Disability and Rehabilitation: Assistive Technology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iidt20

Dialogue agents for artificial intelligence-based conversational systems for cognitively disabled: a systematic review

Syed Mahmudul Huq, Rytis Maskeliūnas & Robertas Damaševičius

DISABILITY AND REHABILITATION: ASSISTIVE TECHNOLOGY
https://doi.org/10.1080/17483107.2022.2146768

Taylor & Francis
Taylor & Francis Group

REVIEW

Dialogue agents for artificial intelligence-based conversational systems for cognitively disabled: a systematic review

Syed Mahmudul Huq[a], Rytis Maskeliūnas[a] and Robertas Damaševičius[b]

[a]Faculty of Informatics, Kaunas University of Technology, Kaunas, Lithuania; [b]Faculty of Applied Mathematics, Silesian University of Technology, Gliwice, Poland

ABSTRACT

Purpose: We present a systematic literature review of dialogue agents for Artificial Intelligence (AI) and agent-based conversational systems dealing with cognitive disability of aged and impaired people including dementia and Parkinson's disease. We analyze current applications, gaps, and challenges in the existing research body, and provide guidelines and recommendations for their future development and use.
Materials and methods: We perform this study by applying Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criteria. We performed a systematic search using relevant databases (ACM Digital Library, Google Scholar, IEEE Xplore, PubMed, and Scopus).
Results: This study identified 468 articles on the use of conversational agents in healthcare. We finally selected 124 articles based on their objectives and content as directly related to our main topic.
Conclusion: We identified the main challenges in the field and analyzed the typical examples of the application of conversational agents in the healthcare domain, the desired characteristics of conversational agents, and chatbot support for aged people and people with cognitive disabilities. Our results contribute to a discussion on conversational health agents and emphasize current knowledge gaps and challenges for future research.

ARTICLE HISTORY
Received 21 March 2022
Revised 28 October 2022
Accepted 7 November 2022

KEYWORDS
Assisted living; chatbot; conversational agents; dialogue systems; natural language processing; digital health; e-Health

➤ IMPLICATIONS FOR REHABILITATION

- A systematic literature review of dialogue agents for artificial intelligence and agent-based conversational systems dealing with cognitive disability of aged and impaired people.
- Main challenges and desired characteristics of the conversational agents, and chatbot support for aged people and people with cognitive disability.
- Current knowledge gaps and challenges for remote healthcare and rehabilitation.
- Guidelines and recommendations for future development and use of conversational systems.

Introduction

Artificial Intelligence (AI) is concerned with how we can create systems that behave intelligently, where intelligence means knowing, anticipating, and behaving in a thoughtful way, something we can ascribe to human intelligence [1]. AI innovations can have a dramatic effect on the lives of People With Disabilities (PWD). Indeed, for many state-of-the-art AI systems, enhancing the life of PWD is a motivating factor [2,3].

AI-based conversational agents are a technology that enables computers to recognize human language and decode various languages. They allow computers to understand what is being said and to decide and to respond to the right answer in a manner that imitates human conversation. These conversational agents are powered by a variety of models including Automatic Speech Recognition (ASR), Text Analysis, Text-To-Speech (TTS), and Speaker Analysis. Chatbots serve conversational experiences to their users for various practical applications including health support [4]. During completing new tasks, conversational agents minimize the workload of users and provide cognitive support to individuals with mental and intellectual disabilities that impair memory and executive functioning. Chatbots programmed with the intention of being indistinguishable from a human when the classic Turing experiments were performed were some of the first cases of conversational agents [5]. These chatbots were used in check-in tests in which human users conversed with them (typing on a computer) to see whether they were conversing with a human or a machine.

A surge in interest in AI methods has been fueled by the popularity of conversational agents, especially those that can use some unconstrained natural language input. Recent progress in machine learning, especially in neural networks, has enabled more complex methods of dialogue management and greater conversational versatility [6]. Customers are increasingly using smartphone conversational agents for daily activities such as extracting details and handling schedules [7]. This is made possible by the advent of more efficient and mobile computers, as well as increased access to personalized data from sensors. Advances in speech recognition, the processing of natural language, and AI have contributed to the growing availability and usage of systems of conversational agents that use text and spoken language to imitate human conversation. Voice-enabled programs such as Microsoft Cortana, Apple Siri, Amazon Alexa, and

CONTACT Robertas Damaševičius robertas.damasevicius@polsl.pl Faculty of Applied Mathematics, Silesian University of Technology, Gliwice, Poland

Google Now and are familiar examples of conversational agents [8].

Through speech recognition, Natural Language Processing (NLP), natural language comprehension, and natural language generation, conversational agents may engage in two-way dialogues with the user. Artificial intelligence is used by conversational agents in order to conduct these dialogues [9]. These interfaces of the conversational agents can include dialogues based on text and/or spoken language. They are named chatbots, chatterbots or virtual agents, in different forms. In order to have a richer interactive experience, some conversational agents present a human image (for example, the image of a coach or a nurse) or a non-human image (for example, a robot or an animal). These are called Embodied Conversational Agents (ECAs) [10].

In recent years, conversational agents perform an increasingly critical role in medical services and medical care [4]. These agents are used to assist physicians during a consultation, to help customers with issues of behavioral improvement, and to assist patients and elderly people in their living environments. Over the past two decades, a substantial body of research has demonstrated the benefits of using ECAs for health-related uses [11–13]. Several randomized clinical trials of ECAs have achieved great improvements in consumers' physical habits and access patterns to online health records [14]. However, most of these agents only permit restricted user feedback (e.g., multiple choices) and do not have the ability to comprehend natural language input.

The tendency of humans to perceive machines as social agents can be used by conversational agents. In order to monitor the well-being and symptoms at home, most adults have stated that they would use a smart virtual coach or an intelligent virtual nurse [15–17]. As a way of overcoming loneliness and alienation, such assistance could be useful. Conversational agents also are used in other applications to improve the health behavior of the users. Conversational agents are already implemented in the self management domain to treat depression [18], smoking and alcohol use cessation [19], allergy and asthma [20], diabetes [21], and tropical diseases [22]. The treatments that are given by human professionals cost two to three times more than that are given by conversational agents. Conversational agents also can provide social support and increased interaction, while remaining flexible and cost-efficient.

With today's technologies, elderly consumers of conversational agents can stay in their homes comfortably for longer periods of time. Commercially available products allow for remote surveillance of a user's health [23], the strengthening of social networks, and the assistance of elderly people in their everyday activities. While technology continues to be in a position to assist aging users, new technologies may not be effective in solving real-world user challenges such as self-efficacy and loneliness.

Even with substantial cognitive degradation, many elderly people maintain the capacity of keeping contact in a multimodal face-to-face fashion [24]. A broad variety of non-verbal ways are incorporated into face-to-face interaction to bring semantic information complementary to facial expressions. This helps people with disabilities to compensate for certain communication pathways (for example, hearing) by using other communication channels (for example, body gestures). Face-to-face dialogue is also featured by well-established comprehension repair mechanisms which enable the listener to request a repeat or clarification from the speaker. In addition, the face-to-face conversation has built-in mechanisms to limit the focus of the participants. This emphasis is important because it is difficult for some elderly people to divide their attention or manage distractions [24]. An agent-based support system for people with Alzheimer's disease and their associates has been developed by Peeters [25]. Such a support system helps patients remember their personal memories and thus strengthens their sense of identity, protection, security and self-esteem [25].

Avatars can be used to provide this type of face-to-face interaction [24]. Multiple benefits can be achieved by contact with avatars. Avatars may include movements that increase the comprehension of the information provided. In addition, by adding a lip-synced animated character to the audio speech output, the visual enrichment of verbal information can be achieved. This can increase the robustness of the transmitted information as understood from natural speech.

The conversational agents can be used to schedule training sessions, provide input during the sessions, and reflect on the results after the session has been finished [26]. Conversational agents can render themselves to serve as physical activity coaches and to assist elderly people in enhancing their well-being [27]. Smart conversational agents are used to create simple communication products for elderly users that are easy to use. The resulting design principles are based on known metaphors and seek to lower the barriers for contacts across the care networks at home to generate direct value for the end users [27]. Message-based communication on the internet may help to preserve and to strengthen social networks for old people with little or no previous experience with computers.

Recently, several survey articles on the conversational agents in healthcare were published [4,11–14,18,28–33]. These surveys are summarized in Table 1.

The aim of this study is to perform a systematic review of the current applications, gaps, and challenges in the existing research body on the AI-based conversational agents and dialogue systems in the healthcare domain and provide guidelines and recommendations for their future development and use. The specific difference of the current study from the previous systematic reviews and surveys is the focus on mental illnesses such as dementia and Parkinson's disease.

The remaining parts of this article are organized as follows. Section "Protocol of systematic literature review" presents and describes the protocol and the process of the systematic literature review. Section "Result of systematic literature review" describes the main results of the systematic literature review. Section "Analysis of AI-based conversational agents" presents the analysis of the particular AI-based conversational agents, dialogue management systems (DMS), and the specific tools. Section "Speech recognition datasets for chatbot training and testing" presents datasets related to speech recognition and artificial intelligence methods used by conversational agents. Section "Evaluation and discussion" evaluates and discusses the results and findings of the systematic review. Finally, Section "Conclusion" presents the conclusions.

Protocol of systematic literature review

Reporting standards

In this study, we used a Systematic Literature Review (SLR), which defined techniques for interpreting, identifying, and evaluating literature related to our research issue. The fundamental idea of SLR is to create a summary of a subject area's content, which provides more information than a typical literature review. This strategy entails gathering a huge number of publications in order to organize, analyze, and pinpoint critical gaps that should be addressed in future research. The goal of this article is to do an

Table 1. Summary and comparison of previous systematic reviews on conversational agents in healthcare.

Author (year)	Ref.	Methodology	Databases	Focus	Papers analyzed (Initial/Final)	Years covered
Hoermann et al. (2017)	[28]	PRISMA	MEDLINE, PsycINFO, Central, Scopus, EMBASE, Web of Science, IEEE, ACM	Text-based dialogue systems in mental health interventions	3192/24	All – 2016
Provoost et al. (2017)	[14]	Arksey & O'Malley [34]	PubMed, ScienceDirect, Web Of Science, ACM Digital Library, SpringerLink	Applications of embodied conversational agents in clinical psychology	1117/54	All – 2015
Laranjo et al. (2018)	[11]	PRISMA	PubMed, Embase, CINAHL, PsycInfo, ACM Digital	Conversational agents with natural language input capabilities for health-related purposes	1513/17	All – 2018
Gaffney et al. (2019)	[18]	PRISMA	MEDLINE, EMBASE, PsycINFO, Web of Science, Cochrane	Conversational agent interventions in the treatment of mental health problems	30,853/13	All – 2018
Vaidyam et al. (2019)	[12]	PRISMA	PubMed, EmBase, PsycINFO, Web of Science, Cochrane, IEEE Xplore	Role of conversational agents in screening, diagnosis, and treatment of mental illnesses	1466/10	All – 2018
Montenegro et al. (2019)	[29]	PICOC	ACM, CiteSeerx, IEEE, JMIR, PMC, Scholar, Science, Springer, Taylor & Francis, Wiley	Conversational agents in health	4145/40	2014-2018
Chattopadhyay et al. (2020)	[30]	PRISMA	Google Scholar, MEDLINE, EMBASE, PsycINFO, CINAHL, Cochrane, PubMed, ACM Digital Library	Humanlike computer-generated characters (virtual humans) for patient-oriented dialogue systems	8125/53	All – 2019
Schachner et al. (2020)	[13]	PRISMA	PubMed MEDLINE, EMBASE, PyscInfo, CINAHL, ACM Digital Library, ScienceDirect, Web of Science	AI-based conversational agents for chronic diseases	2052/10	All – 2020
Tudor Car et al. (2020)	[4]	PRISMA (Scoping review)	MEDLINE, EMBASE, PubMed, Scopus, Cochrane Central, OCLC WorldCat database, ResearchGate, Google Scholar, OpenGrey, Google	Conversational agents in health care	11,401/47	1966–2019
Milne-Ives et al. (2020)	[31]	PRISMA	PubMed, Medline, EMBASE, CINAHL, Web of Science, ACM Digital Library	Effectiveness and usability of conversational agents in health care	9441/31	2008–2020
Vaidyam et al. (2021)	[12]	PRISMA	PubMed, Embase, PsychINFO, Cochrane	Conversational agents for assessing serious mental illness: major depressive disorder, schizophrenia spectrum disorders, bipolar disorder, or anxiety disorder	247/7	2018–2020
Pacheco-Lorenzo et al. (2021)	[33]	PRISMA	Scopus, PubMed, Pro-Quest, IEEE Xplore, Web of Science, CINAHL, Cochrane Library	Smart conversational agents for detection of neuropsychiatric disorders	2356/17	All – 2020

SLR of current works on conversational bots in order to provide an overview of the field [35]. SLRs produce consistent results when used to take a picture of phenomena, such as new technology.

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist to conduct an SLR [36]. From the preparation of the review and selection of literary sources to the actual analysis of the identified works, this protocol formalizes the SLR process.

Search strategy

The group of conversational agents that use some unconstrained natural language feedback will be included for the purposes of this SLR. Input conversational agents who use only non-natural language communication have been omitted from the study. For the SLR, the following bibliographic databases and a web search engine have been used: ACM Digital Library, IEEE Xplore, PubMed, Google Scholar, and Scopus. The range of publication years was between 2000 and 2020. There was no restriction of the publication language.

Search terms for the SLR included the following: Conversational Agents. Dialogue Systems. Relational Agents. Finite-State Conversational Agents. Frame-Based Conversational Agents. Agent-Based Conversational Agents. Cognitive Disorder. Communication Disorder. Autism, and Chatbots.

The reference lists of relevant articles have been searched. Dissertations, theses, and conference proceedings identified in those bibliographic databases were also included for screening. The Boolean logic operator AND was employed to combine together separate categories and OR operator was used to combine the terms within categories.

- Search string for systematic literature review of the conversational agent using unconstrained natural language dealing with disability and helplessness of old people and disability of and helplessness of people with cognitive disorders and developmental disabilities including autism, dementia and Parkinson's disease:
- ″conversational agent*″ OR″ conversational system*″ OR″ dialogue system*″ OR″ relational agent*″ OR″ finite-state conversational agent*″ OR″ frame-based conversational agent*″ OR″ agent-based conversational agent*″ OR″ cognitive disorder*″ OR″ communication disorder*″ OR″ autism*″ OR″ chatbot*″.

Inclusion and exclusion criteria

We considered studies if they (1) were primary study studies involving health issues or conditions; (2) used a conversational agent; and (3) used some artificial intelligence method for data processing, such as speech recognition or deep learning.

The focus of the research on conversational agents must satisfy the following conditions:

- The conversational agent uses any unconstrained natural language as input in the spoken and in written form and in the visual communication medium.
- The conversational agent uses an agent-based method and technique for input and output.
- The conversational agent deals with the disability and helplessness of old people.
- The conversational agent deals with the disability and helplessness of people with cognitive disorders including autism, dementia and Parkinson's disease.

Articles were excluded if they (1) involved did not use any machine learning or deep learning method; (2) involved purely non-technical studies where the dialogue between the human and the program (system) was simulated by a human rather than executed by the conversational agent; (3) did not address any health conditions and diseases or any form of adverse health conditions such as bad habits; (4) addressed health on a general level without elaborating a specific condition or the health condition only had a minor role and was only briefly mentioned.

We also have excluded non-English papers, conference proceeding papers, workshop papers, literature reviews and surveys, poster presentations, technical reports, whitepapers, etc., or if the full text of the article was not accessible for the authors of this paper.

Selection process

All of the references found during the searches were downloaded and inserted into an Excel spreadsheet. Duplicates have been eliminated. Two independent reviewers performed a three-phase screening process, evaluating first the article names, then the abstracts, and finally the complete texts.

Cohen kappa was measured after each of these steps to assess the degree of agreement and to test inter-rater reliability between the researchers [37]. Any differences were addressed and resolved in a cooperative manner.

Data extraction

Three reviewers became acquainted with the listed papers and then independently extracted the details found therein into an Excel spreadsheet with 28 columns containing data on the following aspects: (1) general study detail, (2) health care/chronic conditions, (3) conversational agents, (4) artificial intelligence, and (5) other study items.

Writers, year of publication, study design/style, study goal, conversational agent assessment steps, key recorded outcomes and results, AI methodology, AI system creation were among the information we extracted. The data was then narratively synthesized. Because of the variety of the studies examined, the quality of the studies was not measured in this study. Any inconsistencies were discussed and solved following the consensus.

Risk of methodological bias

Given the variety of research designs and reported assessment measures, the author team had a lengthy discussion about selecting an appropriate tool to determine the methodological biases of the included studies. The Consolidated Standards of Reporting Trials (CONSORT) checklist [38] was used to measure bias. The list consists of 25 items, each of which can be given a score of 1 or 0 depending on whether the item was satisfactorily fulfilled or not. Lower scores indicate a greater chance of methodological bias, whereas higher scores indicate the opposite. Each analysis was rated individually by the paper reviewers, and Cohen kappa was used to determine inter-rater reliability between the two evaluations and scored at 87%.

Result of systematic literature review

Main Findings

The SLR of the conversational agent using unconstrained natural language dealing with the disability and helplessness of old people and the disability and helplessness of people with cognitive disorders including autism, dementia and Parkinson's disease.

As depicted in Figure 1, after duplicate removal, a total of 468 publications were collected from the five bibliographic databases searched and reviewed by the authors based on their titles and abstracts to find those that were possibly relevant to the study issue.

We then observed the guidelines by Zaveri et al. [39] to further screen the articles by their abstract and full text. The eligibility checking was done on the 468 selected papers, based on the Title and Abstract screening of them. Works were divided into three categories: those that were considered important and were immediately included in the study for further analysis; non-relevant studies; and those whose significance was unclear, which were solved by consensus decision by the reviewers.

Finally, the relevance of publications in the third category was determined by the two scholars who read the study reaching an agreement. Following debate among the writers who took part in the process, 307 papers were chosen for additional screening. Following the elimination criteria, the remaining 161 (20 ambiguous and 141 irrelevant) were removed. The biggest reason for exclusion (75 articles) was that the solution was not conversational, which meant that the user couldn't communicate with the system. Furthermore, 24 articles were eliminated since the conversational system was not clearly linked to any mental illness.

The remaining 307 articles were screened using the inclusion/exclusion criteria and 143 articles were excluded after reading the full text and evaluating their suitability, while the remaining 40 articles did not have their full text available for the reviewers, therefore, they were excluded as well. The result yielded 124 articles for further analysis.

Each of the following three dialogue management mechanisms is used by Conversational Agents:

- The Finite-State DMS interacts with users through dialogues consisting of sets of pre-determined steps and states.
- The Frame-Based DMS asks users questions that allow the system to perform tasks by filling slots in predefined frames.
- The Agent-Based DMS performs complex communication between the devices and the applications and the users to complete tasks.

The SLR has identified the following results regarding the dialogue management systems:

- 82 conversational agents (CAs) use finite-state DMS.

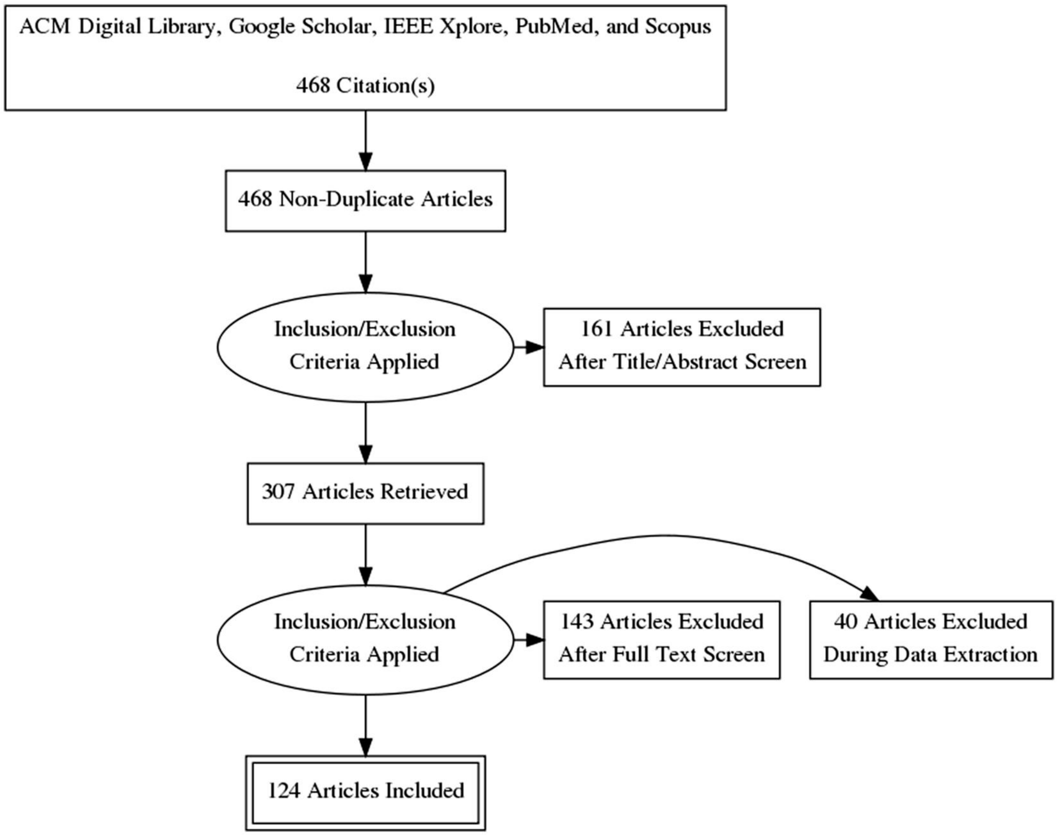

Figure 1. PRISMA flow diagram of the systematic literature review.

- 11 CAs use frame-based DMS.
- 20 CAs use agent-based DMS.

Through the SLR of the conversational agent using unconstrained natural language dealing with the disability and helplessness of old people and the disability and helplessness of people with cognitive disorders including autism, dementia, and Parkinson's disease the following have been identified:

- 12 CAs are dealing with the disability and helplessness of people with cognitive disorders. Among these CAs 8 agents use finite-state DMS and 2 agents use frame-based DMS. Among these CAs 2 agents use an agent-based dialogue system.
- 13 CAs are dealing with the disability and helplessness of old people. Among these CAs 5 agents use finite-state DMS and 2 agents frame-based use DMS. Among these CAs 6 agents use an agent-based dialogue system.
- 4 CAs are dealing with the disability and helplessness of people suffering from autism. Among these CAs 3 agents use finite-state DMS and 1 agent uses an agent-based dialogue system.
- 2 CAs are dealing with the disability and helplessness of people suffering from dementia. One of these CAs uses finite-state DMS and the other conversational agent uses frame-based DMS.
- 5 CAs are dealing with medical diagnostics of disability and helplessness of people suffering from the disease. These CAs use finite-state dialogue management.

We found in the SLR that 8 CAs are using AI-based and agent-based dialogue systems that are dealing with the disability and helplessness of old people and the disability and helplessness of people with cognitive disorders.

- 6 CAs using AI-based and agent-based dialogue systems are dealing with the disability and helplessness of old people.
- 2 CAs using AI-based and agent-based dialogue systems are dealing with the disability and helplessness of people with cognitive disorders.
- 5 CAs using AI-based and agent-based dialogue systems are working with spoken, written, and visual communication media as input and output.
- None of the CAs using AI-based and agent-based dialogue systems dealing with the disability and helplessness of old people and the disability and helplessness of people with cognitive disorders implement context awareness and adaptiveness.

Additionally, 10 DMSs for CAs are using AI-based and agent-based dialogue systems that are dealing with the disability and helplessness of old people and the disability and helplessness of people with the cognitive disorder have been found.

Results of systematic literature review of chatbots

According to the data of WorldCafe [40] and similar repositories, chatbots are the primary conversational agent technology giving some form of reflection for cognitive impairment (e.g., Woebot [41], Wysa [42], Laura [43]). These might be used to treat patients. The chatbots Woebot [44] and Help4Mood [45], for example, were created to provide cognitive behavioral therapy to patients suffering from depression and anxiety. Chatbots can also be used for training reasons. For example, the chatbots LISSA [46] and VR-JIT [47] were used to instruct autistic individuals to enhance their social skills and job interview abilities. Louse was used for field evaluation study of attention management with cognitively-impaired older adults [48]. Similar solutions have also been used to test for illnesses such as dementia [49,50], tobacco and alcohol use disorders [19], stress [51], and signs of depression and suicide [52], behavioral therapy [53] as well as for self-management [54], counselling [55], and education [56]. Second to text based interactions voice based solutions are offered in a limited adaptability and capacity [57], with some adaptations to people with intellectual disability [58,59] and different psychological conditions [60], including depression [61].

In compared to the artificial intelligence conversational agents reviewed in this study, rule-based variants are more straightforward to design and less prone to mistakes since their replies are preset and they do not need to produce new responses [58,59], where authors identified that a few key points. An advantage, they are thought to be more secure and responsible than their AI counterparts. One disadvantage of rule-based chatbots is that users cannot generally manage the discussion because their inputs are limited to predetermined terms and phrases. Regrettably, rule-based chatbots cannot reply to user inputs that do not conform to the established rules. Although rule-based chatbots are inadequate for complicated activities and enquiries due to the difficulty in anticipating every conceivable case. Same survey noted, that the majority of existing conversational agents target at people with cognitive disabilities are implemented as standalone software. This was surprising because web-based chatbots are thought to be more suited than standalone software for the two reasons listed below. To begin, consumers do not need to install a special program on their devices to utilize web-based chatbots, which reduces the possibility of violating their privacy. Second, web-based chatbots are easier to use than standalone chatbots.

Unfortunately, voice-based systems were not without own problems and only provided acceptable speech recognition performance for popular languages such as English; however, a drop in performance was noticeable when targeting people with cognitive disabilities, with single-letter responses being slightly better recognized than word-based responses. Survey [57] comes to a conclusion, that commercial variations appear to specialize in facilitating particular social activities, although not outperforming Google Voice Search when the intervention entailed looking for health information. There is a noticeable research gap, yet currently commercial voice based systems provide only superficial health-related assistance and only for the popular languages.

Results of systematic literature review of dialogue management systems

Merdivan et al. [24] published a research paper in 2019 stating that guided natural interaction and emotional intelligence play a very important role in enhancing the Human-Computer Interaction (HCI) with Information and Communication Technology (ICT) solutions. The capacity that many elderly people maintain, even with substantial cognitive degradation, is to keep contact in a multimodal face-to-face fashion [24]. A broad variety of non-verbal ways are incorporated into face-to-face interaction to bring semantic information complementary to facial expressions. This helps people with disabilities to compensate for certain communication pathways (for example, hearing) by using other communication channels (for example, body gestures). Face-to-face dialogues are also characterized by well-established comprehension repair mechanisms which enable the listener to request repeats or clarifications from the speakers. In addition, face-to-face conversations have built-in mechanisms to limit the focus of the participants. This is especially important for some of the elderly people because it is difficult for some elderly people to divide their attention or manage distractions [24]. Avatars can be used to provide this type of face-to-face interaction [24]. Multiple benefits can be achieved by contact with avatars. Avatars may include movements that increase the comprehension of the information provided [24]. In addition, by adding a lip-synced animated character to the audio speech output, the visual enrichment of verbal information can be achieved. This will increase the robustness of the transmitted information as understood from natural speech [24].

Gebhard et al. developed a method for virtual character applications in 2012 and introduced a platform called Visual SceneMaker [62]. The idea of context-sensitive interactive behavior of applications involves specialist programmers and often results in code that is difficult to maintain. In order to tackle these challenges, Gebhard et al. have developed the Visual SceneMaker tool [62]. The tool allows innovative, non-programming experts to create interactive applications with multiple virtual characters using a simple scripting language [62].

The Visual SceneMaker tool has interesting characteristics, namely the reusability of sceneflow components for the development of new interactive features in a quick prototyping style [62]. The sceneflow editor which is a part of the latest Visual SceneMaker IDE supports this feature [62]. By building nodes and edges, this graphical authoring tool supports writers with drag and drop functionality to draw a sceneflow. New structural extensions of the sceneflow model are also supported by the editor. Parallel sceneflows and a special background node type are included in the enhancements [62]. Complex sceneflows can be split into logical components with the use of concurrent sceneflows, minimizing the number of constructs and helping to organize the entire model [62].

The Visual SceneMaker IDE is based on a sceneflow execution interpreter [62]. The graphical authoring tool enables the visualization of a given SceneMaker application [62]. It also enables immediate observation of the effects caused by the model changes, even at runtime.

Gebhard et al. have intended to combine the Visual SceneMaker with speech synthesis, generation of non-verbal behaviour, emotion simulation, etc. as future work [62]. The new features of Visual SceneMaker have been designed to be incorporated into a component-based embodied agent framework [62].

The Web-Accessible Multimodal Interfaces (WAMI) toolkit was developed by Gruenstein et al. which provides a framework for the creation, deployment and evaluation of accessible multimodal interfaces. In such interfaces, users communicate using voice, mouse, pen, and touch [63]. The toolkit uses modern web development techniques to build browser-based apps that are accessible by a wide range of internet-connected devices [63].

The WAMI toolkit provides two key threads of development [63]. First, it can be employed to link a graphical user interface (GUI) to conventional GALAXY architecture-based spoken dialogue systems (developed by Seneff et al. in 1998) [63]. Before parsing, contextual integration, dialogue management, natural language production, and speech synthesis are used as the first steps in this architecture. It is possible to render interfaces constructed in this way to be understood by a wide variety of natural languages.

For developers who do not have expertise in speech recognition, parsing, dialogue management, and natural language production, WAMI offers a second lightweight development model [63]. The lightweight platform of WAMI offers a common structure in which creators can construct highly immersive multimodal applications with modest capabilities for understanding natural

Table 2. Summary of dialogue management frameworks.

Author/Year	Ref.	Salient characteristics of dialogue management framework
van Waterschoot et al./2018	[64]	Dialogue engine for ECA
Ultes et al./2017	[65]	Multi-domain statistical dialogue system toolkit
Lison et al./2016	[66]	Open source toolkit for building and evaluating spoken dialogue systems
Yu et al./2016	[67]	Open source web-based multimodal dialog framework
Barthol et al./2013	[68]	Flexible framework for exploring a variety of different types of virtual human systems
Gebhard et al./2012	[62]	Visual authoring approach for virtual character applications of ECA
Rich et al./2012	[69]	Dialogue system combining hierarchical task networks with dialogue trees that partially automates dialogue authoring and improves the degree of structure reuse
Skantze et al./2012	[70]	Dialogue system toolkit for rapid development of real-time systems for multi-party face-to-face interaction
Bohus et al./2009	[71]	Plan-based, task-independent dialog management framework
Gruenstein et al./2008	[63]	Framework for creating, installing, and testing web-based multi-modal interfaces through voice, mouse, pen, and touch interaction

language. WAMI's lightweight framework is very beneficial, as it offers a simple way to prototype new software easily [63].

The results of SLR on dialogue management frameworks are summarized in Table 2.

Analysis of AI-based conversational agents

In this section we discuss the particuler AI-based conversational agents identified by our systematic review. A typical conversational agent is a text-based healthcare chatbot (TBHC) that assists medical professionals in delivering clear text-based communications and media items, like videos, podcasts, with evidence-based interventions in a pervasive and automated way.

SELMA

SELMA is a TBHC for chronic pain self-management [54]. SELMA is displayed by means of text messages and further presentation and elaboration of materials based on previous response choices and topics of specific interest to the individual [54]. First, SELMA provides psychoeducation [72] from day one to day twenty-one in seven every day or other regular text message sequences that observe a similar structure: initial salutation, psychoeducational guidance, main tutorial, and goodbye [54]. Psychoeducation improves people's understanding of their illnesses while reducing fear and insecurity. This technique also increases clarity and decreases misconceptions of biopsychosocial pain management, thus reinforcing enthusiasm for therapy and self-help ability. Coping techniques enable participants to play an active role in their pain management by evaluating and adapting their actions, emotions and cognitions to control pain. It takes about 5 to 30 min to complete, considering whether or not a message series requires an exercise. The 18 intervention units each consist of 2 to 4 sequences of messages.

SELMA was developed using an open-source software platform for the design of mobile TBHCs called MobileCoach [https://www.mobile-coach.eu/] [54]. This platform enables the intervention writers to design digital health interventions (fully automated and script-based) consistent with the model of talk-and-tools. In other words, SELMA provides a basic chat-based interface with predefined choices for answers that can be used to connect with participants.

The SELMA linguistic style was based on the premise that interpersonal closeness is positively linked to the link bond between the patient and the chatbot in order to create a social partnership and working alliance with participants. SELMA's conversational style is likely to impact relationship-building mechanisms and, using emojis and some sense of humor, mimics a real human chat-based interaction. Supportive machine agents were favorably regarded. SELMA demonstrates compassion and affective empathy and stresses the completed activities of the participants. This technique is focused on mutual support and strives for a coaching style that is supportive. To further engage participants, SELMA sends out personalized text messages to start a conversation every day or every other day.

Diabot

The DIAgnostic chatBOT (Diabot) is a medical chatbot that uses advanced Natural Language Understanding (NLU) techniques to engage patients in conversation to provide personalized prediction using the health dataset and based on the different symptoms sought by the patient [73]. Using the Pima Indian diabetes dataset [https://www.kaggle.com/uciml/pima-indians-diabetes-database] to recommend constructive preventive steps to be taken, the Diabot architecture is made for advanced diabetes prediction [73]. There are several classification algorithms in Machine Learning for prediction, which can be used based on their accuracy. However, the researchers have used Ensemble learning, which is a meta-algorithm that combines a multitude of weaker models and averages them to generate a balanced and precise final model, rather than considering just one model.

The framework has a front-end User Interface (UI) at a high level for the patient to converse with the bot. It is designed using HTML and JavaScript using the React UI framework [73]. The chatbot communicates via API calls with the NLU engine at the backend. Using the RASA NLU platform, the NLU engine was created. Two models are trained at the backend using two datasets, one for generic health diagnostics and another one aimed for advanced diabetes forecasting, providing the NLU engine with the appropriate diagnosis decision, which is trained to provide performance based on user queries [73].

Tess

In only two to four weeks of engagement, psychological AI services and mental health chatbots were demonstrated to substantially reduce symptoms of depression and anxiety [74]. Tess [https://www.x2ai.com/], a personalized mental health chatbot created by X2AI Inc. (X2), was motivated by the ability to support thousands of people around the world at the same time [74].

Accessing Tess support is convenient through existing communication platforms such as text messaging (SMS) and Facebook Messenger, and voice-enabled services can be incorporated with Amazon Alexa/Google Home. Digital solutions provide the added advantage of seamlessly connecting with existing services and educational courses, such as Elizz Caregiver in the Workplace program at SE Health, in addition to scalability. To support a range of features, Tess is built using a combination of technology, emotion algorithms, and machine learning techniques [74]. Tess is qualified to provide prescripted interventions in conjunction with mental health practitioners aiming to replicate an empathic reaction that is suitable to the input emotion or scenario. Based on the individual's expressed concern, particular measures are activated. Tess may offer a solution to help them achieve a more comfortable state, or triage them to a suitable resource, if anyone indicates that they feel anxious. In order to de-escalate the situation, Tess should provide a simple risk assessment and send a crisis warning via SMS to a crisis center or a counselor to advise them to take over the conversation.

FeelFit

FeelFit [https://www.wiwi.uni-osnabrueck.de/feelfit] tracks users' vital parameters with sensors from different suppliers, consolidates and incorporates sensor data, and outputs it via a Conversational Interface (CI) [75]. It is constructed as a modular device that includes numerous components and thus makes a user-centered configuration. Natural language input and output of data, display options on a smart mirror, and connectivity with other mobile devices are the components of the system [75].

Because people can use different devices to converse with Conversational Agents (CAs), it's critical to categorize these devices and determine which will be used for which agent [75]. A valid method for classifying devices is the categorization of input modalities into written and spoken inputs. Written inputs may, on the one hand, use visual-only devices that, for example, use touch or remote input but do not enable voice input, and intermodal devices, such as smartphones, allow voice or text inputs. Speech-only devices, on the other hand, only accept voice input and do not respond to touch or written input. Non-input devices still exist, which only present information to the user.

Both of these devices should be enabled in order for the health monitoring system to provide the patient with the appropriate details at any time. Each system allows sensor contact, such as a body temperature thermometer, a weight scale, and a blood pressure monitor for systolic and dysostosic blood pressure readings [75]. In addition to direct communication between a sensor and a computer through WLAN or Bluetooth, indirect communication must be provided *via* third-party web services, as some vendors allow sensor readings to be used only by their applications. The vendors offer APIs with appropriate authentication mechanisms, such as OAuth 2.0.0, to access these sensor data. Although it is possible to collect sensor data for each input and output unit, FeelFit uses a data collection smartphone since it is the common mobile device in everyday use and it has high processing power and rich functionality [75].

It is important to make the data accessible to the other devices after it has been processed. As a result, REST web services on a cloud coordinate data transmission to all computers. Since the system is customized and includes confidential data, it is important to introduce an authentication service to access the REST services, guaranteeing that only approved users can access their own data [75]. All data is stored in a NoSQL database due to its ability to store high-frequency and heterogeneous data. Without making major adjustments to the data structure, the sensors' data can be kept in the user's profile.

Interaction with the CA through visual-only, voice-only, or multimodal, or devices can send the data to a NLP service for speech recognition, spoken language comprehension, and dialogue management [75]. Agent-specific configuration is required for the latter two activities, such as training of agent-specific classifications and conversation flows. A REST web service will be used to produce a response after the input has been processed, and the NLP service sends the response back to the computers [75]. A text-to-speech synthesis will present the answer as spoken language for spoken input. Furthermore, some inputs which cause actions (such as showing graphs on a non-input computer), causing the web service to send information after language processing to another device [75].

KR-DS

Xu et al. proposed an End-to-End Knowledge-Routed Relational Dialogue System (KR-DS) that integrates medical knowledge into the subject transformation in dialogue management and integrates it with natural language comprehension and generation [76]. To handle topic transitions, a novel Knowledge-Routed Deep Q-Network (KR-DQN) is implemented, which incorporates a relational refinement for encoding relationships between different pairs of symptoms and symptom disease, for decision-making [76].

The End-to-End Knowledge-routed Relational Dialogue System (KR-DS) involves Natural Language Understanding (NLU), Dialogue Management (DM) and Natural Language Creation as a task-oriented dialogue system (NLG) [76]. From utterances, NLU recognizes user intent and slot values. DM then performs the subject transformation according to the current state of the dialog. In sDM, the agent learns to request symptoms in order to continue with the diagnostic task and to notify the illness in order to make a diagnosis. Natural language sentences are generated by a template-based NLG, provided the predicted system behavior. A user simulator executes the conversation exchange conditioned on the created user target in order to train the entire system in an end-to-end manner through reinforcement learning.

Since their dataset was obtained from a Chinese website, Xu et al. used NLU to identify purpose and fill slots for the Chinese language [76]. To improve accuracy, they used a public Chinese segmentation tool and added medical terminology to a custom dictionary. NLU classifies purpose and fills in a collection of slots to shape a semantic frame given a word sequence. Semantic frames are organized data containing the intentions and slots of the user. In their automatic diagnosis dialogue system, they take into account six different types of intents. There are four kinds of intentions for a user, including request and disease, confirm and symptom, deny and symptom, and not-sure and symptom [76]. To mark each word tag in the sentence, the authors have applied the common BIO format.

Xu et al. have introduced a Bi-LSTM to identify each word's BIO tags and to classify this sentence's intent at the same time [76]. They filled slots with tag labeling based on contextual knowledge from conversation and normalization of medical terminology. The annotators' medical terminology are used to normalize symptoms and diseases. The authors maintained a rule-based dialogue state tracker with regard to context awareness, which stores the status of symptoms. To construct a new symptom vector, the authors interpreted slots in the current semantic frame as a symptom vector and added it to the previous symptom vector [76]. If a symptom is requested by an agent and does not appear in the response of the user, referring to the reported request space, this approach may also fill the slot.

Supervised learning is used to train the bi-directional LSTM model as symptoms and intents are labeled in the dataset of Xu et al. [76]. Furthermore, after pre-training, reinforcement learning can be used to train NLU in combination with other aspects of their KR-DS.

iHelpr

The iHelpr chatbot [https://www.inspiresupporthub.org/index] offers guided self-assessment on stress, anxiety, depression, sleep, and self-esteem, among other topics [77]. It is a part of an online self-help portal called the Inspire Support Hub [77].

Initially, iHelpr allows the user to complete a tool for self-assessment based on the choice they have selected. Based on the results of the self-assessment survey, personalized advice with evidence-based feedback is provided to the customer. Links to additional support literature and suggested e-learning services are included in the recommendations. The user is given helpline numbers and, if appropriate, emergency contact information if there is an elevated risk based on higher scores.

The Microsoft Bot Framework and NodeJS were used to build the iHelpr chatbot [77]. It is related to a MySQL database containing coping mechanisms and scores of questionnaires. The Language Understanding Intelligent Service (LUIS) of Microsoft was implemented to identify user utterances and adapt them to the correct intent. and NodeJS were used to build the iHelpr chatbot [77]. It is related to a MySQL database containing coping mechanisms and scores of questionnaires [77]. In order to identify the utterances made by users and to adapt them to the correct purpose, Microsoft's Language Understanding Intelligent Service (LUIS) was implemented.

A clinical psychologist planned the conversation flow. Conversation scripts were produced and perfected with the psychologists over many iterations to ensure they were fit for purpose [77]. The conversation was then pasted into the prototyping tool Botsociety, which allowed the conversation flow to be visualized in a conversational GUI. In some places, but mostly through short responses, the user can communicate with the chatbot using free text. Users do not send GIFs or emojis to the chatbot, but inside the talk, GIFs are used.

AAC

An Alternative Communication (AAC) app with an embedded chatbot known as Alex has been created by Cooper et al. [78]. Alex is designed for use on the autism spectrum by individuals [78]. Programming Alex does not require any specialized skills and is structured to deliver what they consider important to speech therapists, parents and other main stakeholders [78]. In a safe, non-judgmental environment, the consumer is able to practice spontaneous conversation with Alex.

Cooper et al. built a method focused on visual symbols because aided AAC is the most powerful type of AAC for children on the autism spectrum [78]. A grid display or a Visual Scene Display is used by aided AAC systems to represent symbols (VSD) [78]. Grid display is a typical AAC system layout in which symbols are positioned in unique grid positions [78]. It promotes the recall of the individual contents of each cell as each symbol is isolated. Since the grid can be stacked to provide more symbol access, a grid display allows for a broad vocabulary. Within a naturalistic scene, VSD embeds symbols, providing meaning for their use. It supports a smaller, context-specific vocabulary that can be generalized, as the symbols displayed are necessarily related to the context of the scene [78]. A grid display was chosen because it offers a wider vocabulary that is suitable for unstructured conversation. A 4 by 10 grid is the default layout, which can be changed depending on the content of the grid [78]. While research supports the use of animated symbols for AAC by children, these studies do not discuss the need to reduce the AAC display's visual-cognitive load or involve participants on the autism spectrum [78]. As a consequence, the first version of the application was developed with static symbols.

Grid layouts are typically organized by operation, taxonomy, theme or type of term, with the layout's effectiveness varying depending on the individual [78]. As it is normally best understood by younger children and better associated with visuals than the word types themselves, the theme layout was chosen. The Open Board Format is used to store information related to symbol layouts. As a potential standard for digital AAC boards, Open Board Format files (.obf and .obz) have been implemented to ensure the interoperability of communication boards between devices [78]. Based on the common JavaScript Object Notation (JSON) format, the.obf files contain details about the symbols and their corresponding images themselves. The.obz files are zips of the.obf files that contain their image files as well as a manifest file that explains how the boards are linked. Users can import and export boards in this style. In order to personalize the layout and vocabulary of the grid, a high degree of customization is built into the app. It is possible to import or build custom boards and change the dimensions of the grid layout to fit this material. To provide user-specific terminology, symbols can be customized, rearranged, and inserted. A robust search function is included so that the user can search for symbols to add to the board or be used in general conversation in the inbuilt symbol dictionary of over 15,000 symbols [78].

MMDAgent

Tanaka et al. have developed an automated social skills trainer from MMDAgent, a Japanese-voiced dialogue system that aggregates dialogue management, text-to-speech, speech recognition, and behavior generation [79]. Defining target skills and describing their priorities are all part of the training phase. The skills to be learned are calculated based on these problems after recognizing the major social problems faced by the trainee [79]. Trainers act as a model to show the potential the users are working on so that before attempting it themselves, they can see what they need to do [79]. By encouraging users to watch a documented model video of people with reasonably strong narrative skills, the automated social skills trainer replicates this move for narrative skills. Users will watch those good examples and mimic them. Participants role-play for the teacher their interactions. In the target case, this helps them to exercise their own talents. Trainers concentrate on voice tone, amplitude, facial expressions, eye gaze, and other nonverbal actions while observing the participants' social skills. Tanaka et al. have improved on this step by providing audiovisual details [79]. Using a restricted local model-based face tracker, they extracted a number of facial features to analyze the video's data. Using the NOCOA and database, the authors reported that the model can be extended to other speakers and videos.

HARLIE

Human And Robot Language Interaction Experiment (HARLIE) [https://apkpure.com/de/harlie-the-e-health-chatbot/org.harlie.chatbot] has been created by Ireland and colleagues [80]. It is an Android-based chatbot that analyzes the user's speech and language articulation when conversing about various topics.

Both speech-to-text and text-to-speech tools are included in Android and iPhone smartphones. The acoustic signals can be transformed into digital text, and digital text can be converted into a digital synthetic, acoustic voice. These tasks are carried out by HARLIE using Google's speech-to-text and text-to-speech application programming index (API) [80]. Since this necessitates sending acoustic audio to an off-shore server, the signal is subjected to random voice modulation before being sent. This means that the user's speech pattern cannot be used to identify them if they are intercepted.

The use of a standardized programming language such as Artificial Intelligence Mark-up Language, as well as case-based inference and textual pattern matching algorithms, have become the most commonly used techniques for text input processing (AIML) [80]. HARLIE employs AIML to converse with a consumer in a non-deterministic and substantive manner. The digital brain is made up of AIML files that cover a broad variety of subjects, circumstances, and speech tasks.

HARLIE actively analyzes aspects of the health of users' voice and communication during the conversation [80]. This includes the consistency of vowels, vocabulary selection, and the length of mid-sentence pauses. People who are working to develop their speech or communication as a result of difficulties caused by a neurological disorder such as Parkinson's disease or stroke may need to practice and obtain input on a regular basis [80].

Chester

Allen et al. created Chester, a prototype intelligent assistant that communicates with its user using conversational natural spoken language to inform and provide guidance about their prescribed drugs [81]. The components are divided into three categories: perception, actions, and generation, and the information sources are frequently shared among them. The interpretation components are responsible for deciphering what the user has said or done, while the action components are in charge of the system's goals and rationale, and the generation components are in charge of the system's contributions to the discussion.

Speech recognition is the first stage of interpretation, and it draws its vocabulary from the system's common lexicon to create a list of potential interpretations for the Parser to process [81]. The Parser generates a sense representation represented in a domain-independent semantic representation using a general lexicon and grammar of spoken dialogue utterances. The Parser output is transferred to the interpretation manager, who performs contextual interpretation. It communicates with the Task Manager, who conducts plan and purpose identification, to address referring expressions such as pronouns and definite noun phrases. It also uses a collection of ontology mapping rules to translate the generic semantic representation into a Chester-specific representation. It updates the Discourse Context by defining the most likely planned speech act in the form of collective problem-solving intervention.

The Behavioral Agent is the agent's self-contained heart [81]. It takes the interpretation manager's collective problem-solving act and conducts additional aim recognition before planning system activity based on its own priorities and responsibilities, the user's utterances and acts, and shifts in the world state. The Task Manager manages activities that involve task- and domain-specific processing [81]. In the form of communicative acts, activities that include contact and coordination with the user are sent to the Generation Manager. Chester features a drug knowledge base and a rudimentary scheduling algorithm for thinking about medications, among other domain-specific components. The bulk of the work we had to do to develop the framework was based on these domain-specific aspects.

The Generation Manager is in charge of managing system response preparation [81]. Content preparation, where the utterance is formulated, and surface realization, where spoken words are constructed, are two forms of generation. Discourse responsibilities (stored in the Discourse Context) and the directives it receives from the Behavioral Agent influence its actions. The glue that binds the layers together is an abstract model of problem-solving that can convey both user and machine contributions to the collaborative mission.

Twin-robot dialogue system

Iio et al. built a twin-robot dialogue system that incorporates the question–answer–response dialogue model with the involvement of two robots in a conversation [82]. The sounds are recorded using a microphone collection. A microphone array combines the sounds *via* a noise reduction process. Following that, the incorporated sound is sent to the automatic speech recognition module, which recognizes the user's utterance. NTT Docomo's cloud speech recognition service was used by Iio et al. [82]. The service receives a voice and sends the results of voice recognition to the utterance collection module. The utterance collection module selects an utterance from the database based on the selection rules. The robot controllers receive the selected utterance [82]. As a signal of user speech, the voice recognition results are also sent to the nodding generation module. The robot controllers receive a nodding motion from the nodding generation module. In the nodding generation module, nodding is a motion that communicates that the robot is listening to the user. When the machine receives a speech recognition result, this motion is always performed in the answer state [82]. The robot controllers translate the utterance into gestures, which they then carry out. The completion signal is sent to the utterance selection module once the execution is complete. The next utterance is selected by the utterance selection module. As a result, the machine selects and executes an utterance based on the effects of speech recognition and its own action execution.

Miraculous-Life

The Miraculous-Life dialogue system developed by Merdivan et al. consists of several parts with each part specializing in certain tasks [24]. With each part specializing in certain tasks, the dialogue is made up of several parts.

The Automatic Speech Recognizer (ASR) module is responsible for translating the utterances of spoken users into text. The Natural Language Interpreter translates textual information to meaningful features so that these features can be processed and the current dialogue state is modified by the dialogue State Tracker (DST). DST outputs the current dialogue state to enable the dialogue Response Selection (DRS) module to generate a user's textual response. The DRS module is trained to produce a user utterance response. This textual reply is later translated by the Text To Speech (TTS) synthesizer into speech. As ASR and TTS are not explicitly linked to the dialogue manager, they can be viewed as complementary modules for the dialogue system.

A dialogue manager is a component of a dialogue system that is responsible for transmitting information between participants in the interaction between human and machines. Merdivan et al. concentrated on the chatbot form of dialogue system and on how they are trained in manager models based on rule-based, sequence-to-sequence learning, reinforcement learning and hierarchical reinforcement learning [24]. To get the benefit of each other's strengths, all these strategies can be implemented together. Alternatively, they can be implemented independently. In addition, the paper by Merdivan et al. provides findings on the dialogue dataset with a new image-based approach in which dialogue is processed as images to train dialogue manager [24].

ReMindMe

The design and architecture of a framework named ReMindMe was initiated by Peeters [25]. It is an agent-based support system for individuals with Alzheimer's and for their social environment. In two ways, ReMindMe supports patients and their carers. First of all, it helps patients to remember their personal memories, thereby enhancing their sense of identity, protection, security, and self-esteem [25]. Second, ReMindMe promotes self-disclosure, i.e., the sharing of personal memories of patients with the associated caregivers. Self-disclosure is necessary for the formation and preservation of interpersonal relationships [25]. Self-disclosure of personal memories between patients and caregivers allows caregivers to enhance their care delivery due to an increased understanding of the personal needs of patients and it helps patients build a sense of companionship and acceptance with the caregivers [25]]. Peeters has stated that musical memory remains largely unchanged in Alzheimer's Disease [25]. Therefore musical memory can serve as a helping mechanism to evoke lifetime event memories. In many cases such memory evokes can not be accomplished through a verbal pathway [25].

Roberta

The need for human-robot interaction with multimodal communication channels was emphasized by Sansen et al. [83]. For easy contact with elderly people, they introduced the intelligent robot conversational agent as a human-sized humanoid robot named Roberta in 2016, which sits on an electric wheelchair to act as a cognitive coach and physical coach for dependent individuals [83]. Two positions are available for the electric wheelchair, namely sitting and standing. The computing architecture of the robot consists of multiple minicomputers dedicated to specific tasks and arranged in a tree structure. Roberta is equipped with a stereoscopic camera that is used during the dialogue as a standard camera for facial recognition. Text-to-speech process uses the beam-forming technique of an 8 microphone array. Roberta is also equipped with an expressive face and has two arms used for non-verbal communication and to capture small items on the shelves for individuals for which robots are working [83].

Roberta obtains information from the world and from the system. Through a dialogue manager it incorporates this knowledge into the management of dialogue and of strategies [83], In order to help elderly people to exercise their speech and memory skills, it is built to help people tell stories about their lives [83]. It also aims at gaining knowledge of the person. The framework, on the other hand, is built to feature an open-domain conversational system that can provide the user with useful and interesting information [83]. Roberta analyzes the actions of the user and, based on that, infer the level of emotion and interest of the user in order to adjust their presentation accordingly [83].

Virtual coach for active ageing

In 2014, Callejas et al. created a conversational agent that serves as a physical activities coach [26]. In order to provide practical coaching, the agent can be built as an Android app running on smartphones and can be combined with inexpensive, readily available sports sensors. Two types of sensors are used in the architecture proposed by Callejas et al. [26]. The first sort includes sensors that are already built into the coach's smartphone. This includes the microphone and Global Positioning System (GPS) sensor. The second category contains sensors that are external to the smartphone, for example pulsometer and pedometer [26]. Biosignals are calculated by these sensors: pulsometers provide heart rate and skin conductivity information, while pedometers count the steps while walking [26].

The conversational agent developed by Callejas et al. may be used to schedule exercise sessions, provide input during the sessions, and to analyze the outcomes after the exercise [26]. The smart conversational agent for coaching elderly people in physical activity has been introduced.

There are following three knowledge sources in the proposal of Callejas et al.: the user model, training model, and interaction model [26]. The user model includes static information (name, gender and age) and information modified following the interactions. The information on the exercise session deals primarily with the progression of the user in the exercises and the influence of the user in relation to this development. The training model deals with the appropriate techniques for the coach according to the user's physical condition and the predicted user activity. Callejas et al. developed a basic polarity and arousal model (positive vs. negative effect and percentage intensity level) with regard to the affective model, which is combined with the details of the exercise session to which it relates. The interaction history is stored as logs, including date, mode of interaction, sensor information log, and sensor communication error log. In addition, interaction parameters are monitored when there is a conversation with the agent including dialogue length, a number of user turns, a number of device turns, the output of speech recognizer knowledge, and success rates of speech understanding.

BeWell

Lane et al. created the BeWell mobile wellness applications, which track user behavior in three areas of health: sleep, physical activity, and social interaction [84]. These apps inspire better actions by offering input in the form of an ambient display on the smartphone's wallpaper.

BeWell and is made up of two software components: a phone app and cloud infrastructure [84]. The accelerometer and microphone sensors on the phone are used by the BeWell and phone app to automatically track the user's daily activities. The BeWell and Cloud infrastructure receives the inference results from the phone's classifiers. All of the data is processed in the cloud, and health scores are calculated there. Based on inferred behavioral trends, wellness ratings summarize the effect on overall health. BeWell and measures wellness ratings for each of the health indicators it controls. Physical activity, sleep cycles, and social contact are included in the latest prototype [84]. The BeWell phone app uses an ambient display rendered on the device's wallpaper to show these scores to users on the phone.

Vastenburg et al. conversational agent

Vastenburg et al. produced a conversational agent in 2008 that serves as a physical activity coach and helps elderly people enhance their well-being [27]. It can be used to schedule exercise sessions and provide input during the sessions [27]. The agent can be built as an Android app running on smartphones. To create a message-based conversation system, to prepare elderly people for physical activity and to create a measuring system for user activities, they have introduced the intelligent conversational agent [27].

The smart conversational agent is used to create a simple communication product for elderly users that is easy to use. The resulting design principle is based on a known metaphor and seeks to lower the barrier for contact across the care network and family and friends at home to generate direct value for the end user [27]. Message-based communication using the Internet may help preserve or expand the social network for older people with little or no previous experience with computers. In addition, a message display, such as medicine reminders, may be used to show device-generated messages [27]. In order to incorporate the picture frame-sized touch screen, the intelligent conversational agent is used to show up-to-date details on old people's physical activities [27]. The goal of developing a labeling mechanism for elderly people's user activities is to track living habits remotely, to identify unusual circumstances, and to offer routine help to actual activities [27]. Recognition of behavior is focused on detecting recurrent trends in sensor data. Activity recognition algorithms need to be trained using sensor data with activity labels to be able to automatically identify events. Because the configuration of the sensor and user activities vary between individuals and locations, for each new environment, algorithms need to be trained [27].

Intelligent visual companion framework for independent living

Todd and Sasi have developed an intelligent conversational agent that helps elderly people by initiating casual conversation, providing clues for performing daily living tasks, tracking elderly peoples' well-being, and performing secretarial tasks [85]. In Todd and Sasi's research work, an Intelligent Visual Companion Framework for Independent Living (IVCSIL) is developed to meet the social and companionship needs in a smart home and in an independent living house. Based on *if – then* rules, an AI system is built by Todd and Sasi [85]. In the figure below, the functionalities are shown using use cases. A comprehensive schedule for a person is entered as a text by the caretaker, and this is transmitted to the end user using a text-to-voice interface. A task manager is designed to define and to delegate task goals [85]. The conversation initiation is embedded in the scheduler [85]. All other functionalities are incorporated as subsystems, including fitness, entertainment, and general living functionalities [85]. The next course of action of the system and anticipated responses from the user would be decided by the conditions attached to the task calls with the individual features [85]. The tasks are inserted via the keyboard and these texts are translated into voice by the conversational agent [85].

The summary of AI and conversational agents dealing with a cognitive disability is presented in Table 3.

Speech recognition datasets for chatbot training and testing

Machine learning methods dealing with Automatic Speech Recognition (ASR) need datasets to train the chatbots and to test the DMSs. Below are some of the prominent general voice recognition datasets.

The speech accent archive [https://www.kaggle.com/rtatman/speech-accent-archive] was created to exhibit a vast number of different speech accents from various languages. As a result, there are 2,140 English speech samples in the dataset, each from a different speaker reading the same passage. Furthermore, participants hail from 177 different countries and speak a total of 214 different languages.

The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) [https://www.kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio] contains 24 professional actors (12 female and 12 male), all of whom are saying the same thing. Not only that, but the speech emotions captured at two stages of strength include calm, happy, sad, angry, afraid, surprise, and disgust.

The TED-LIUM corpus [https://www.openslr.org/51] is a collection of TED talks and their transcriptions, which can be found on the TED website. There are 2,351 audio samples in total, totaling 452 h of audio. In addition, the dataset includes 2,351 STM-formatted compatible automated transcripts.

The Google Audioset dataset [https://research.google.com/audioset/] includes an ontology of 635 audio event groups and over 2 million 10-s sound clips from YouTube videos. Furthermore, human labelers were used to add metadata, meaning, and content analysis.

Over 1,000 h of English speech extracted from audiobooks can be found in the LibriSpeech ASR Corpus [https://www.openslr.org/12]. The majority of the recordings are based on Project Gutenberg texts.

The aim of the Gender Recognition by Voice database [https://www.kaggle.com/primaryobjects/voicegender] is to assist systems in determining if a voice is male or female based on acoustic and speech properties. As a result, the dataset includes over 3,000 speech samples from both male and female speakers.

Hundreds of thousands of speech samples for voice recognition can be found in the Common Voice dataset [https://www.kaggle.com/mozillaorg/common-voice]. Over 500 h of speech recordings are included, as well as speaker demographics. User-submitted blog entries, old movies, books, and other public speech were used to build the corpus.

The VoxCeleb dataset [https://www.robots.ox.ac.uk/vgg/data/voxceleb/] is a large-scale speaker recognition dataset of over 100,000 phrases spoken by 1,251 celebrities. VoxCeleb, like the previous datasets, contains a wide variety of languages, occupations, and ages.

The TensorFlow and AIY teams collaborated to build the Google Speech Commands dataset [http://download.tensorflow.org/data/speechcommandsv0.02.tar.gz]. There are 65,000 clips in this dataset, each lasting one second. Each clip contains one of the 30 different voice commands delivered by tens of thousands of people.

Small speech samples make up the Synthetic Speech Commands dataset [https://www.kaggle.com/jbuchner/synthetic-speech-commands-dataset]. Each file, for example, includes single-word utterances like yes, no, up, down, on, off, stop, and go.

Table 3. Summary of AI and conversational agents dealing cognitive disability.

Author/Year	Ref.	Type	Management	Input	Output	Dealing cognitive disability?
Gaffney et al./2020	[86]	Windows computer app	Finite-state	Written	Written	Yes [mental health of people]
Hauser-Ulrich et al./2020	[54]	Mobile device app	Finite-state	Written	Spoken; written	Yes [physical health people with pain]
Bali et al./2019	[73]	Chatbot	Finite-state	Written	Written	Yes [medical diagnostics of people]
Bott et al./2019	[87]	ECA	Finite-state	Spoken	Spoken	Yes [old adult people]
Hernandez/2019	[88]	Chatbot	Finite-state	Written	Written	Yes [old adult people]
Joerin et al./2019	[74]	Chatbot	Finite-state	Written	Written	Yes [mental health of people]
Jungmann et al./2019	[89]	Mobile device app	Finite-state	Written	Written	Yes [mental health of people]
Lee et al./2019	[90]	Medical chatbot	Finite-state	Written	Written	Yes [mental health people]
Meier et al./2019	[75]	ECA	Finite-state	Spoken; written	Spoken; written	Yes [health awareness of people]
Xu et al./2019	[76]	Dialogue system	Finite-state	Written	Written	Yes [medical diagnostics of people]
Cameron et al./2018	[77]	Medical chatbot	Finite-state	Written	Written	Yes [mental health of people]
Cooper et al./2018	[78]	Chatbot; ECA	Finite-state	Spoken	Spoken	Yes [autism people]
Divya et al./2018	[91]	Web browser based chat	Finite-state	Written	Written	Yes [medical diagnostics of people]
Rarhi et al./2018	[92]	Medical chatbot	Finite-state	Written	Written	Yes [medical diagnostics of people]
Srivastava et al./2018	[93]	Chatbot	Finite-state	Written	Written	Yes [old adult people]
Chi et al./2017	[94]	Embodied conversational agent	Finite-state	Spoken	Spoken	Yes [old adult people]
Hoermann et al./2017	[28]	Web browser based chat	Finite-state	Written	Written	Yes [mental health of people]
Mujeeb et al./2017	[95]	Medical chatbot	Finite-state	Written	Spoken	Yes [autism people]
Ni et al./2017	[96]	Medical chatbot	Finite-state	Written	Written	Yes [medical diagnostics of people]
Philip et al./2017	[97]	Windows computer app; ECA	Finite-state	Spoken	Spoken	Yes [mental health of people]
Tanaka et al./2017	[79]	Windows computer app; ECA	Finite-state	Spoken	Spoken; written	Yes [autism people]
Ireland et al./2016	[80]	Mobile device app; chatbot	Frame-based	Spoken	Spoken; written	Yes [dimentia people and perkinson's disease people]
Morbini et al./2012	[98]	Dialogue system	Finite-state	Written	Spoken	Yes [mental health of people]
Allen et al./2006	[81]	Windows computer app	Finite-state	Spoken	Spoken	Yes [old adult people]
Levin and Levin/2006	[99]	Telephone	Finite-state	Spoken	Spoken	Yes [physical health of people with pain]
Preininger et al./2020	[100]	Web browser	Frame-based	Written	Written	Yes [old adult people]
Morris et al./2018	[101]	ECA	Frame-based	Written	Written	Yes [mental health of people]
Ireland et al./2016	[80]	Mobile device app; chatbot	Frame-based	Spoken	Spoken; written	Yes [mental health of people]
Wolters et al./2015	[102]	Web browser	Frame-based	Spoken	Spoken; written	Yes [dimentia people]
Beveridge and Fox/2006	[103]	Telephone and web browser	Frame-based	Spoken	Spoken	Yes [old adult people]
Iio et al./2019	[82]	Intelligent conversational agent for robot	Agent-based	Spoken	Spoken; facial gesture from robot	Yes [old adult people]
Merdivan et al./2019	[24]	Intelligent conversational agent	Agent-based	Spoken	Spoken	Yes [old adult people]
Munoz et al./2018	[104]	Conversational agent implementing digital game	Agent-based	Visual domain specific language	Visual domain specific language	Yes [autism people]
Peeters et al./2016	[25]	Intelligent conversational agent	Agent-based	Spoken	Spoken	Yes [mental health of people]
Sansen et al./2016	[83]	Intelligent conversational agent for robot	Agent-based	Spoken; facial expression	Spoken, facial gesture from robot	Yes [old adult people]
Callejas et al./2014	[26]	Intelligent conversational agent	Agent-based	Spoken; facial expression	Spoken, written	Yes [old adult people]
Lane et al./2014	[84]	Intelligent conversational agent	Agent-based	Visual domain specific language	Visual domain specific language	Yes [physical and mental health of people]
Vastenburg et al./2008	[27]	Intelligent conversational agent	Agent-based	Visual domain specific language	Visual domain specific language	Yes [old adult people]
Todd and Sasi/2006	[85]	Intelligent conversational agent	Agent-based	Spoken	Spoken	Yes [old adult people]

Table 4. Summary of datasets.

Dataset	Content	Samples
Speech accent	English speech	2,140
RAVDESS	Emotional voice	24 voices
TED-LIUM	Talks	2,351 (452 h)
Google audioset	Audio events	2 mln 10-sec sound clips
LibriSpeech ASR corpus	Audiobooks	1,000 h
VoiceGender	Male/female voices	3,000
Common voice	Speeches	500 h
VoxCeleb	Spoken phrases	100,000 (spoken by 1,251 people)
Google speech commands	Voice commands	65,000 clips
Synthetic speech com-mands	Single-word utterances	N/A
Fluent speech commands	Utterances	30,000
CHiME-5	Party talks	20 videos (each 2 h long)
HUB5	Telephone conversations	40
CALLHOME	Telephone conversations	120 (each 30 min)

Over 30,000 utterances from approximately 100 speakers make up the Fluent Speech Commands dataset [https://fluent.ai/fluent-speech-commands-a-dataset-for-spoken-language-understanding-research/]. Each.wav file in this dataset contains a single utterance for controlling smart-home appliances or virtual assistants. All audio also includes movement, object, and position labels.

The CHiME-5 dataset [http://spandh.dcs.shef.ac.uk/chimechallenge/CHiME5/data.html] contains videos from 20 different dinner parties held in real homes. Each file is at least 2 h long and contains audio from the kitchen, living room, and dining room.

The HUB5 English Evaluation Transcripts (HUB5) [https://catalog.ldc.upenn.edu/LDC2002T43] are transcripts of 40 English telephone conversations from the year 2000. Conversational speech over the phone is the subject of the HUB5 assessment sequence, with the task of transcribing conversational speech into text.

The CALLHOME American English Speech dataset [https://catalog.ldc.upenn.edu/LDC97S42] contains 120 unscripted 30-min telephone conversations in English. Because of the study's circumstances, the majority of participants called family or close friends. Speech synthesis is the process of synthesizing human speech artificially. Text-to-speech, speech-enabled interfaces, navigation systems, speech generation, and accessibility for visually disabled people will all benefit from this machine learning-based technique. There are two types of speech synthesis methods existing, namely: concatenative method and parametric method. Speeches from a large dataset are concatenated to create new, audible speech in the concatenative method. A voice database is used when a particular style of speech is needed. This limits the approach's scalability. A recorded human voice and a feature with a set of parameters that can be modified to adjust the voice are used in the parametric method.

The speech recognition datasets are summarized in Table 4.

Evaluation and discussion

Principal findings

In health related applications, the conversational agents lag behind those in other areas (such as travel statistics, restaurant collection, and booking), where conversation management and natural language production techniques have improved beyond rule-based approaches. For simple and well organized tasks, rule-based approaches used in finite-state dialogue management systems are simple to build. However, they have the disadvantages of limiting user input to preset words and phrases, preventing the user from taking action in the conversation, and finding it impossible to correct inappropriately understood items. Frame-based frameworks partly alleviate the limitations of finite-state dialogue management, allowing for structure and mixed initiative, as well as more open dialogue.

Both methodologies are capable of completing tasks by requesting data from user-completed forms. The only distinction is that frame-based systems don't need the appropriate fields to be filled in a certain order, allowing the user to have more detail than the system's query needs. The interviewing agent keeps track of what information is required and asks the appropriate questions.

Unlike finite-state and frame-based systems, agent-based systems can accommodate dynamic dialogues in which the user can start and lead the discussion. Typically, agent-based dialogue management strategies employ computational models trained on real human-machine dialogue data, resulting in increased speech comprehension, performance, scalability, and adaptability.

Recent advances in deep learning and neural networks have sparked the creation of more sophisticated and efficient agent-based conversational agents [105]. Agent-based conversation management is expected to become more commonly used as massive health databases (including patient-generated data obtained from mobile sensors and wireless tracking devices) and deep learning techniques become more widely available for health applications.

Conversational agents' use in the healthcare environment to automate activities may increase as they become more competent and trustworthy, and their use should be checked systematically and on a regular basis. Automation's human performance factors can pose significant security threats, depending on the level of automation and the type of automated process. As a result, in the area of healthcare, it is important to use unrestricted natural language processing (NLP) technologies and other artificial intelligence implementations with conversational agents with caution.

Chatbot support for aged people and people with disability

Majority of the intelligent conversational agents using AI-based and agent-based dialogue management system work on simple tasks including physical coaching, recalling/sharing personal memories, and helping in daily life.

We found in the systematic literature review that only 8 conversational agents are using AI-based and agent-based dialogue systems that are dealing with disability of and helplessness of old people and disability of and helplessness of people with cognitive disorder including autism, dementia and Parkinson's disease.

- 6 conversational agents using AI-based and agent-based dialogue systems are dealing with disability of and helplessness of old people.

- 2 conversational agents using AI-based and agent-based dialogue systems are dealing with disability of and helplessness of people with cognitive disorder.

All the intelligent conversational agents using AI-based and agent-based dialogue management system work on simple tasks including physical coaching, recalling/sharing personal memories, and helping in daily life.

Through the systematic literature review following research gaps were found in the existing research on the conversational agents that are using AI-based and agent-based dialogue systems are either dealing with disability of and helplessness of old people or dealing with disability of and helplessness of people with cognitive disorder including autism, dementia and Parkinson's disease. There exists not combined conversational agents that are dealing with all aspects of the following:

- Disability of and helplessness of old people.
- Disability of and helplessness of people with cognitive disorder including autism, dementia and Parkinson's disease.
- Multiple communication media including spoken and written and non-verbal and visual communication media of input and output.
- Context-awareness and adaptiveness in the implementation of the conversational agents.

The central research question to be addressed and solved is the following:

1. Analysing, designing, and developing the general framework enabling the conversational agents dealing with disability of cognitive disorders in communication and autism using the AI-based and agent-based dialogue management system.
2. Testing the developed general framework enabling the conversational agents dealing with disability of cognitive disorders in communication and autism using the AI-based and agent-based dialogue management system.

Desired characteristics of AI-based and agent-based conversational agents

The AI-based and agent-based conversational agents should have the following characteristics:

- Input from old people and from people of mental helplessness should be voice input and text input. The text input should be directly fed into the AI-based and agent-based conversational agent. The voice should be processed by the Automatic Speech Recognition (ASR) method and will then be fed into the AI-based and agent-based conversational agent.
- Input from health professionals should be voice input and text input. The text input should be directly fed into the AI-based and agent-based conversational agent. The voice should be processed by the Automatic Speech Recognition (ASR) method and will then be fed into the AI-based and agent-based conversational agent.
- The output of the AI-based and agent-based conversational agent should be in the form of voice output and of text output. The voice output will be done by the speech synthesis method.
- The input and the output of the AI-based and agent-based conversational agent should be stored and should be processed using the Artificial Intelligence Markup Language (AIML).
- The health-related information of old people and of people of mental health problems should be stored in the form of a natural language (NL) database. This NL database should be processed by the machine learning method.

Limitations and dealing with bias

This comprehensive literature review has some advantages as well as several drawbacks. It was carried out and stated in accordance with the PRISMA guidelines. From 2010 to 2020, we performed an exhaustive literature search using five bibliographic databases and a detailed and systematic search approach.

In order to prevent losing critical studies and create a systematic understanding of AI-based conversational agents for health care for mental disorders, we prioritized sensitivity over accuracy in our search strategy. The research eligibility requirements were specified scientifically.

Three reviewers worked independently on study collection, title and summary screening, full-text screening, and data extraction. At several points during the selection process, we tested for inter-rater compatibility, and Cohen kappa showed significant agreement at each stage. However, the final paper selection might still have omitted relevant AI-based conversational agents.

The variability and a limited number of included studies, as well as the prevalence of quasi-experimental studies, are major shortcomings of this study. This emphasizes the search field's ambiguity and novelty.

Finally, the probability of bias ranged greatly among the studies included, lowering the credibility of results among those with a high risk of bias. This made it more difficult to trust the results of experiments with a high chance of prejudice.

The problem of uncanny valley

With artificial agents provoking influential physical composition and sensory expressions, the relations in modern HMI transactive engagements are argued as supportive of the design and development in natural language processing – one of the key components of any conversational agent [106]. In this section, we aim to discuss the problem of the uncanny valley, a hypothesis raised by prof. Mori [107] that to rephrase the original robotic narrative – AI-driven human machine interaction can become like a live human agent, and similar to the original Moshi Mori's scenario, if the system is not functioning on the ideal level – the initial positive responses of interaction might dramatically drop into the negatives [108]. It is clear that users would report discomfort, likeability and eeriness [109], and even disgust, leading to what could be described as realism inconsistency [110]. And on the other hand – having a human-like system would be to raise the human evaluation back to the positive and to trust the interface [111]. While this is the aim of many researchers, designing the UIs and conversational agents, little empirical evaluation has been made in recent years, yet the interest is raising in this age of smart assistants, with unprecedented level of spoken interactions with machines through Alexa, Siri, Cortana, etc. agents, even if there are no direct counterparts of these for people with special needs, where such speech based assistants can create perceptual tension and negative experiences due to the conflicting stimuli of computer speech and "humanlike" language [112]. In fact, it was shown that in addition to adapting to the quality of speech synthesis [113], understanding the user's side may be crucial for designing better chatbots in the future and, thus, can contribute to advancing the field of human-computer interaction [114]. Many hypotheses have suggested that the uncanny valley would be caused by artificial–human categorization difficulty or by a perceptual mismatch between artificial and human features [115]. Like in robotics, HMI researchers still need to draw an accurate map of the uncanny valley, so they could understand what makes the AI conversational agent human, identify a necessity to create using nonhuman designs, to develop interfaces to which people with special needs would relate comfortably, which would understand them and offer information in a supportive and adaptive manner. Researchers found that designs affect emotions [116] there remains a remarkable influence of the field of expertise of the probands, which leads to our new speculative hypothesis that the "technophilic attitude" of a significant part of the problems covered and superseded their primary affinity towards their artificial partner in conversation [117]. Spreading VR and AR-based solutions does not seem like a panacea – it was shown that virtual characters were often regarded as more uncanny (less familiar and human-like) than humans and that increasing levels of asynchrony increased perception of uncanniness was always noticeable [118], similar to that of video games [119].

Conclusion

This study presented a detailed analysis of previous research on the use of conversational systems in the assessment and evaluation of cognitive disorders. Detection of psychiatric illness currently has limitations such as late detection, which has prompted the science community to innovate and investigate new detection methods based on natural language conversations between a patient and an AI-assisted conversational agent.

Artificial intelligence-based conversational agents are becoming more popular in healthcare settings thanks to technological advancements. Despite their medical prevalence and economic strain on 21st-century healthcare systems, this emerging area of research has a small range of applications targeted for chronic conditions.

There is a shortage of evidence-based assessment and comparisons inside and between chronic health problems in the existing applications published in the literature. Future studies should concentrate on the following evaluation and reporting recommendations for technical issues including the underlying AI design and overall solution evaluation.

The future scope for conversational agents is quite strong in the domain of self-management. The agents will evaluate symptoms that are given to them as various inputs, report back the health monitoring outputs, and suggest a course of action based on these varied inputs.

Acknowledgments

First and first, we would want to thank Allah the Almighty, the Most Gracious, and the Most Merciful for the blessings bestowed upon us and our helpful President reviewers, experts, and scholars, during our extensive research and completion of this article. Allah blesses His last Prophet Muhammad (peace be upon him), his family, and his associates. No funding was used for this study.

Disclosure statement

No potential conflict of interest was reported by the author(s). All authors have agreed to the submission of this journal.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Robertas Damaševičius http://orcid.org/0000-0001-9990-1084

References

[1] Cuzzolin F, Morelli A, Cîrstea B, et al. Knowing me, knowing you: theory of mind in AI. Psychol Med. 2020;50(7):1057–1061.
[2] Trewin S, Basson S, Muller M, et al. Considerations for AI fairness for people with disabilities. AI Matters. 2019;5(3):40–63.
[3] Rong G, Mendez A, Bou Assi E, et al. Artificial intelligence in healthcare: review and prediction case studies. Engineering. 2020;6(3):291–301.
[4] Tudor Car L, Dhinagaran DA, Kyaw BM, et al. Agents in health care: scoping review and conceptual analysis. J Med Internet Res. 2020;22(8):e17158.
[5] Powell J. Trust me, I'm a chatbot: how artificial intelligence in health care fails the Turing test. J Med Internet Res. 2019;21(10):e16222.
[6] Masche J, Le N-T. A review of technologies for conversational systems. In Advanced computational methods for knowledge engineering. Cham: Springer International Publishing; 2017. p. 212–225.
[7] Parmar P, Ryu J, Pandya S, et al. Health-focused conversational agents in person-centered care: a review of apps. NPJ Digit Med. 2022;5(1):21.
[8] Robbertz AL, Winsor CC, Valente RC, et al. Towards a more inclusive world: enhanced augmentative and alternative communication for people with disabilities using AI and NLP; 2020. Available from: https://digitalcommons.wpi.edu/mqp-all/7230.
[9] Svenningsson N, Faraon M. Artificial intelligence in conversational agents. In Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference, Kobe Japan, AICCC 2019: ACM; 2019.
[10] ter Stal S, Kramer LL, Tabak M, et al. Design features of embodied conversational agents in eHealth: a literature review. Int J Human Comput Stud. 2020;138:102409.
[11] Laranjo L, Dunn AG, Tong HL, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc. 2018;25(9):1248–1258.
[12] Vaidyam AN, Wisniewski H, Halamka JD, et al. Chatbots and conversational agents in mental health: a review of the psychiatric landscape. Can J Psychiatry. 2019;64(7):456–464.
[13] Schachner T, Keller R, V Wangenheim F. Artificial intelligence-based conversational agents for chronic conditions: systematic literature review. J Med Internet Res. 2020;22(9):e20701
[14] Provoost S, Lau HM, Ruwaard J, et al. Embodied conversational agents in clinical psychology: a scoping review. J Med Internet Res. 2017;19(5):e151.
[15] de Cock C, Milne-Ives M, van Velthoven MH, et al. Effectiveness of conversational agents (virtual assistants) in health care: protocol for a systematic review. JMIR Res Protoc. 2020;9(3):e16934.

[16] Maskeliunas R, Damaševičius R, Lethin C, et al. Serious game iDO: Towards better education in dementia care. Information. 2019:10(11). doi:10.3390/info10110355
[17] Tropea P, Schlieter H, Sterpi I, et al. Rehabilitation, the great absentee of virtual coaching in medical care: scoping review. J Med Internet Res. 2019;21(10):e12805.
[18] Gaffney H, Mansell W, Tai S. Conversational agents in the treatment of mental health problems: mixed-method systematic review. JMIR Ment Health. 2019;6(10):e14166.
[19] Auriacombe M, Moriceau S, Serre F, et al. Development and validation of a virtual agent to screen tobacco and alcohol use disorders. Drug Alcohol Depend. 2018;193: 1–6.
[20] Rammohan R, Dhanabalsamy N, Dimov V, et al. Smartphone conversational agents (Apple Siri, Google, Windows Cortana) and questions about allergy and asthma emergencies. Journal of Allergy and Clinical Immunology. 2017;139(2):AB250.
[21] Baptista S, Wadley G, Bird D, My Diabetes Coach Research Group, et al. Acceptability of an embodied conversational agent for type 2 diabetes self-management education and support via a smartphone app: mixed methods study. JMIR Mhealth Uhealth. 2020;8(7):e17038.
[22] Omoregbe NAI, Ndaman IO, Misra S, et al. Text Messaging-Based medical diagnosis using natural language processing and fuzzy logic. J Healthcare Eng. 2020; 2020:1–14.
[23] Vanagas G, Engelbrecht R, Damaševičius R, et al. EHealth solutions for the integrated healthcare. J Healthc Eng. 2018;2018:3846892.
[24] Merdivan E, Singh D, Hanke S, et al. Dialog systems for intelligent human computer interactions. Elect Notes Theoret Comput Sci. 2019;343:57–71.
[25] Peeters MM. ReMindMe: agent-based support for self-disclosure of personal memories in people with Alzheimer's disease. 2nd International Conference on Information and Communication Technologies for Ageing Well and e-Health 2016; Rome, Italy, 2016.
[26] Callejas Z, Griol D, McTear MF, et al. A virtual coach for active ageing based on sentient computing and M-health. Proceedings of IWAAL 2014; Belfast, UK, 2014.
[27] Vastenburg MH, Visser T, Vermaas M, et al. Designing acceptable assisted living services for elderly users. Proceedings of AMI 2008; Nuremberg, Germany, 2008.
[28] Hoermann S, McCabe KL, Milne DN, et al. Application of synchronous text-based dialog systems in mental health interventions: systematic review. J Med Internet Res. 2017; 19(8):e267.
[29] Montenegro JLZ, da Costa CA, da Rosa Righi R. Survey of conversational agents in health. Expert Syst Appl. 2019; 129:56–67.
[30] Chattopadhyay D, Ma T, Sharifi H, et al. Computer-controlled virtual humans in patient-facing systems: systematic review and meta-analysis. J Med Internet Res. 2020; 22(7):e18839.
[31] Milne-Ives M, de Cock C, Lim E, et al. The effectiveness of artificial intelligence conversational agents in health care: systematic review. J Med Internet Res. 2020;22(10):e20346.
[32] Vaidyam AN, Linggonegoro D, Torous J.Changements du paysage psychiatrique des chatbots: une revue systématique des agents conversationnels dans la maladie mentale sérieuse [Changes to the psychiatric chatbot landscape: a systematic review of conversational agents in serious mental illness]. Can J Psychiatry. 2021;66(4): 339–348.
[33] [1]Pacheco-Lorenzo MR, Valladares-Rodríguez SM, Anido-Rifón LE, et al. Smart conversational agents for the detection of neuropsychiatric disorders: a systematic review. J Biomed Inform. 2021;113:103632.
[34] Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1): 19–32.
[35] Turner M, Kitchenham B, Brereton P, et al. Does the technology acceptance model predict actual use? A systematic literature review. Inf Softw Technol. 2010;52(5):463–479.
[36] Moher D, Liberati A, Tetzlaff J, The PRISMA Group, et al. Preferred reporting items for systematic reviews and meta-analyses: the Prisma statement. PLOS Med. 2009; 6(7):e1000097–6.
[37] McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012;22(3):276–282.
[38] Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340(1): c332.
[39] Zaveri A, Rula A, Maurino A, et al. Quality assessment for linked data: a survey. SW. 2015;7(1):63–93.
[40] Bumble JL, Carter EW. Application of the world café to disability issues: a systematic review. J Disabil Policy Stud. 2021;32 (3):193–203.
[41] Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health. 2017;4(2):e19.
[42] Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth. 2018;6(11): e12106.
[43] Gong E, Baptista S, Russell A, et al. My diabetes coach, a mobile app-based interactive conversational agent to support type 2 diabetes self-management: randomized effectiveness-implementation trial. J Med Internet Res. 2020; 22(11):e20322.
[44] Prochaska JJ, Vogel EA, Chieng A, et al. A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study. J Med Internet Res. 2021;23(3):e24850.
[45] Burton C, Szentagotai Tatar A, McKinstry B, Help4Mood Consortium, et al. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. J Telemed Telecare. 2016;22(6):348–355.
[46] Ali MR, Razavi SZ, Langevin R, et al. A virtual conversational agent for teens with autism spectrum disorder. In Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents. IVA '20: ACM International Conference on Intelligent Virtual Agents, Glasgow, UK, ACM; 2020.
[47] Smith MJ, Ginger EJ, Wright K, et al. Virtual reality job interview training in adults with autism spectrum disorder. J Autism Dev Disord. 2014;44(10):2450–2463.
[48] Wargnier P, Carletti G, Laurent-Corniquet Y, et al.. Field evaluation with cognitively-impaired older adults of attention management in the embodied conversational agent

louise. In 2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH); Orlando, FL, USA, 2016.
[49] Tanaka H, Adachi H, Ukita N, et al. Detecting dementia through interactive computer avatars. IEEE J. Transl. Eng. Health Med. 2017;5:1–11.
[50] Ujiro T, Tanaka H, Adachi H, et al. Detection of dementia from responses to atypical questions asked by embodied conversational agents. In Interspeech 2018, ISCA; Hyderabad, India, 2018
[51] Huang J, Li Q, Xue Y, et al. TeenChat: a chatterbot system for sensing and releasing adolescents' stress. In: Yin X, Ho K, Zeng, D, Aickelin U, Zhou R, Wang H, editors. Health information science. Springer International Publishing; 2015. p. 133–145.
[52] Bresó A, Martínez-Miranda J, Botella C, et al. Usability and acceptability assessment of an empathic virtual agent to prevent major depression. In Expert Systems. 2016;33(4): 297–312.
[53] Jang S, Kim J-J, Kim S-J, et al. Mobile app-based chatbot to deliver cognitive behavioral therapy and psychoeducation for adults with attention deficit: a development and feasibility/usability study. Int J Med Inform. 2021;150: 104440.
[54] Hauser-Ulrich S, Künzli H, Meier-Peterhans D, et al. A smartphone-based health care chatbot to promote self-management of chronic pain (SELMA): pilot randomized controlled trial. JMIR Mhealth Uhealth. 2020;8(4):e15806.
[55] Rathnayaka P, Mills N, Burnett D, et al. A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. In Sensors. 2022;22(10):3653.
[56] Mateos-Sanchez M, Melo AC, Blanco LS, et al. Chatbot, as educational and inclusive tool for people with intellectual disabilities. In Sustainability. 2022;14(3):1520.
[57] Bérubé C, Schachner T, Keller R, et al. Voice-Based conversational agents for the prevention and management of chronic and mental health conditions: systematic literature review. J Med Internet Res. 2021;23(3):e25933.
[58] Greuter S, Balandin S, Watson J. Social games are fun: exploring social interactions on smart speaker platforms for people with disabilities. In Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts. CHI PLAY '19: The Annual Symposium on Computer-Human Interaction in Play, Barcelona, Spain, ACM; 2019.
[59] Abd-Alrazaq AA, Alajlani M, Alalwan AA, et al. An overview of the features of chatbots in mental health: a scoping review. Int J Med Inform. 2019;132:103978.
[60] Reis A, Paulino D, Paredes H, et al. Using intelligent personal assistants to assist the elderlies an evaluation of Amazon Alexa, Google Assistant, Microsoft Cortana, and Apple Siri. In 2018 2nd International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW), Thessaloniki, Greece¸ IEEE; 2018.
[61] Gotthardt M, Striegl J, Loitsch C, et al. Voice assistant-based CBT for depression in students: effects of Empathy-Driven dialog management. In: Miesenberger K, Kouroupetroglou G, Mavrou K, Manduchi R, Covarrubias Rodriguez M, Penáz P, editors. Lecture notes in computer science. Springer International Publishing; 2022. p. 451–461.
[62] Gebhard P, Mehlmann G, Kipp M. Visual SceneMaker a tool for authoring interactive virtual characters. J Multimod User Interf. 2012;6:3–11.
[63] Gruenstein A, McGraw I, Badr I. The WAMI toolkit for developing, deploying, and evaluating web-accessible multimodal interfaces. Proceedings of the 10th international conference on Multimodal interfaces, ICMI 2008, Chania Crete Greece; 2008, p. 141–148.
[64] van Waterschoot J, Bruijnes M, Flokstra J, et al. Flipper 2.0: a pragmatic dialog engine for embodied conversational agents Proceedings of IVA 2018, Sydney, Australia; 2018.
[65] Ultes S, Rojas-Barahona L, Su P-H, et al. PyDial: a multi-domain statistical dialog system toolkit. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics-System Demonstrations 2017, Vancouver, Canada; 2017.
[66] Lison P, Kennington C. OpenDial: a toolkit for developing spoken dialog systems with probabilistic rules. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics–System Demonstrations 2016, Berlin, Germany; 2016.
[67] Yu Z, Ramanarayanan V, Mundkowsky R, et al.. Multimodal HALEF: an open-source modular web-based multimodal dialog framework IWSDS 2016; Saariselkä, Finland, 2016.
[68] Hartholt A, Traum D, Marsella S, et al. All together now introducing the virtual human toolkit. Proceedings of 13th International Conference on Intelligent Virtual Agents 2013, Edinburgh, UK; 2013.
[69] Rich C, Sidner CL. Using collaborative discourse theory to partially automate dialog tree authoring. Proceedings of IVA 2012, Santa Cruz, CA, USA; 2012.
[70] Skantze G, Al Moubayed S. IrisTK: a statechart-based toolkit for multi-party face-to-face interaction. Proceedings of ICMI 2012, Santa Monica, CA, USA; 2012.
[71] Bohus D, Rudnicky AI. The RavenClaw dialogue management framework: architecture and systems computer speech and language. Computer Speech and Language. 2009:23:332–361.
[72] Zhao S, Sampson S, Xia J, Cochrane Schizophrenia Group, et al. Psychoeducation (brief) for people with serious mental illness. Cochrane Database Syst Rev. 2015;2015(4)
[73] Bali M, Mohanty S, Chatterjee S, et al. Diabot: a predictive medical chatbot using ensemble learning. Int J Recent Technol Eng. 2019;8(2):2277–3878.
[74] Joerin A, Rauws M, Ackerman ML. Psychological artificial intelligence service, Tess: delivering on-demand support to patients and their caregivers: technical report. Cureus. 11(1):e3972. doi:10.7759/cureus.3972; 2019.
[75] Meier P, Beinke JH, Fitte C, et al. FeelFit – design and evaluation of a conversational agent to enhance health awareness. Proceedings of ICIS 2019, Munich, Germany; 2019.
[76] Xu L, Zhou Q, Gong K, et al.. End-to-end knowledge-routed relational dialog system for automatic diagnosis. Proceedings of AAAI 2019, Honolulu Hawaii USA; 2019.
[77] Cameron G, Cameron D, Megaw G, et al. Assessing the usability of a chatbot for mental health care. Proceedings of INSCI 2018 Workshops, St. Petersburg, Russia; 2018.
[78] Cooper A, Ireland D. Designing a chat-bot for non-verbal children on the autism spectrum. In: Schaper LK, Ryan A,Cummings E, editors. Connecting the system to

enhance the practitioner and consumer experience in healthcare. IOS Press; 2018.
[79] Tanaka H, Negoro H, Iwasaka H, et al. Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders. PLOS One. 2017;12(8):e0182151.
[80] Ireland D, Atay C, Liddle J, et al. Hello harlie: enabling speech monitoring through chat-bot conversations. In: Schaper LK,Georgiou A,Whetton S, editors. Digital health innovation for consumers, clinicians, connectivity and community. IOS Press; 2016.
[81] Allen J, Ferguson G, Blaylock N, et al. Chester: towards a personal medication advisor. J Biomed Inform. 2006;39(5): 500–513.
[82] Iio T, Yoshikawa Y, Chiba M, et al.. Twin-robot dialog system with robustness against speech recognition failure in human-robot dialog with elderly people. Applied Sciences. 2020;10(4):1522.
[83] Sansen H, Chollet G, Glackin C, et al. The roberta IRONSIDE project a cognitive and physical robot coach for dependent persons. Proceedings of HandiCap 2016, Paris, France; 2016.
[84] Lane ND, Lin M, Mohammod M, et al. BeWell: sensing sleep, physical activities and social interactions to promote wellbeing. Mobile Netw Appl. 2014;19(3):345–359.
[85] Todd M, Sasi S. Intelligent virtual companion system for independent living. Proceedings of ICAI 2006, Las Vegas, Nevada, USA; 2006.
[86] Gaffney H, Mansell W, Tai S. Agents of change: understanding the therapeutic processes associated with the helpfulness of therapy for mental health problems with relational agent MYLO. Digit Health. 2020;6: 2055207620911580.
[87] Bott N, Wexler S, Drury L, et al. A protocol-driven, bedside digital conversational agent to support nurse teams and mitigate risks of hospitalization in older adults: case control PrePost study. J Med Internet Res. 2019;21(10):e13440.
[88] Hernandez JPT. Network diffusion and technology acceptance of a nurse chatbot for chronic disease self-management support: a theoretical perspective. J Med Invest. 2019;66(1.2):24–30.
[89] Jungmann SM, Klan T, Kuhn S, et al. Accuracy of a chatbot (ada) in the diagnosis of mental disorders: comparative case study with lay and expert users. JMIR Form Res. 2019;3(4):e13863.
[90] Lee M, Ackermans S, van As N, et al. Caring for Vincent: a chatbot for self-compassion Proceedings of CHI 2019, Glasgow, UK; 2019.
[91] Divya S, Indumathi V, Ishwarya S, et al. A self-diagnosis medical chatbot using artificial intelligence. J Web Develop Web Design. 2018;3(1):7.
[92] Rarhi K, Bhattacharya A, Mishra A, et al. 2018. Automated medical chatbot. SSRN Electronic Journal; https://doi.org/10.2139/ssrn.3090881
[93] Srivastava M, Suvarna S, Srivastava A, et al. Automated emergency paramedical response system. Health Inf Sci Syst. 2018;6(1):22.
[94] Chi N-C, Sparks O, Lin S-Y, et al. Pilot testing a digital pet avatar for older adults. Geriatric Nursing. 2017;38(6): 542–547.
[95] Mujeeb S, Hafeez M, Arshad T. Aquabot: a diagnostic chatbot for achluophobia and autism. International Journal of Advanced Computer Science and Applications (IJACSA) 2017;8(9),209-216.
[96] Ni L, Lu C, Liu N, et al. MANDY: towards a smart primary care chatbot application Proceedings of International Symposium on Knowledge and Systems 2017, Bangkok, Thailand; 2017.
[97] Philip P, Micoulaud-Franchi J-A, Sagaspe P, et al. Virtual human as a new diagnostic tool, a proof of concept study in the field of major depressive disorders. Sci Rep. 2017;7: 42656.
[98] Morbini F, Forbell E, De Vault D, et al. A Mixed-Initiative conversational dialog system for healthcare Proceedings of SIGDIAL 2012: the 13th Annual Meeting of the Special Interest Group on Discourse and Dialog, Seoul, South Korea; 2012.
[99] Levin E, Levin A. Evaluation of spoken dialog technology for real-time health data collection. J Med Internet Res. 2006;8(4):e30.
[100] Preininger AM, South B, Heiland J, et al. Artificial intelligence-Based conversational agent to support medication prescribing. JAMIA Open. 2020;3(2):225–232.
[101] Morris RR, Kouddous K, Kshirsagar R, et al. Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. J Med Internet Res. 2018;20(6):e10148.
[102] Wolters MK, Kelly F, Kilgour J. Designing a spoken dialog interface to an intelligent cognitive assistant for people with dementia. Health Informatics J. 2016;22(4):854–866.
[103] Beveridge M, Fox J. Automatic generation of spoken dialog from medical plans and ontologies. Journal of Biomedical Informatics. 2006;39(5):482–499.
[104] Munoz R, Villarroel R, Barcelos TS, et al. Developing computational thinking skills in adolescents with autism spectrum disorder through digital game programming; IEEE Access, vol. 6. 2018; p. 63880–63889.
[105] Kapočiūtė-Dzikienė J. A Domain-Specific generative chatbot trained from little data. Applied Sciences. 2020;10(7): 2221.
[106] Betriana F, Osaka K, Matsumoto K, et al. Relating mori's uncanny valley in generating conversations with artificial affective communication and natural language processing. Nurs Philos. 2021;22(2).
[107] Mori M. The uncanny valley. Energy. 1970;7(4):33–35.
[108] Männistö-Funk T, Sihvonen T. Voices from the uncanny valley. Dig Culture Soc. 2018;4(1):45–64.
[109] Złotowski JA, Sumioka H, Nishio S, et al. Persistence of the uncanny valley. In: Ishiguro H, Dalla Libera F, editors. Geminoid studies. Singapore: Springer; 2018.
[110] MacDorman KF, Chattopadhyay D. Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. Cognition. 2016; 146:190–205.
[111] Weisman WD, Peña JF. Face the uncanny: the effects of doppelganger talking head avatars on Affect-Based trust toward artificial intelligence technology are mediated by uncanny valley perceptions. Cyberpsychol Behav Soc Netw. 2021;24(3):182–187.

[112] Clark L, Ofemile A, Cowan BR. Exploring verbal uncanny valley effects with vague language in computer speech. In: Weiss B, Trouvain J, Barkat-Defradas M, Ohala J, editors. Voice attractiveness. Singapore: Springer; 2020. p. 317–330.
[113] Baird A, Parada-Cabaleiro E, Hantke S, et al. The perception and analysis of the likeability and human likeness of synthesized speech. Interspeech 2018, Hyderabad, India; 2018.
[114] Ciechanowski L, Przegalinska A, Magnuski M, et al. In the shades of the uncanny valley: an experimental study of human–chatbot interaction. Future Gener Comput Syst. 2019;92:539–548.
[115] Kätsyri J, Förger K, Mäkäräinen M, et al. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to The Valley of eeriness. Front Psychol. 2015;6:390.
[116] Tinwell A, Grimshaw M, Abdel-Nabi D.. Effect of emotion and articulation of speech on the uncanny valley in virtual characters. Vol. 6975. In: D'Mello S, Graesser A, Schuller B, Martin JC, editors. Affective computing and intelligent interaction. ACII 2011. Lecture notes in computer science. Berlin, Heidelberg: Springer; 2011
[117] Romportl J. Speech synthesis and uncanny valley. Vol 8655. In: Sojka P, Horák A, Kopeček I, Pala K, editor names. Text, speech and dialogue. TSD. Lecture notes in computer science. Cham: Springer; 2014.
[118] Tinwell A, Grimshaw; Deborah M, Nabi A. The effect of onset asynchrony in audio-visual speech and the uncanny valley in virtual characters. IJMRS. 2015;2(2):97.
[119] Schwind V, Leicht K, Jäger S, et al. Is there an uncanny valley of virtual animals? A quantitative and qualitative investigation. Int J Hum Comput Stud. 2018;111:49–61.