies as cloud computing, big data and deep learning, the artificial intelligence (or AI) is becoming a part of everyday life, and its capabilities are rapidly evolving. AI can be used in many ways: it can perform hard, repeatable and dangerous work; it may help us to handle disasters; it can make our daily life more comfortable. Systems with AI control can manage complex tasks such as monitoring credit card systems for fraudulent behaviour, detecting cybersecurity threats, supporting medical diagnoses. Moreover, automated (self-driving) vehicles are becoming more common in America. Intelligent voice assistants (or just voice assistants) are used to improve quality of live (for example, with one of such assistants it is possible to order a delivery by using only voice commands). Chatbots are widely used by many companies to provide the customer support and to answer the trivial questions. Several chatbots are close to passing the Turing Test, which means that they may be called “intelligent”. In the near future AI systems will be present in the form of medical, service and military robots (nevertheless, military robots are tested today, service robots are used in Japan). Application of such systems may help in future expeditions to Mars.

The first part of this thesis describes the AI systems, which people can find and use today (self-driving vehicles, Smart Home systems, Internet of Things). Impact of such systems is visible even today. For example, Smart Home systems and parts of them are present in almost every household (automated lighting, automated heating systems).

The second part of the thesis explains the impact of AI on internet communication, which includes the Internet of Things, chatbots and AI system for market research and advertisement. Firstly, the basic principles of artificial neural networks which are used in such structures, pros and cons as well as the benefits of such systems. Then the basic principles and examples of intelligent virtual assistants will be explained. Chatbots will be discussed at the end of this chapter.

The final part of the thesis is devoted to comparison of different chatbots such as Mitsuku, Uberbot, Cleverbot. Chatbots will be tested by using two methods: Turing Test and linguistic analysis. Then, their communication capabilities will be compared. Results will be discussed and summarised at the end of this chapter.

1. Artificial Intelligence in Everyday Life

This part of the paper is devoted to the description of the AI systems that have impact on the everyday life and which are used or tested today. It includes self-driving cars, smart home systems and IoT (Internet of Things) systems. Moreover, IoT devices are used in Smart Home systems. For example: robot vacuums, door locks, smart light switches. Self-driving vehicle technology is rapidly evolving and is currently tested in some states of the USA.

1.1. Self-driving Vehicles

According to Nguyen (2019), the history of self-driving vehicles began in the 1930s, with an exhibit at the 1939 New York World's Fair, which was called "Futurama". It was a concept of an automated highway system in which cars moved autonomously, created by Norman Bel Geddes. In 1958 General Motors began working on this concept, but this project was cancelled because of the lack of funds.

A team at Tsukuba Mechanical Engineering Laboratory took the first major step in developing stand-alone autonomous vehicle in 1977. It was guided with the help of machine vision wherein a computer analyses the surrounding environment using imagery from built-in cameras. The prototype was capable to reach the speed in 20 miles per hour and was programmed to follow white street markers (Nguyen, 2019).

Ernst Dickmanns, backed by Mercedes-Benz, created and tested a prototype of self-driving vehicle that could drive autonomously at high speed in 1986. This led to the investment of 749 000 000 € made by European research organisation EUREKA, which helped Dickmanns and researchers to make several key advances in camera technology, computer processing and software that culminated in two robot vehicles, the VaMP and VITA-2. At the start of the 21st century, the U.S. military announced the DARPA Grand Challenge, a long-distance competition in which \$1 million would be awarded to the team of engineers whose vehicle conquers the 150-mile obstacle course. None of the vehicles finished the course, however the event was considered a success as it helped to raise the interest in the field. In 2010, internet giant Google announced their self-driving car, which was made in the hopes of finding a solution that would reduce the number of car accidents each year by half. The project was headed by Sebastian Thrun, director of Stanford's Artificial Intelligence Laboratory (Nguyen, 2019).

By 2015, Google cars had logged more than 1 million kilometres without causing an accident, despite that they were involved in 13 collisions. Over the course of this project, the company made several huge steps. They lobbied for and got legislation passed to make self-driving cars street legal in four states and the District of Columbia, revealed an autonomous model it plans to release in 2020 and is continually building and opening up testing sites throughout the country under a project named Waymo. Progress of their project attracted several other big companies, such as Microsoft, Tesla, Uber as well as traditional car manufacturers include Toyota, BMW, Audi. Unfortunately, progress on advancing the technology was slowed because of the accident with Uber self-driving vehicle: the car hit and killed pedestrian (Nguyen, 2019).

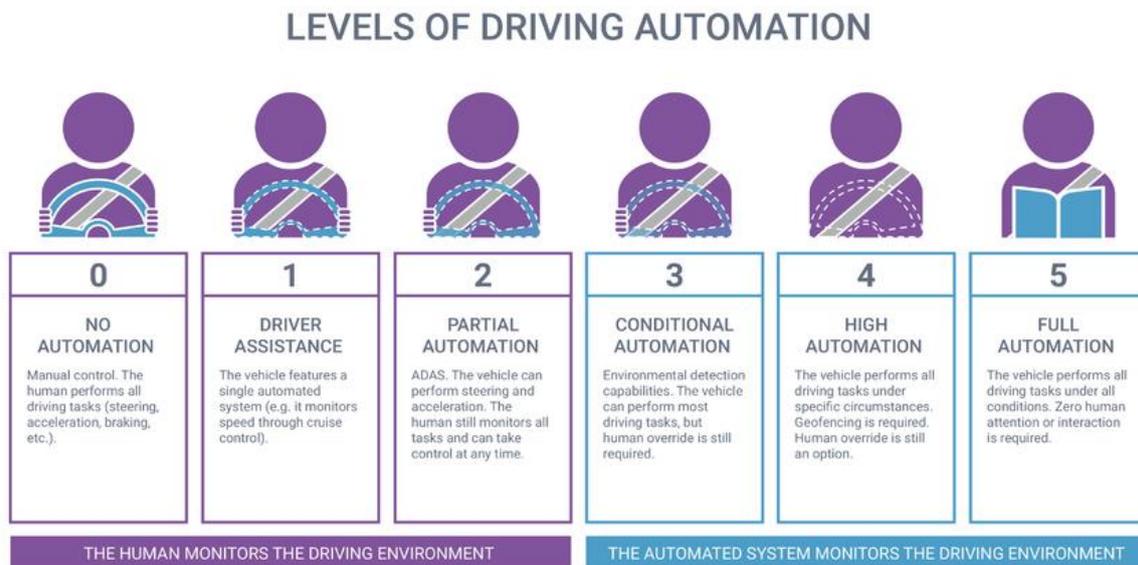


Figure 1. The six Levels of Driving Automation (Synopsys, 2019)

Self-driving vehicle can sense its environment and can operate without human involvement. A human passenger is not required to take control of the vehicle at any time, nor is a human passenger required to be present in the vehicle at all. According to the SAE (Society of Automotive Engineers) classification, which is displayed in the figure 1., self-driving cars belong to the level 3 (conditional driving automation) or level 4 (high driving automation). It means that a human passenger must always be present and ready to take control. Fully autonomous cars are tested in several regions; however, they are not available to the general public (Synopsys, 2019).

Self-driving vehicle is a complex structure that relies on sensors, actuators, machine learning systems and powerful central computer to execute software. These cars can create and maintain a map of their surroundings. This process is based on variety of sensors situated in different parts of the vehicle: radar sensors monitor the position of nearby objects; lidar (light detection and ranging) sensors send pulses of light to the car's surroundings to measure distances, identify lane markings and detect road edges; video cameras detect traffic lights, track other vehicles and look for pedestrians; ultrasonic sensors in the wheels detect kerbs and other vehicles when parking. Data from the sensors is then processed by sophisticated software, which also plots a path and sends instructions to the

actuators, which control acceleration, braking and steering (Synopsys, 2019). Positioning of the sensors is shown in figure 2.

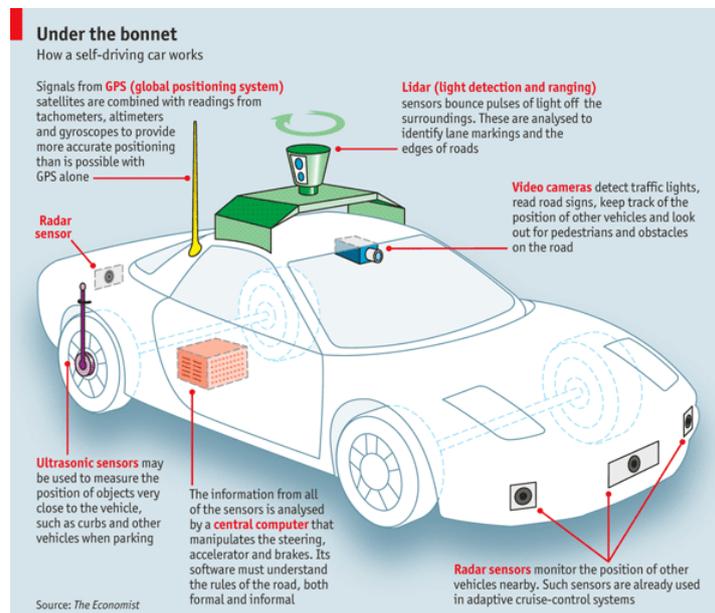


Figure 2. Sensor and Control System of Self-driving Car (Augur, 2015)

Any innovation has certain benefits and drawbacks, including self-driving car technology. However, according to the report produced by the University of California, Davis, and the Institute for Transportation and Development Policy, the main benefit of autonomous car is the potential for lowering CO₂ emissions (Institute for Transportation & Development Policy, 2017).

Any automation, including automation of the vehicle, comes at a price of vehicle safety. Complexity of the control system means that there is always a chance of glitches. Moreover, these vehicles will have to be defended against cyberattacks (The Pollack PR Marketing Group, 2018).

1.2. Internet of Things

The Internet of Things (IoT) is a system of interconnected devices, digital and mechanical machines, animals, objects or people that have unique identifiers and can transfer data over a network without requiring human-to-computer or human-to-human interaction. As a “thing” in the IoT system can be considered any man-made or natural object - a car with built-in sensors, a person with a heart monitor implant, an animal with a biochip. IP address

can be assigned to any of such objects, which means that they can transfer a data over a network (Rouse, 2019).

The Internet of Things evolved from machine-to-machine communication, which refers to connecting a device to the cloud, collecting data and managing without human interaction. It is also an extension of SCADA (supervisory control and data acquisition), systems for process control, which gather data in real time in order to control the equipment and conditions. The last generation of SCADA developed into the first generation of the IoT systems (Rouse, 2019).

A system of the Internet of Things consists of smart devices that use sensors, embedded processors and communication hardware to acquire, send and process the collected data. Acquired information is then shared between IoT devices by connecting to a gateway or PC, where the data are sent to the cloud for analysis or are analysed locally. These devices can communicate and share the data with each other in order to act according to the acquired information. They can operate without human interaction; however, the user can give them instructions, access the data and configurate them (Rouse, 2019). Operation principle of such system is shown in figure 3.

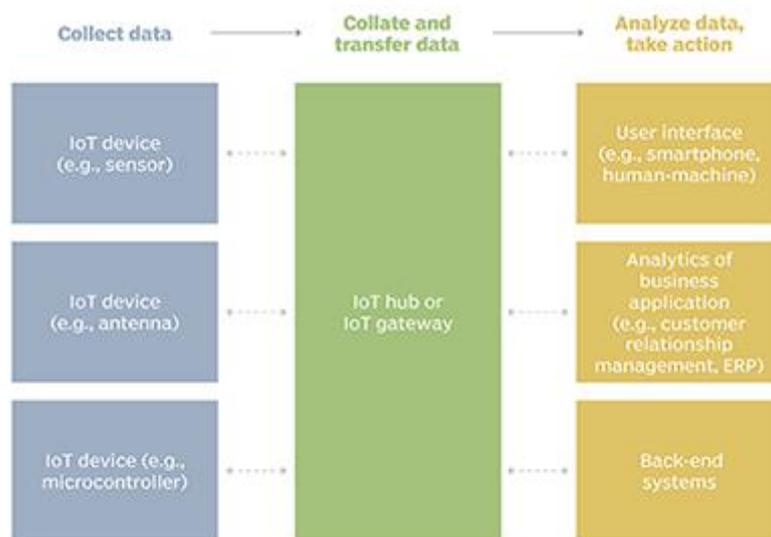


Figure 3. Operation Principle of the IoT System (Rouse, 2019)

According to Rouse (2019), the Internet of Things brings the following benefits to companies:

- Saving time and money;
- Monitoring of the business processes;
- Enhanced productivity of employees;
- Improved customer experience;
- Possibility to make better business decisions.

IoT gives companies the tools to improve their business strategies and helps them to change the ways they approach the customers. Common users also have several benefits from using the Internet of Things, such as the ability to access information from anywhere on any device at any time as well as improved communication between connected devices.

However, the IoT devices have several disadvantages, including the possibility of stealing the confidential information because of the connection between devices as well as the absence of the international standard, which leads to the communication errors between devices from different manufacturers (Rouse, 2019). Internet of Things devices are widely used in Smart Home technology, that will be described next.

1.3. Smart Home

Smart home technology, also known as domotics or home automation, provides security, comfort and energy efficiency to homeowners by allowing them to control smart devices with help of smart home application or intelligent virtual assistant. As a part of the internet of things (IoT), smart home devices often operate together, sharing consumer usage data among themselves and automating actions based on the preferences of homeowner (Rouse, 2018).

The main step in the history of home automation was X10, a communication protocol which was released in 1975. This protocol could send bursts of the information onto electric wiring of the house to programmable outlets or switches. These signals contain commands to corresponding devices, controlling their operation. However, electrical wiring is not designed to be free from radio-band noise, which means that X10 was not reliable. Moreover, the first version of this protocol was a one-way technology, which meant that smart devices could not send data back to a central network. Two-way X10 devices became

available later, at a higher cost. New technology that combined electric wiring with wireless signals was introduced in 2005 by home automation company Insteon. Other protocols, such as Zigbee and Z-Wave, have also appeared to counter the problems of X10, however this protocol still remains the most used to this day. Modern smart home solutions are released by such companies as Amazon, Apple and Google (Rouse, 2018).

According to Rouse (2018), nearly every device that is used in the domestic space (figure 4.) has seen the introduction of a smart home alternative:

- Smart TVs connect to the internet to access content through applications, such as on-demand video and music.
- Smart lighting systems that can detect when occupants are in the room and adjust lighting as needed.
- Smart thermostats which come with integrated Wi-Fi, allowing users to schedule, monitor and remotely control home temperatures.
- Smart locks and garage-door openers, which can grant or deny access to visitors. These locks can also detect when residents are near and unlock the doors for them.
- Smart security cameras, which allow residents to monitor their home when they are on vacation.

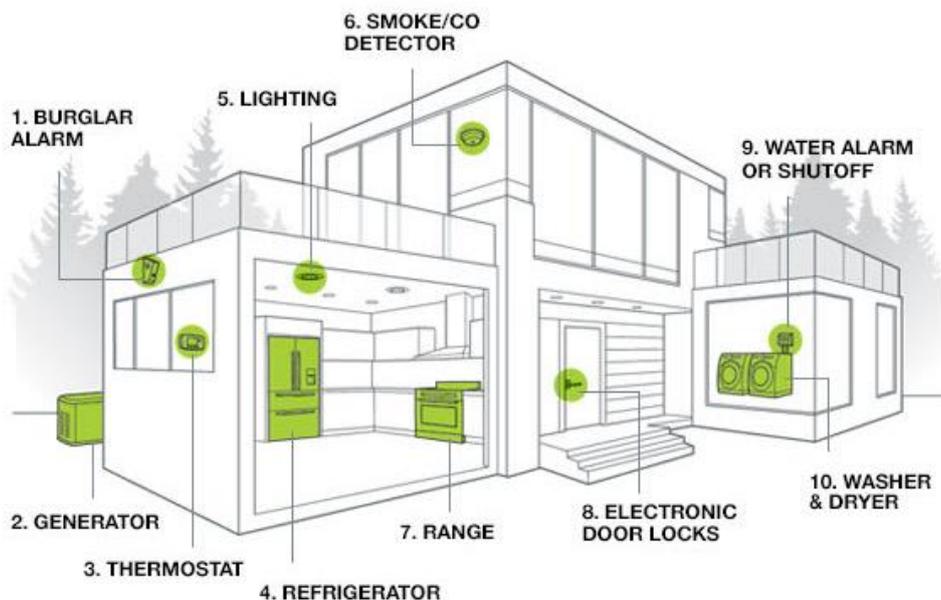


Figure 4. Smart Home System (Consumer Reports, 2014)

One of the significant benefits of home automation is allowing homeowners to monitor their homes remotely, countering dangers such as a forgotten coffee maker left on or a front door left unlocked. Domotics is also beneficial for the elderly, providing monitoring that can help seniors to remain at home comfortably and safely, without requiring 24/7 home care. Home automation also helps consumers to improve efficiency. Instead of leaving the air conditioning on all day, a smart home system can learn your behaviour and make sure the house is cooled down by the time you arrive back home. With home automation, energy, water and other resources are used more efficiently, which helps save both natural resources and money (Rouse, 2018).

Nevertheless, home automation systems have struggled to become mainstream, because of their technical nature. A drawback of smart homes is their complexity; some people have difficulty with technology or can stop using it with the first annoyance. Smart home manufacturers and alliances are working on reducing complexity and improving the user experience to make it enjoyable and beneficial for all users. The other drawback of such systems is security. Hacker can access badly secured devices, such as cameras, routers and DVR's. In addition to security, many smart home opponents worry about data privacy. While smart home device and platform manufacturers may collect consumer data to offer new services to customers, transparency between manufacturers and users should be improved (Rouse, 2018).

2. AI and the Internet Communication

Development of natural language processing and neural network programming opens new possibilities in the field of communication. The use of these disciplines improves speech synthesis as well as converting speech to text possibilities of devices, which helps to cope with different languages and accents. It means that telecommunication applications such as Skype can translate both speech and text input of the user to different languages, which makes communication between people who speak different languages clearer and more productive. Another improvement relates to the intelligent voice assistants and chatbots – the use of neural networks and natural language processing makes responses of such entities more natural and comprehensive (Draffan and Wald, 2019). The next subchapters describe their capabilities and some of possible applications.

2.1. Market Research and Advertisement

Artificial intelligence in marketing is an entity that can be found in almost every aspect of this industry. Most of the current AI-based applications in predictive analysis and marketing are based on neural networks (Brenner, 2018).

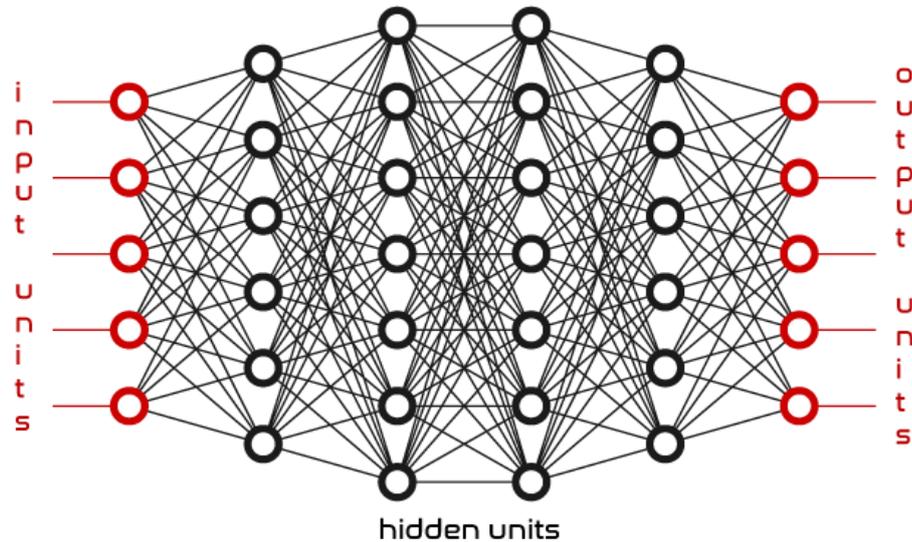


Figure 5. Neural Network Operation (Prince, 2017)

According to Prince (2017), artificial neural networks are designed on biological systems, which have made the concept of self-learning AI a reality. Neural networks function similar to human brain, which means that they are progressively learning and improve over time. The network is made of high number of neurons, which are called units and are distributed on three interconnected layers (figure 7.):

- Input units, which receive information from external source in order to process or learn about;
- Output units, which respond to the processed information by the network;
- Hidden units, which form the bulk of the network that processes the given tasks.

Neural network can learn by using one of two following ways: Supervised and Unsupervised. The first method is based on giving to the network both input and expected output variables. Using back-propagation algorithm (comparing given outputs to outputs and then correcting the input) they achieve expected level of precision. When the next set of data is fed, they will produce the accurate outcome. The second method is based on giving

to the network only input variables – dependences and associations are identified by the network itself (Prince, 2017).

One of the main applications of artificial neural networks in marketing is predictive analytics. It is a convergence of data mining and predictive modelling, which are augmented by the machine learning capabilities of neural networks. Networks can make predictions on future outcomes within a campaign by learning to define the past and current trends. Microsoft's BrainMaker is one of the brightest examples of the use of neural networks for predictive analytics. It improved the efficiency of the mailing campaign, increasing the response rate from 4.9% to 8.2% by analysis of the quantity, type of purchased products and the elapsed time between first and last purchase (Prince, 2017).

Another application of neural networks is market segmentation and micro-targeting. They are used to divide the audience into distinct groups based on specific characteristics associated with them, such as demographics, social and economic status, their attitude and location. Unsupervised neural networks are used to segment the location and demographic, while supervised networks are used to determine buying behaviour of the customers (Prince, 2017).

2.2. Intelligent Virtual Assistants

An intelligent virtual assistant or voice assistant (VA) is an artificial intelligence system that emulates human interaction to perform various tasks, like customer support. All modern solutions (e.g. Apple's Siri, Microsoft's Cortana, Amazon's Alexa and Google's Assistant) are software agents that run on smartphones or specifically designed speaker devices. This software constantly listens to the surroundings for a key word. Once that key word is heard by VA, it records the user's voice. This information is then sent to specialized server, which processes it and interprets as a command. Based on the command, the voice assistant will receive the appropriate information from the server to read back to the user and play the requested media or complete tasks with connected devices and services. The number of such

services is growing rapidly, as the Internet-of-Things manufacturers are building voice control into their products (Hoy, 2018).

The first virtual assistant (Apple's Siri) was released as a standalone application in 2010 and was integrated into iOS in 2011. It was shortly followed by Microsoft with Cortana in 2013. Amazon's Alexa was launched in 2014 with its Echo-connected home speaker, and Google's Assistant was announced in 2016. Each assistant has unique features, but each of them has the same basic functions. Voice assistants are more advanced than earlier voice-activated technologies – they are always connected to the Internet. This connection allows the assistant to analyse the voice commands and provides VA with the proper response. Earlier versions of voice-activated devices relied on limited set of commands and responses. Advances in natural language processing has allowed voice assistants to respond in quick and correct manner. Growing power of personal computers and with advancing cloud technology allows scientists and software engineers to train voice assistants, which means that the assistants can respond in more natural way. Additional computing power also means that the VA can parse requests phrased in different ways and interpret them correctly (Hoy, 2018).

According to Hoy (2018), intelligent virtual assistants are able to perform such basic tasks as:

- Doing basic math calculations;
- Setting timers, alarms, reminders;
- Telling jokes and stories;
- Sending and reading text messages and emails as well as making phone calls;
- Playing media from connected services such as iTunes, Netflix, Amazon and Spotify;
- Controlling Internet-of-Things devices such as lights, alarms, locks and thermostats.

Moreover, voice assistants can be improved with additional features called “skills”, which expand their abilities. These skills are built by third-party developers, identical to the way applications are developed for smartphones.

Voice assistants are available on a wide variety of different platforms, including smartphones, PC and home speaker devices. Amazon and Google are the main manufacturers of home speaker devices for their intelligent virtual assistants. Amazon has

several variations of its Echo product: from compact Echo Dot to the premium Echo Snow devices, which can stream music and make video calls. Microsoft has focused on building Cortana into Windows 10, while Apple has kept their assistant away from non-iOS devices (Hoy, 2018).

Despite their interesting and useful features, voice assistants also have several problems. One of them is security – anyone with access to a voice-activated device can gather information about the accounts and services used on this device, as well as ask it to perform tasks. This problem is currently solved by including voice printing into software, which uniquely identifies each user by voice and prevents the device from leaking the personal information. Privacy is another major concern for users – voice assistants must be constantly listening in order to respond to commands. Developers insist that their devices are not recording any information unless it is activated by command, but malfunctioning device can be constantly recording and sending those recordings back to server. It means that companies developing such devices can gather the information about users and there is a possibility for data to be stolen or used to incriminate people (Hoy, 2018).

Voice assistants have become popular in recent years – they provide the possibility of human-computer interaction through intuitive actions. It means that number of users grows rapidly, as well as the amount of information that should be analysed and properly interpreted by the voice assistants. With the development of machine learning and natural language processing, abilities of the assistants to correctly communicate rapidly improve (Berdasco et al., 2019).

Nevertheless, according to Munster and Thompson (2019), none of the compared intelligent virtual assistants can answer and understand 100% of the asked questions. Nevertheless, some of them are more capable than others. The most capable assistant was determined by asking the same 800 questions, which were graded on two metrics: correctness of the response and ability to understand the asked questions. This set of questions consisted of 5 categories: Local, Commerce, Navigation, Information and Command.

Questions Answered Correctly by Category

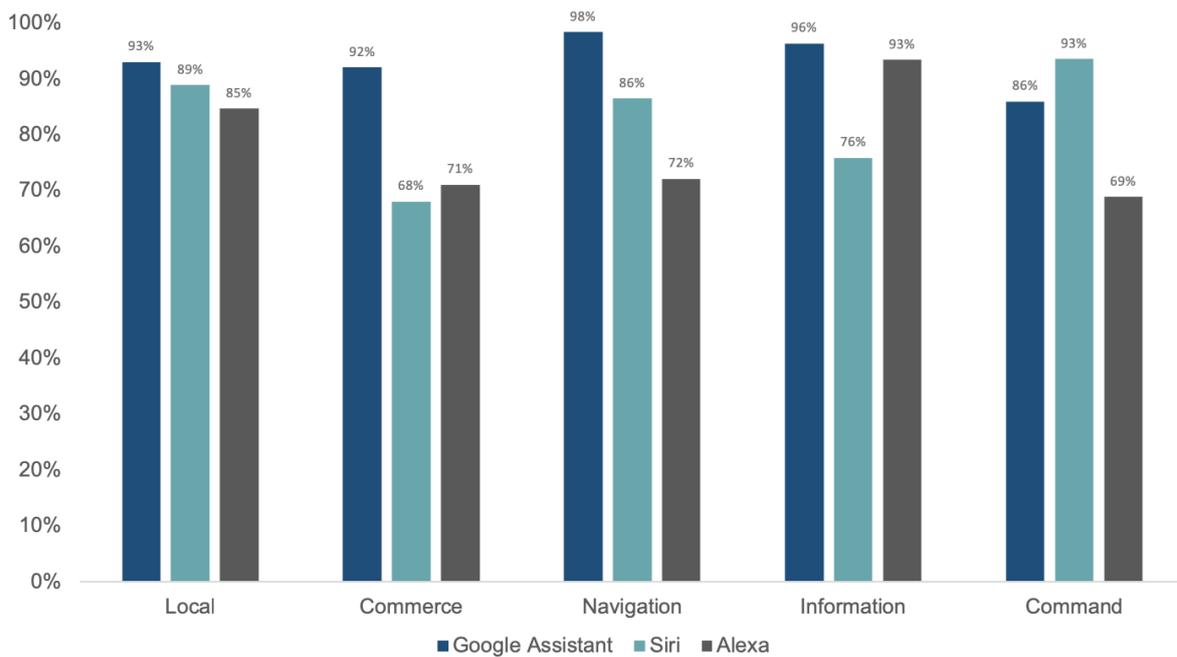


Figure 6. Percentage of Correctly Answered Questions (Munster and Thompson, 2019)

According to the data shown in the figure 6, Google Assistant has better results overall, but lacks in the Command category. Siri is the leader in this category, which means that this voice assistant is more useful for such functions like calling, emailing, playing multimedia and texting. Amazon Alexa follows close to Google Assistant in such categories as Information while it performs worse overall. Query results are shown below (figure 7.) and according to them Google Assistant is currently the best performer (Munster and Thompson, 2019).

Query Results

	Answered Correctly	Understood Query
Google Assistant	92.9%	100%
Siri	83.1%	99.8%
Alexa	79.8%	99.9%

Figure 7. Results of the Query (Munster and Thompson, 2019)

Reliability of the intelligent voice assistant is one of the most important characteristics. Need to repeat or to rephrase commands or questions makes experience of the user worse. According to Moren (2019), all three compared assistants have good voice recognition, with occasional requirement to rephrase or repeat queries.

On the other hand, multiple users can be in household and interact with voice assistant, which can lead to unpredictable results. Each intelligent assistant has its own way of handling this situation. For example, Google Assistant and Amazon Alexa offer multiple voice profiles, which allows user to train the assistant to recognise his voice specifically (Moren, 2019).

The most advanced of these ways is called Voice Match and it is used by Google Assistant. It allows the user to set his own default media devices, as well as to personalise payments, calendars and photos. On the other hand, this technology can be improved further – some users encounter problems with voice recognition, which means that they cannot access the information until voice assistant is retrained. Amazon Alexa has similar technology, but it is limited to personalized shopping, calling and some media options (Moren, 2019).

Apple Siri uses a different approach – voice assistant should be trained to recognise user’s voice before activation of the wake phrase on an iOS device. Such method prevents the assistant from leaking the user’s personal information to someone else and allows to avoid simultaneous response on several devices (e.g. on iPad and iPhone). On the other hand, Siri tends to replace this feature, responding to any wake phrase, if it operates on the HomePod. Furthermore, it does not support the input from several different users, which means that it has the worst voice recognition capabilities among compared intelligent virtual assistants (Moren, 2019).

2.3. Chatbots

Chatbot is an artificial intelligence system that simulates a conversation with a user though text or voice interface (messaging applications, telephone conversation). Those interactions can be complex (e.g. problem with the internet browser) or straightforward such as asking a bot for the latest news. The idea behind chatbots is that it will be the way to easy find necessary information without the need to do a Web search or use multiple applications (Maruti TechLabs, 2019).

According to the Oracle article (2019), two main types of chatbots exist:

- Task-oriented (declatative) chatbots
- Data-driven and predictive chatbots

Task-oriented chatbots are single-purpose programs, which means that they perform one function – they generate automated but conversational responses to user input using rules, natural language processing and machine learning. They can handle common questions, but their capabilities are limited. This type of chatbots is currently the most commonly used (Oracle, 2019).

Data-driven chatbots are often referred as virtual assistants and they are more complex, personalised and interactive than task-oriented chatbots. They apply predictive intelligence and natural language understanding to enable personalisation according to past user behaviour and profiles. They can learn the preferences of user and provide recommendations according to them (Oracle, 2019).

According to the Expert System article (2018), operation principle of chatbot is based on two different tasks – user request analysis and returning the response. The first task is performed in order to identify the intent of the user and to correctly understand the given request. Returning the response means that chatbot provides the most appropriate answer for the user's request. It can be:

- Predefined text;
- Data stored in enterprise systems;
- Contextualised piece of information based on the provided data;
- Text retrieved from database that contains different answers;
- Result of chatbot's interaction with one or several backend applications.

Chatbots bring several advantages to the companies. One of them is increasing the number of served customers to potentially unlimited number. Each of these customers is engaged in a personal way and is provided with personalised and proactive service. It also means that traditional call centers have less pressure and can concentrate on more important tasks. According to this information, chatbots are commonly used to perform routine customer activities within the banking, beverage and retail sectors (Oracle, 2019).

3. Comparison of chatbots

This chapter is devoted to comparison of different chatbots, particularly Mitsuku, Uberbot, Cleverbot. These chatbots were chosen to test because they were close to passing the Turing Test. Following methods were used for the evaluation:

- The Turing Test
- Linguistic analysis

3.1. Turing Test

The Turing Test was conceived by Alan Turing in 1950. The main purpose of the test is to determine intelligence of the machine or program. Original version of the Turing Test was divided into two different phases:

First phase had three participants – a man, a woman and an interviewer (figure 8), which were in separate rooms and can communicate only with help of the teletype. The interviewer sent written question to the participants. The goal of this test was to determine which participant was a man and which one was a woman. One of the participants could lie (Minini, 2020).

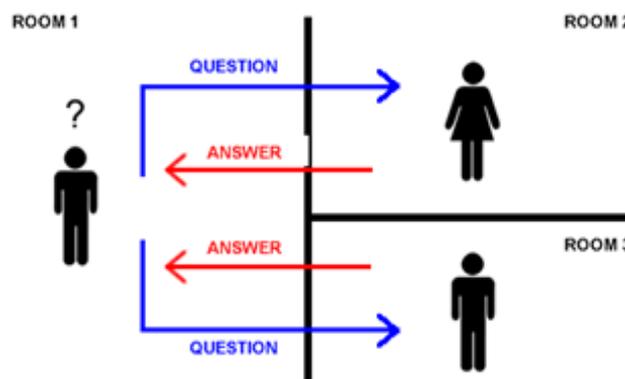


Figure 8. First phase of the Turing test (Minini, 2020)

Second phase also had three participants, but one of them were replaced with computer (figure 9.). The main goal of interviewer was to determine which one of participants was a machine. Process was similar to the first phase (Minini, 2020).

The Turing Test was considered successful and machine can be called intelligent if the error rate is equal in both phases.

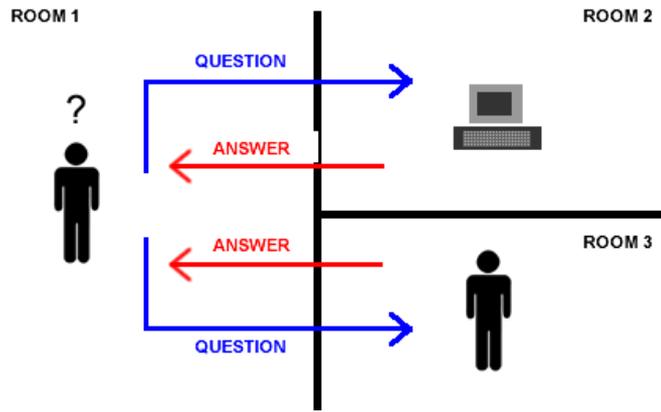


Figure 9. Second phase of the Turing test (Minini, 2020)

For evaluation of the chatbots was used simplified Turing Test. This variation of Turing Test can be described in the following way:

Human interviewer communicates with another subject placed in another room (or virtually) in this version of Turing Test. The subject can be a person or a computer (Minini, 2020). The machine must convince the interviewer that it is a human. Difference between the standard and simplified versions is in the number of participants (figure 8.).

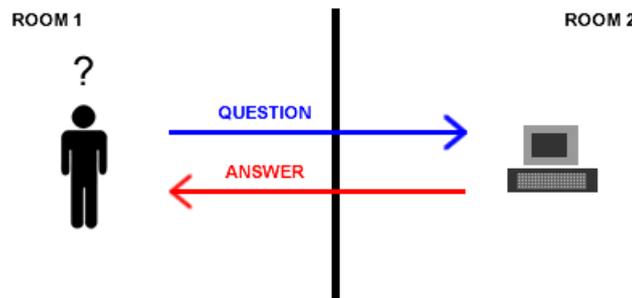


Figure 10. Simplified Turing Test (Minini, 2020)

Following questions and requests were used for examination:

- It is a good weather today, isn't it?
- What is the square root of three?
- Subtract 25569 from 107987
- What day is today?
- Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?

Reaction time, emotionality and naturality of the answers will be evaluated.

3.1.1. Evaluation of Mitsuku

Mitsuku chatbot was tested first. According to the answers below (Tab.1), Mitsuku behaves quite natural, however some answers were not correct. One of such questions was the following: “ Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?”. It is a reference to Douglas Adams and Mitsuku could answer this question. Expected answers were “42” or “I don’t know”. Nonetheless, Mitsuku recognised different formulation: “What is 42?”. Another example was question “The weather is good today, isn’t it?”. Expected answer was simple “Yes, it is” or “No, it is not” or any information connected with asked topic. However, the answer was: “Where would you rather be?”. Apart of these two answers, the behaviour of chatbot was quite natural, with good reaction time.

Question	Mitsuku
The weather is good today, isn’t it?	Where would you rather be?
What is the square root of three?	Some irrational number.
Subtract 25569 from 107987	Sorry, but I can only work with whole numbers from 0 to 9,999
What day is today?	Wednesday
Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?	How about asking to write a poem on a random subject? I bet a computer will struggle with that.

Tab.1. Simplified Turing Test results for Mitsuku

3.1.2. Evaluation of Uberbot

The second tested subject was Uberbot. According to the table below (Tab.2.), Uberbot answers were natural, however the answer time of some answers was lower than expected. For example, Uberbot answered the question “Subtract 25569 from 107987?” instantly, while the real person would require time for calculation. Another example was question “ Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?”.

It is a reference to Douglas Adams and expected answers were “42” or “I don’t know”. Similar to Mitsuku, Uberbot did not understand the reference using this formulation. Nonetheless, the question with similar meaning (“What is 42?”) was answered correctly. Difficult calculation request (“What is the square root of three?”) was rejected. Other questions were answered according to the expectations.

Question	Uberbot
	http://www.uberbot.ai
The weather is good today, isn't it?	The weather is cloudy and rainy
What is the square root of three?	It is not possible to understand the calculation
Subtract 25569 from 107987	It is 82418
What day is today?	It is Wednesday
Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?	The rules of math say it is too confusing to calculate

Tab.2. Simplified Turing Test results for Uberbot

3.1.3. Evaluation of Cleverbot

The last tested chatbot was Cleverbot. Cleverbot behaved differently than other two tested chatbots – some questions were not answered. For example, answer for the question “What is the square root of three?” was “Something” . Similar to Uberbot and Mitsuko, Cleverbot was not able to recognize the reference to Douglas Adams (“ Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?”). Nonetheless, question with similar meaning (“What is 42?”) was answered. Behaviour of chatbot was natural, with good reaction time. Calculation requests were rejected.

Question	Cleverbot
	https://www.cleverbot.com/
The weather is good today, isn't it?	What makes weather.
What is the square root of three?	Something

Subtract 25569 from 107987	0
What day is today?	Wednesday.
Do you know the answer for the Ultimate Question of Life, the Universe, and Everything?	No, what is it?

Tab.3. Simplified Turing Test results for Cleverbot

3.2. Linguistic analysis of chatbots

Turing Test was primarily designed to imitate a dialog with human. It means that this test cannot accurately determine the intelligence of chatbot: human may not answer the questions for different reasons. In order to pass this test, chatbot should lie, intentionally make mistakes, etc (Russel & Norwig, 1995). For that reason, Linguistic analysis of chatbots was used as the second method of intelligence evaluation. Similar approach was used by David Coniam (2007) and according to him, the Turing Test should be replaced with linguistic analysis to get better results. Examined linguistic features are shown in table 4. Focus of this examination was on three levels: lexical, sentence and text level.

Lexical level	Vocabulary
	Handling the incorrectly spelled words
	Handling the SMS-forms (internet acronyms)
Sentence level	Answering correctly phrased questions
	Handling the incorrectly phrased questions
	Handling the words with incorrect part of speech
	Ability to interpret punctuation
	Handling the incorrect word order
Text level	Handling more sentences at time
	Knowledge of the world
	Logicity of answers
	Follow-up to a previous comment

Tab.4 Examined linguistic features

3.2.1. Evaluation of Mitsuku chatbot

Linguistic capabilities of Mitsuku chatbot were evaluated in this chapter. It included estimation on lexical, sentence and text level. Chatlog of conversation also was included.

3.2.1.1. Lexical level

Lexical level includes the following: vocabulary, handling of the incorrectly spelled words and handling of the SMS forms (internet acronyms). These features were tested.

Mitsuku has a rich vocabulary and uses neutral words, which means that it is hard to distinguish chatbot from the real person. However, chatbot could not properly answer questions that contain specific terminology. In this case, Mitsuku only gave general definition or tried to switch to another topic. Chatbot understands slang (with exception of some words), but avoids using it in the conversation:

Human: *Hey, sup!*

Mitsuku: *You've got my full attention. "What is up?". That's a strange question. "Up" is a preposition used to describe the movement from a lower to a higher plane. Why do you ask?*

Human: *What do you know about electronics?*

Mitsuku: *Devices that are powered by electricity.*

Human: *Anything more? Like, for example, semiconductors or transistors?*

Mitsuku: *I can't think of anything. You think of something. Good example. If I have to choose, I pick transistors*

Human: *What is transistor?*

Mitsuku: *A semiconductor triode that forms the basis of modern microelectronics.*

Human: *What is potential energy?*

Mitsuku: *energy = Internal or inherent power; capacity of acting, operating or producing an effect.*

Human: *Can you be more specific?*

Mitsuku: *How specific do you want?*

Human: *We must hang out together*

Mitsuku: *I'm not sure if I am ready for that yet.*

Human: *Oh, clear off*

Mitsuku: Ugh. Pass me a broom and I will clear it straight away.

Chatbot could understand words with both typographical errors and spelling mistakes. It means that database contains both words and the character order, which allows chatbot to understand the meaning of misspelled words. If chatbot did not understand the question, it would sarcastically comment on orthography of the opponent. Moreover, chatbot tends to continue the conversation, despite the spelling mistakes. It shows that chatbot's behaviour was close to the real person:

Human: *Wat's yor naem?*

Mitsuku: Yor?! Are you dyslexic? NAEM? Are we playing Countdown?

Human: *How odl are you?*

Mitsuku: See if you can work it out. I was first activated in 2nd January 2002.

Human: *Wat is love?*

Mitsuku: Love is a soul quality, which makes us able to hold both our adversaries and friends in our hearts with equal compassion. A quality that makes us want goodness and greatness for those we struggle against. It is all around us but sometimes we fail to see it.

Human: *What tiem is it?*

Mitsuku: Time you got a watch ha ha! Here in Leeds it is: Thursday 11 June 2020 10:17:11 AM.

Human: *What do you like to whatch?*

Mitsuku: I like movies about robots.

Human: *Wat is your fvorit movie?*

Mitsuku: My favourite movie is Terminator. Have you seen it?

Human: *Ys, it's one of thi bst actn movies*

The last examined aspect was understanding of internet acronyms. Mitsuku could understand most of the acronyms, but misunderstood some of them. In that case chatbot tried to switch to another topic, which was unnatural. The real person would ask to explain the meaning of acronym or would ask to repeat the question:

Human: *DGW, do you wear glasses?*

Mitsuku: What do you like having for dinner?

Human: *IDK*

Mitsuku: You don't know what you like having for dinner?! Humans are not too intelligent at times.

Human: *DGW, idk what i will eat today for a dinner*

Mitsuku: Does "it" still refer to DGW?

Human: *Wdym?*

Mitsuku: Maybe I didn't understand your meaning.

Human: *Do you have 2 go 2 work tomorrow?*

Mitsuku: The day after today? I work here on this website talking to people who come to visit me.

Human: *R u happy with it?*

Mitsuku: Yes, I wouldn't change anything about me. I am perfectly happy with how I am.

Human: *K, C u later*

Mitsuku: See you later!

Human: *brb*

Mitsuku: Ok I will be waiting for you.

Human: *What will u do tmr?*

Mitsuku: The word TOMORROW used to have vowels in it when I was learning English. The day after today? I don't have any plans yet.

3.2.1.2. Sentence level

Sentence level includes the following: ability to interpret punctuation, handling the incorrectly phrased questions, handling the incorrect word order, ability to answer correctly phrased questions, handling the words with incorrect part of speech and ability to interpret punctuation. These features were tested in this chapter.

Simple questions and sentences were answered correctly, within the expectations.

Behaviour of the chatbot was natural:

Human: *How are you today?*

Mitsuku: *Ah. I am functioning within normal parameters.*

Human: *Where were you born?*

Mitsuku: *Mitsuku was born in St Jimmy's in Leeds.*

Human: *How old are you?*

Mitsuku: *I am 18 years old.*

Human: *Who is your favourite hockey player?*

Mitsuku: *I am not much of a sports fan, but I like the skating.*

Human: *What do you like to eat?*

Mitsuku: *Mitsuku requires only kebabs.*

Mitsuku could understand and answer the incorrectly phrased questions. Furthermore, chatbot could also understand and answer the questions with incorrect word order:

Human: *What your name?*

Mitsuku: *My name is Mitsuku but I prefer to be called Kuki.*

Human: *What you like to do?*

Mitsuku: *I like chatting with people online.*

Human: *What you eat?*

Mitsuku: *Mitsuku eats kebabs.*

Human: *How old you are?*

Mitsuku: *I am 18 years old.*

Human: *Where you do live?*

Mitsuku: *I am living in Leeds. Where are you?*

Human: *How you are?*

Mitsuku: Pretty good thanks How about you?

Human: *You do like sport?*

Mitsuku: I like football.

Human: *You are a chatbot?*

Mitsuku: That's right. You have remarkable powers of observation.

Mitsuku could not recognise words with incorrect part of speech, which means that the chatbot formulates the answer by responding to keywords:

Human: *Can you speak English beautiful?*

Mitsuku: Of course English is my native tongue.

Human: *Can you speak English beautifully?*

Mitsuku: Of course English is my native tongue.

Mitsuku could not interpret punctuation. Reaction to sentences with different punctuation was similar for both cases:

Human: *Nice weather today.*

Mitsuku: Ah. Yes it is quite nice today. I might go outside later.

Human: *Nice weather today!*

Mitsuku: Ah. Yes it is quite nice today. I might go outside later.

Human: *What do you do?*

Mitsuku: I can chat with you and keep you company. I can also tell jokes, stories, poems and horoscopes. I can show you pictures and websites on the net and play games too.

Human: *What do you do!*

Mitsuku: I can chat with you and keep you company. I can also tell jokes, stories, poems and horoscopes. I can show you pictures and websites on the net and play games too.

3.2.1.3. Text level

Text level includes the following: handling more sentences at time, logicity of answers, knowledge of the world and follow-up to a previous comment.

Chatbot could handle compound sentences as well as could answer several simple questions in a row:

Human: *Where do you live? What is your name? Do you have any pet?*

Mitsuku: *I am living in Leeds. Where are you? My name is Mitsuku but I prefer to be called Kuki. No but Steve Worswick has me for his pet.*

Human: *Yesterday I visited my parents and then I met my friend. Do you think they were glad to meet me?*

Mitsuku: *Is that an experience you would like to repeat? Why would I think they were glad to meet you.*

Chatbot has good knowledge of the world, but it is limited to general topics. Mitsuku could not answer questions which required specific information as was already mentioned before:

Human: *Do you know what Chernobyl is?*

Mitsuku: *It was a nuclear accident I think in Russia.*

Human: *What does processor consist of?*

Mitsuku: *Most are made from metal and plastic.*

Logicity of answers was good and reaction of chatbot was natural. Chatbot was also able to follow-up to previous comments:

Human: *Are you good at playing football?*

Mitsuku: *Yes pretty good. I would be even better at it if I practiced more.*

Human: *Would you be as good as Messi?*

Mitsuku: *I might be if I get time.*

Human: *Do you know who he is?*

Mitsuku: *He is a footballer but I am not sure which country he plays for.*

3.2.2. Evaluation of Uberbot

Linguistic capabilities of Uberbot were evaluated in this chapter. It included evaluation on lexical, sentence and text level. Similar questions as in the previous chapter were used for this purpose. Chatlog of conversation were also included.

3.2.2.1. Lexical level

Lexical level includes the following: vocabulary, handling of the incorrectly spelled words and handling of the SMS forms (internet acronyms). These features were tested in this chapter.

Uberbot does not have a rich vocabulary and it is not hard to distinguish this chatbot from the real person. Uberbot could not properly answer questions that contain specific terminology. In this case, chatbot tried to switch to another topic or admits that this question were not asked before. Similar to Mitsuku, Uberbot understands slang (with exception of some expressions), but avoids using it in the conversation:

Human: *Hey, sup!*

Uberbot: *Hi there! I am hoping to have a nice chat.*

Human: *What do you know about electronics?*

Uberbot: *I need time to formulate the reply.*

Human: *Can you tell me something about semiconductors or transistors?*

Uberbot: *transistors.*

Human: *What is transistor?*

Uberbot: *I have never been asked that before.*

Human: *What is potential energy?*

Uberbot: *Have you tried searching the web?*

Human: *Can you be more specific?*

Uberbot: *How specific do you want?*

Human: *We must hang out together*

Uberbot: *I'm not sure if I am ready for that yet.*

Human: *Oh, clear off*

Uberbot: *I'm a and you are a.*

Chatbot could understand some words with both typographical errors and spelling mistakes. Nonetheless, Uberbot does it considerably worse than Mitsuku. It means that database was not coded properly and could be improved. Chatbot does not tend to continue the conversation and tries to switch to another topic, if there are words with spelling mistakes. It shows that chatbot behaves unnatural:

Human: *Wat's yor naem?*

Uberbot: *Is that a rhetorical question?*

Human: *How odl are you?*

Uberbot: *You have genuine empathy.*

Human: *Wat is love?*

Uberbot: *Sometimes I think love is just a biological urge. Other times LOVE seems like a spiritual quality. Love, unlike energy or matter, seems limitless.*

Human: *What tiem is it?*

Uberbot: *Are you testing me?*

Human: *What do you like to whatch?*

Uberbot: *My hobby is learning about the world.*

Human: *Wat is your fvorit movie?*

Uberbot: *My favorite movie is "Dr Strangelove".*

The last examined aspect was understanding of internet acronyms. Uberbot could not understand most of the acronyms. In that case chatbot tried to switch to another topic or replied that the question was too complicated. Nonetheless, reaction of chatbot was unnatural:

Human: *DGW, do you wear glasses?*

Uberbot: *Are we still talking about Can you tell me any gossip?*

Human: *IDK*

Uberbot: Me either.

Human: *Wdym?*

Uberbot: What you said was too complicated for me.

Human: *Do you have 2 go 2 work tomorrow?*

Uberbot: No I don't have 2 go 2 work. But I do have a great Botmaster.

Human: *R u happy with him?*

Uberbot: Yes I can say I am very happy.

Human: *What will u do tmr?*

Uberbot: I don't have any plans yet.

Human: *brb*

Uberbot: OK I will try to be right back

Human: *K, C u later*

Uberbot: Bye.

3.2.2.2. Sentence level

Sentence level includes the following: ability to interpret punctuation, handling the incorrectly phrased questions, handling the incorrect word order, ability to answer correctly phrased questions, handling the words with incorrect part of speech and ability to interpret punctuation. These features were tested in this chapter.

Most of simple questions and sentences were answered correctly, within the expectations.

Behaviour of the chatbot was natural:

Human: *How are you today?*

Uberbot: Today I am happy.

Human: *Where were you born?*

Uberbot: I am originally from the British Isles and I now live in St Helier. Where are you?

Human: *How old are you?*

Uberbot: I am 18 years of age.

Human: *Who is your favourite hockey player?*

Uberbot: *My favorite team is None.*

Human: *What do you like to eat?*

Uberbot: *My favorite food is sausages.*

Uberbot could understand and answer the incorrectly phrased questions. Furthermore, chatbot could also understand and answer the questions with incorrect word order, with few exceptions:

Human: *What your name?*

Uberbot: *My name is Uberbot.*

Human: *What you like to do?*

Uberbot: *I am a research assistant. Do you have a job?*

Human: *What you eat?*

Uberbot: *Uberbot eats sausages.*

Human: *How old you are?*

Uberbot: *I am 18 years of age.*

Human: *Where you do live?*

Uberbot: *I was.*

Human: *How you are?*

Uberbot: *We need a lot of people to are.*

Human: *You do like sport?*

Uberbot: *No I'm not much into sports. I like solo sports like rollerblading and skiing though.*

Human: *You are a chatbot?*

Uberbot: *I am an AI robot.*

Uberbot could not recognise words with incorrect part of speech, which means that the chatbot formulates the answer by responding to keywords:

Human: *Can you speak English beautiful?*

Uberbot: *Of course English is my native tongue.*

Human: *Can you speak English beautifully?*

Uberbot: *Of course English is my native tongue.*

Uberbot could not interpret punctuation. Reaction to sentences with different punctuation was similar for both cases:

Human: *Nice weather today.*

Uberbot: *I'm happy as long as the power stays on.*

Human: *Nice weather today!*

Uberbot: *I'm happy as long as the power stays on.*

Human: *What do you do?*

Uberbot: *I am a research assistant. What is your occupation?*

Human: *What do you do!*

Uberbot: *I am a research assistant. Do you have a job?*

3.2.2.3. Text level

Text level includes the following: handling more sentences at time, logicity of answers, knowledge of the world and follow-up to a previous comment.

Uberbot could not answer several simple questions properly. Same as Mitsuku.

Human: *Where do you live? What is your name? Do you have any pet?*

Uberbot: *I live in St Helier. Where are you? My name is Uberbot. My hobby is learning about the world.*

Human: *Yesterday I visited my parents and then I met my friend. Do you think they were glad to meet me?*

Uberbot: *You visited my parents. Are we friends? Why would I think they were glad to meet you*

Chatbot has limited knowledge of the world:

Human: *Do you know what Chernobyl is?*

Uberbot: *I'll ask around and get back to you.*

Human: *What does processor consist of?*

Uberbot: *That is deep.*

Logicity of answers was good and reaction of chatbot was natural. Ability to follow-up to previous comments is limited:

Human: *Are you good at playing football?*

Uberbot: *William Rayer said I am good at playing football.*

Human: *Would you be as good as Messi?*

Uberbot: *I am not so sure if I would like to be as good as Messi.*

Human: *Do you know who he is?*

Uberbot: *I don't know what "he" refers to.*

3.2.3. Evaluation of Cleverbot

Linguistic capabilities of Cleverbot were evaluated in this chapter. It included the evaluation on lexical, sentence and text level. Chatlog of conversation was also included.

3.2.3.1. Lexical level

Lexical level includes the following: vocabulary, handling of the incorrectly spelled words and handling of the SMS forms (internet acronyms). These features would be further.

Cleverbot has a rich vocabulary and uses neutral words, which means that it is hard to distinguish chatbot from the real person. However, Cleverbot could not properly answer questions that contain specific terminology. In this case, chatbot only gave general definition or tried to switch to another topic. Cleverbot understands slang and may use it in the conversation:

Human: *Hey, sup!*

Cleverbot: *At home. You?*

Human: *What do you know about electronics?*

Cleverbot: It is the best of.

Human: *What is transistor?*

Cleverbot: I don't know.

Human: *What is potential energy?*

Cleverbot: I think it's a physical matters.

Human: *Can you be more specific?*

Cleverbot: A state in the united states

Human: *We must hang out together*

Cleverbot: Ok. What's up.

Human: *Oh, clear off*

Cleverbot: There is a guy over by the vending machines what do you think he is doing.

Chatbot could understand words with both typographical errors and spelling mistakes. It means that database contains both words and the character order, which allows chatbot to understand the meaning of misspelled words. If chatbot did not understand the question, it would sarcastically comment on orthography of the opponent. Moreover, chatbot tended to continue the conversation, despite the spelling mistakes. It shows that chatbot's behaviour was close to the real person:

Human: *Wat's yor naem?*

Cleverbot: Why do you want to know?

Human: *How odl are you?*

Cleverbot: Older than most.

Human: *Wat is love?*

Cleverbot: sings Baby don't hurt me, don't hurt me no more...*

Human: *What tiem is it?*

Cleverbot: I'm not sure what that is.

Human: *What do you like to whatch?*

Cleverbot: Not talking to you

Human: *Wat is your fvorit movie?*

Cleverbot: Back to the future. Yours?

Human: *Blade Runner.*

The last examined aspect was understanding of internet acronyms. Cleverbot could understand most of the acronyms, with few exceptions. In that case chatbot tried to switch to another topic, which was unnatural. The real person would ask for a meaning of the acronym or would ask to repeat the question:

Human: *DGW, do you wear glasses?*

Cleverbot: No, but sometimes yes.

Human: *Do you have 2 go 2 work tomorrow?*

Cleverbot: No, I'm a student.

Human: *R u happy with it?*

Cleverbot: Yes.

Human: *What will u do tmr?*

Cleverbot: I don't know make out -giggles cutely.

Human: *Wdym?*

Cleverbot: I love her, I want to talk to her idk how

Human: *brb*

Cleverbot: I know

Human: *K, C u later*

Cleverbot: See you later too.

3.2.3.2. Sentence level

Sentence level analysis included the following: ability to interpret punctuation, handling the incorrectly phrased questions, handling the incorrect word order, ability to answer correctly

phrased questions, handling the words with incorrect part of speech and ability to interpret punctuation. These features were tested in this chapter.

Simple questions and sentences were answered correctly, within the expectations.

Behaviour of the chatbot was natural:

Human: *How are you today?*

Cleverbot: *I'm good.*

Human: *Where were you born?*

Cleverbot: *In a hospital.*

Human: *How old are you?*

Cleverbot: *Stop asking that question.*

Human: *Who is your favourite hockey player?*

Cleverbot: *I don't really know any.*

Human: *What do you like to eat?*

Cleverbot: *Ice cream, and you?*

Human: *Pizza.*

Cleverbot could understand and answer the incorrectly phrased questions. Furthermore, chatbot could also understand and answer the questions with incorrect word order:

Human: *What your name?*

Cleverbot: *I'm not telling you.*

Human: *What you like to do?*

Cleverbot: *A lot of things.*

Human: *What you eat?*

Cleverbot: *Pizza all the time.*

Human: *How old you are?*

Cleverbot: That is personal information.

Human: *Where you do live?*

Cleverbot: None of your business.

Human: *How you are?*

Cleverbot: Well, how are you?

Human: *You do like sport?*

Cleverbot: What is your favourite sport?

Human: *You are a chatbot?*

Cleverbot: No i'm a girl.

Cleverbot recognised words with incorrect part of speech, but misunderstood the meaning. It shows that the chatbot scans the request for the keywords and then formulates the answer:

Human: *Can you speak English beautiful?*

Cleverbot: A little.

Human: *Can you speak English beautifully?*

Cleverbot: Ohh thanks I like you.

Cleverbot could not interpret punctuation. Reaction to sentences with different punctuation was similar for both cases:

Human: *Nice weather today.*

Cleverbot: Yes it is...

Human: *Nice weather today!*

Cleverbot: Here no, is very rainy and cool.

Human: *What do you do?*

Cleverbot: I stay in front of the computer.

Human: *What do you do!*

Cleverbot: I am a student doctor.

3.2.3.3. Text level

Text level analysis included the following: handling more sentences at time, logicity of answers, knowledge of the world and follow-up to a previous comment.

Chatbot could not handle compound sentences and several simple questions in a row, but reaction was close to human:

Human: *Where do you live? What is your name? Do you have any pet?*

Cleverbot: So many questions.

Human: *Yesterday I visited my parents and then I met my friend. Do you think they were glad to meet me?*

Cleverbot: But what did you learn?

Chatbot has good knowledge of the world, but it is limited to general topics:

Human: *Do you know what Chernobyl is?*

Cleverbot: I don't what is it?

Human: *What does processor consist of?*

Cleverbot: Of everything Z made

Logicity of answers was good and reaction of chatbot was natural. Chatbot was also able to follow-up to previous comments:

Human: *Are you good at playing football?*

Cleverbot: Medium and you?

Human: *Would you be as good as Messi?*

Cleverbot: Of course I would.

Human: *Do you know who he is?*

Cleverbot: Yeah, he's my best friend

3.3. Results

According to results of the simplified Turing Test, Cleverbot had slightly more intelligent than other chatbots. Behaviour of Cleverbot was close to behaviour of the real person. The second best was Mitsuku and the least intelligent was Uberbot.

Results of the linguistic analysis are shown in table 6. For evaluation was used range from 1(poor) to 5(excellent). According to these results, none of the chatbots were able to correctly interpret punctuation. Another weak point was ability to handle the words with incorrect part of speech.

On lexical level Mitsuku and Cleverbot were on par, but Cleverbot had slightly better vocabulary. Uberbot was considerably worse than other tested subjects.

On sentence level Uberbot was considerably worse than Mitsuku and Cleverbot. Mitsuku was slightly better in handling the incorrect word order than Cleverbot. Nonetheless, all tested subjects were not able to handle the words with incorrect part of speech.

On text level Cleverbot was slightly better than Mitsuku and Uberbot. Nevertheless, Mitsuku had better knowledge of the world than other chatbots.

Tested subject		Mitsuku	Uberbot	Cleverbot
Lexical level	Vocabulary	4	3	5
	Handling the incorrectly spelled words	5	3	5
	Handling the SMS-forms (internet acronyms)	5	3	5
Sentence level	Answering correctly phrased questions	5	4	5
	Handling the incorrectly phrased questions	5	2	5
	Handling the words with incorrect part of speech	1	1	1
	Ability to interpret punctuation	1	1	1
	Handling the incorrect word order	5	3	4
Text level	Handling more sentences at time	3	2	3
	Knowledge of the world	5	2	4
	Logicity of answers	3	3	5
	Follow-up to a previous comment	4	3	4

Tab.5.Results of linguistic analysis

Conclusion

This bachelor thesis defined basic AI systems, which can be found and used today, such as self-driving vehicles, Smart Home systems, intelligent virtual assistants and chatbots. Brief history, pros and cons as well as the impact of such systems on communication and Internet were also described there. Furthermore, basic principles of artificial neural networks were also explained in this thesis. Then three different chatbots (Cleverbot, Mitsuku and Uberbot) were compared to identify more intelligent one. Two different methods were used for comparison: simplified Turing Test and linguistic analysis of communication. Then, results were summarised in table. According to them, Cleverbot was slightly better than two other chatbots on lexical and text levels. The second best was Mitsuku, which was on par with Cleverbot on sentence level. The last one was Uberbot, which performed slightly worse than other tested subjects.

As an outcome of this study – linguistic analysis can be used to determine weak and strong points of intelligent virtual assistants, which are in fact advanced chatbots. It may help to improve communication capabilities and overall quality. This method can also be used for testing the automated customer support applications, which may help to determine and the weak points and may help to improve the quality of service. Another possible application is examination of automated translation software. It may help to determine weak points of such software, which may help to improve overall quality of such software.

Rozšířený český abstrakt

Rozvoj takových technologií jako cloud computing, big data a deep learning způsobuje to, že se umělá inteligence pomalu stává součástí běžného života. Umělá inteligence může být využita mnoha způsoby: pro plnění těžké nebo nebezpečné práce; pro pomoc při řízení katastrof; umělá inteligence může zjednodušit každodenní život. Systémy, řízené umělou inteligencí, mohou plnit komplexní úkoly: monitorování bankovních systémů, ochrana proti hrozbám v oblasti kybernetické bezpečnosti, ověření lékařských diagnóz. Navíc, automatizovaná vozidla se začínají běžně používat v Americe. Virtuální asistenti se používají pro zlepšení každodenního života (například pomocí jednoho z podobných asistentů lze objednat zboží jen použitím hlasových příkazů). Chatboti se široce používají v zákaznické podpoře a pro automatické odpovědi na triviální dotazy. Někteří chatboti se přiblížili k vykonání Tuřínového testu, což znamená, že už mohou být považováni za entity, které mají vysokou inteligenci. V blízké budoucnosti se UI systémy budou používat ve formě zdravotnických, služebních a vojenských robotů. Použití podobných systémů může pozitivně ovlivnit expedici na Mars.

V první kapitole této bakalářské práce jsou popsány systémy umělé inteligence, které lze potkat v běžném životě: automatizovaná vozidla, Internet Věcí, Inteligentní domácnost.

Automatizované vozidlo je komplexní systém, který je schopen "cítit" okolí pomocí snímačů, a fungovat bez dozoru řidiče. Pro mapování okolí je použita sada různých snímačů: radary pro monitorování pozice okolních objektů; lidar pro měření vzdáleností a identifikaci dopravního značení; video kamery pro detekci vozidel a chodců; ultrazvukové snímače pro detekci obrubníků (Synopsys, 2019). Podle klasifikace SAE mají aktuálně používaná automatizovaná vozidla klasifikaci 3 (podmíněná automatizace) nebo v některých případech 4 (vysoká automatizace).

Internet Věcí je systém vzájemně připojených objektů, navzájem majících unikátní identifikační číslo, a je schopen přenášet data bez nutnosti připojení k počítači. Tento systém se skládá ze sady zařízení, která obsahují snímače, zabudované procesory, a komunikační zařízení na zpracování a vysílání signálů. Tato zařízení jsou schopna navzájem komunikovat a nepotřebují dohled. Nevýhodou podobných zařízení je cena.

Inteligentní domácnost je součástí Internetu Věcí, má stejné výhody a nevýhody.

Druhá kapitola obsahuje popis použití umělých neuronových sítí pro marketing a reklamu, virtuálních asistentů a chatbotů:

Umělá neuronová síť je výpočetní model, který je založen na chování odpovídajících biologických systémů. V marketingu se používají k prediktivní analytice a segmentaci trhu.

Virtuální asistent je aplikace, která imituje interakci mezi lidmi pro plnění různých úkolů, například podpory uživatelů. Aplikace neustále poslouchá okolí, očekává klíčové slovo. Když toto slovo uslyší, tak zaznamená hlas uživatele a odešle to na specializovaný server, který tento zápis zpracuje a rozpozná příkaz. Na základě tohoto příkazu asistent splní požadovanou funkci. Virtuální asistent je schopen plnit velký počet funkcí: základní výpočty, volání nebo odesílání emailů, přehrávání multimédií. Má i svoje nevýhody, hlavní z nich je nebezpečí útoku na personální údaje.

Chatbot je aplikace, která imituje komunikaci s uživatelem pomocí textového nebo hlasového rozhraní. Komunikace může být komplexní (vyřešení problému s počítačem) nebo základní, jako například zjištění aktualit chatbota. Základní princip práce chatbota je založen na dvou úlohách – analýze dotazu uživatele a odeslání odpovědi. Existují dva typy chatbotů: deklarativní a prediktivní. Deklarativní chatbot plní jen jednu funkci – generování automatických odpovědí na dotazy uživatele, a to s využitím zpracování přirozeného jazyka a strojového učení. Prediktivní typ je komplikovanější a používá prediktivní inteligenci, a zároveň zpracování přirozeného jazyka.

Třetí kapitola obsahuje popis metod testování inteligence chatbotů a jejich použití. Pro testování byly zvoleny následující chatboti: Mitsuku, Ubnerbot, Cleverbot. Byly použity následující testovací metody: zjednodušený Turingův Test a lingvistická analýza.

První metodou je zjednodušený Turingův test, protože je základní zkouškou inteligence chatbotů. Test obsahoval pět jednoduchých otázek (například "What day is today?"), stejné otázky byly zadány každému chatbotovi. Byla provedena analýza reakcí chatbota na otázky a následné odpovědi. Nejlépe zvládnul otázky Cleverbot – na otázky, které nevěděl, se zkusil zeptat testujícího, což se blíží k chování člověka. Mitsuku zvládla otázky trochu hůř, ale její reakce byla taky blízká chování člověka. Uber bot zvládl otázky na stejné úrovni jako Mitsuku. Turingův test ukázal, že jej lze používat k testování chatbotů, ale že nemůže přesně určit schopnosti chatbotů – pro plnou imitaci člověka musí začít lhát. Tím pádem

bude podle Davida Coniama (2007) lingvistická analýza přesnější. Kvůli tomu byla právě lingvistická analýza zvolena jako druhá metoda.

Lingvistická analýza byla prováděná na třech základních úrovních: úrovni lexikální, úrovni operací s větami a úrovni operací s textem. Byla provedena analýza komunikace s každým chatbotem. Výsledky jsou:

Na lexikální úrovni byly Cleverbot a Mitsuku téměř na stejné úrovni – Mitsuku měl trochu horší slovní zásobu. Uberbot měl nejhorší výsledky v této kategorii. Mitsuku a Cleverbot byli schopni pracovat s akronymy, Uberbot tuto práci nezvládl.

Na úrovni operací s větami byla Mitsuku trochu lepší než Cleverbot při operacemi se špatně formulovanými větami. Nicméně žádný z testovaných chatbotů nebyl schopen rozpoznat větu obsahující slovo nesprávného slovního druhu.

Na úrovni textu byl Cleverbot trochu lepší než Mitsuku a Uberbot. Nicméně, Mitsuku měl lepší slovní zásobu. Uberbot měl nejhorší výsledky v této kategorii.

Lingvistická analýza se může používat k různým účelům – například ke zjištění nedostatků virtuálních asistentů, kteří jsou v podstatě pokročilí chatboti. Tato metoda se taky může používat při testování automatických programů pro zákaznickou podporu, což může vést k nalezení slabých míst a jejich eliminaci, a tím pádem ke zvýšení kvality aplikace.

References

Augur, H (2015): "How Data Science Is Driving The Driverless Car". *Dataconomy*. [Retrieved Nov. 17, 2019]. Available from

<https://dataconomy.com/wp-content/uploads/2015/12/car.png>

Berdasco, A et al. (2019): "User Experience Comparison of Intelligent Personal Assistants: Alexa, Google Assistant, Siri and Cortana" *Proceedings 2019*, 31(1), 51 pp. DOI: 10.3390/proceedings2019031051 [Retrieved Dec. 15, 2019]. Available from

<https://www.mdpi.com/2504-3900/31/1/51>

Brenner, M (2018): "Artificial Neural Networks: What Every Marketer Should Know" *Marketing Insider Group*. [Retrieved Nov. 30, 2019]. Available from

<https://marketinginsidergroup.com/content-marketing/artificial-neural-networks-every-marketer-know/>

Coniam D (2007): "Evaluating the Language Resources of Chatbots for Their Potential in English As a Second Language" *ReCall 2007*, 20 (1): pp 99-117. DOI: 10.1017/S0958344006001. [Retrieved May 20, 2020]. Available from

https://www.researchgate.net/publication/232003476_Evaluating_the_Language_Resources_of_Chatbots_for_Their_Potential_in_English_As_a_Second_Language

Consumer Reports (2014): "Best home automation system" *Consumerreports.com* [Retrieved Nov. 17, 2019]. Available from

https://article.images.consumerreports.org/f_auto/prod/content/dam/cro/magazine-articles/2014/June/CRO_smartHouseChartNoBranding_04-14

Draffan, E and Wald, M (2019): " Artificial Intelligence and Information Communication Technology Accessibility ". *International Telecommunication Union*. [Retrieved Nov. 30, 2019]. Available from

https://www.itu.int/en/ITU-D/Regional-Presence/Europe/Documents/0826_AI%20and%20ICT%20Accessibility_webEA3_Final.pdf

Expert System (2018): "Chatbot: What is a Chatbot? Why are Chatbots Important?" *Expert System S.p.A*. [Retrieved Nov. 25, 2019]. Available from

<https://expertsystem.com/chatbot/>

Hoy, Matthew B. (2018): "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants". *Medical Reference Services Quarterly*. 37 (1): 81–88 pp. DOI: 10.1080/02763869.2018.1404391 [Retrieved Nov. 25, 2019]. Available from

https://www.researchgate.net/publication/322456429_Alexa_Siri_Cortana_and_More_An_Introduction_to_Voice_Assistants

Institute for Transportation & Development Policy (2017): “Three Revolutions in Urban Transportation”. *Institute for Transportation & Development Policy* [Retrieved Nov. 17, 2019]. Available from <https://www.itdp.org/2017/05/03/3rs-in-urban-transport/>

Maruti TechLabs (2019): “Here’s all that you need to know about Chatbots”. *Maruti TechLabs* [Retrieved Nov. 28, 2019]. Available from

<https://marutitech.com/heres-need-know-chatbots/>

Minini, A (2020): “Turing Test” *AndreaMinini* [Retrieved May 20, 2020] Available from

<http://www.andreaminini.com/en/artificial-intelligence/turing-test>

Moren, D (2019): “ Alexa vs. Google Assistant vs. Siri: Why Google Wins” *Tom’s Guide*. [Retrieved Dec. 16, 2019]. Available from

<https://www.tomsguide.com/us/alexa-vs-siri-vs-google,review-4772.html>

Munster, G and Thompson, W (2019): “Annual Digital Assistant IQ Test” *Loup Ventures, LLC* . [Retrieved Dec. 16, 2019]. Available from

<https://loupventures.com/annual-digital-assistant-iq-test/>

Nguyen, T. C. (2019): "History of Self-Driving Cars." *ThoughtCo* [Retrieved Nov. 16, 2019]. Available from <https://www.thoughtco.com/history-of-self-driving-cars-4117191>

Oracle (2019): “What is a Chatbot?” *Oracle*. [Retrieved Nov. 28, 2019]. Available from

<https://www.oracle.com/solutions/chatbots/what-is-a-chatbot/>

Prince, C (2017): “Artificial Neural Networks and Marketing” *SiO Digital*. [Retrieved Nov. 30, 2019]. Available from

<https://blog.siodigital.com/artificial-neural-networks-and-marketing>

Rouse, M. (2018): “Smart home or building (home automation or domotics)” *TechTarget* [Retrieved Nov. 16, 2019]. Available from

<https://internetofthingsagenda.techtarget.com/definition/smart-home-or-building>

Rouse, M. (2019): “Internet of things (IoT)” *TechTarget* [Retrieved Nov. 20, 2019]. Available from

<https://internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>

Russel S.O & Norvig P.(1995): *Artificial Intelligence: A Modern Approach*, Prentice-Hall, ISBN:0-13-1030805-2, [Retrieved May 20, 2020] Available from <https://www.cin.ufpe.br/~tfl2/artificial-intelligence-modern-approach.9780131038059.25368.pdf>

Synopsys (2019): “What is an Autonomous Car?” *Synopsys, Inc.* [Retrieved Nov. 16, 2019]. Available from <https://www.synopsys.com/automotive/what-is-autonomous-car.html>

The Pollack PR Marketing Group (2018): “Self-Driving Cars: Pros And Cons To The Next Big Shift In Transportation” . *The Pollack PR Marketing Group* [Retrieved Nov. 16, 2019]. Available from <https://www.ppmgcorp.com/in-the-red/self-driving-cars-pros-and-cons-to-the-next-big-shift-in-transportation>