Conversational Interfaces in IoT Ecosystems: Where We Are, What Is Still Missing

Simone, Gallo
CNR - ISTI, Pisa, Italy University of Pisa, Pisa, Italy
simone.gallo@isti.cnr.it

Alessio, Malizia
University of Pisa, Pisa, Italy Molde University College, Molde, Norway
alessio.malizia@unipi.it

Fabio, Paternò
CNR - ISTI, Pisa, Italy
fabio.paterno@isti.cnr.it

ABSTRACT

In the last few years, text and voice-based conversational agents have become more and more popular all over the world as virtual assistants for a variety of tasks. In addition, the deployment on the market of many smart objects connected with these agents has introduced the possibility of controlling and personalising the behaviour of several connected objects using natural language. This has the potential to allow people, also those without a technical background, to effectively control and use the wide variety of connected objects and services. In this paper, we present an analysis of how conversational agents have been used to interact with smart environments (such as smart homes). For this purpose, we have carried out a systematic literature review considering publications selected from the ACM and IEEE digital libraries to investigate the technologies used to design and develop conversational agents for IoT settings, including Artificial Intelligence techniques, the purpose that they have been used for, and the level of user involvement in such studies. The resulting analysis is useful to better understand how this field is evolving and indicate the challenges still open in this area that should be addressed in future research work to allow people to completely benefit from this type of solution.

CCS CONCEPTS

• **Human-centered computing** → Human-computer interaction (HCI); Interaction paradigms; Natural language interfaces; • **General and reference** → Document types; Surveys and overviews; • **Computer systems organization** → Embedded and cyber-physical systems; Sensors and actuators.

KEYWORDS

Conversational Agents, Internet of Things, User Experience

ACM Reference Format:
Simone, Gallo, Alessio, Malizia, and Fabio, Paternò. 2023. Conversational Interfaces in IoT Ecosystems: Where We Are, What Is Still Missing. In *International Conference on Mobile and Ubiquitous Multimedia (MUM '23), December 03–06, 2023, Vienna, Austria.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3626705.3627775

MUM '23, December 03–06, 2023, Vienna, Austria

ACM ISBN 979-8-4007-0921-0/23/12.
https://doi.org/10.1145/3626705.3627775

1 INTRODUCTION

The use of smart objects made possible by the advent of the Internet of Things (IoT) is undergoing rapid and steady growth that is leading to their spread and use in the most common daily environments. Recent reports show that the number of objects connected to the Internet is already substantial and will continue to increase in the coming years: in 2021, there were about 11 billion objects, and this figure is projected to increase to 30 billion in 2030 [69]. The everyday use of these technologies finds application in different areas, supporting end users and organizations in performing tasks of different nature and complexity. Five macro-areas can be identified in which IoT devices find use [4]: healthcare, through the use of smart wearables and personal monitoring; environmental, which includes smart farming, smart agriculture, wildlife monitoring and climate change monitoring; smart cities composed of smart homes and buildings, traffic and security monitoring, and in commercial and industrial sectors. Thus, the spread of these devices opens new possibilities for improving people's quality of life, from comfort and sustainability through smart home automation (e.g., [3]) and energy consumption control (e.g., [36]) to assistance for the older adults and impaired (e.g. [44]) or remote monitoring and control of farming systems (e.g., [8]). However, there is a need to increase user confidence in the use of these technologies [25]. Providing the users (especially the non-experts) with useful tools for better understanding and controlling these environments is crucial for implementing the smart living vision.

Using tools such as chatbots or, more in general virtual assistants accessible through natural language can be a promising approach to breaking down barriers between the user and technology given their potential ease of use, as demonstrated by the recent success of ChatGPT or by the diffusion of widely adopted commercial products such as Alexa or Google Assistant. Thus, it is important to understand how the development of conversational systems can empower non-technical individuals to interact intuitively with smart environments (e.g., homes), how such tools are developed and tested, and to what extent they meet the user needs while keeping the interaction simple and clear even for complex tasks.

Despite several research papers providing comprehensive reviews of the literature on conversational agents from various perspectives and application fields, the field concerning conversational systems applied to the Internet of Things has received limited attention. In previous studies, a systematic literature review proposed by Rapp et al. [63] focuses on the interaction between text-based chatbots and users, considering general aspects such as user satisfaction, trust, and acceptance when engaging with chatbots. In another study, Suhaili et al. [53] conducted a comprehensive review of task-oriented chatbots, emphasizing the technical implementation aspects rather than user perception. Their analysis aims to identify the techniques employed in developing the chatbots' capabilities to understand user requests and generate appropriate responses. Regarding specific application domains, a study [58] reviews the state of the art of chatbots for education by exploring in which education sub-field are these solutions employed, how the educational system can benefit from the use of chatbots, what the challenges faced during the implementation (such as ethical and evaluation issues) and which areas of education could potentially benefit for the use of chatbots. The applications of chatbots in healthcare has been considered in [72], with a particular focus on oncology application. In this case, the authors identify six main task categories (diagnosis, treatment, monitoring, support, workflow, and health promotion) in which healthcare chatbots are employed. Based on these categories, the authors indicate how chatbots act in assisting ontology care patients, defining the current limitations and proposing various aspects that could be improved to enhance their efficiency. Overall, there is a lack of contributions that provide a review of the literature on chatbots applied in the Internet of Things domain.

For this purpose, we have carried out this systematic literature review that aims to analyse the various contributions in the field of conversational agents in the context of IoT ecosystems, and then identify areas that need further investigations and aspects that require additional research efforts. In the paper we first introduce the method followed in the systematic review to investigate the use of conversational agents to control IoT ecosystems. We review the relevant papers according to the identified research questions. Then, we discuss the analysis carried out and identify emerging trends in the considered field and areas that require more work. Lastly, we draw some conclusions and provide indications for future work.

2 METHODOLOGY

Following the guidelines introduced by Kitchenham and Charters [46], this review begins with a planning phase that involves conducting an initial analysis of the chatbot literature. This analysis revealed a lack of reviews concerning conversational systems applied to intelligent environments and connected objects. Subsequently, we defined and agreed upon the research questions and the review protocol, which included the search process and the definition of inclusion and exclusion criteria. Once the articles were retrieved from the chosen databases, we started the conduction phase. This stage comprised a two-step screening process: the first screening was based on the title and abstract of the articles, the second screening applied the exclusion criteria. During the second screening, we discussed the relevance of the papers for our study.

Lastly, the obtained articles were analysed to address the research questions and report the results obtained.

2.1 Research questions

The research questions have been defined to investigate the evolution of chatbots in IoT ecosystems and identify areas that require further investigation. The goal is to analyse how intelligent techniques have been exploited, which application domains have stimulated more interest, what interaction modalities have been considered, how conversational breakdowns have been managed and how user experience has been assessed. By addressing such aspects, we can provide a clear picture of the considered area since they cover both the technological aspects and those related to the users and the applications considered. In particular, the research questions for the literature review are:

RQ1: What intelligent technologies have been used to build IoT conversational agents?

Understanding these technologies offers insights into the technical foundations and capabilities of chatbots in the IoT context, enabling a more profound comprehension of their potential and limitations.

RQ2: In which IoT application domains have conversational agents been employed?

Identifying these domains aids researchers in understanding explored areas and potential avenues for further investigation and improvement.

RQ3: What devices and modalities have been considered for accessing conversational agents and how they have been deployed?

Analysing the interaction methods applied can be useful to understand whether there are areas not yet sufficiently explored and possible limitations in the approaches adopted.

RQ4: How have conversational breakdowns been addressed?

Conversational agents must effectively handle challenges such as language ambiguity, user intent misinterpretation, and technical limitations. Resolving breakdowns can enhance user experience and agent reliability in IoT ecosystems.

RQ5: What methodologies have been used to measure the usability and user experience of the proposed conversational agents?

Analysing whether and how usability and user experience evaluation methods have been applied is useful to understand whether such aspects have been sufficiently considered and how they can be further investigated.

2.2 Search process

In order to identify the relevant articles for conducting this systematic review, we selected a set of papers from two digital libraries, those of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE), in May 2023. The articles were obtained by running queries with a string of keywords aimed at finding contributions that addressed both the conversational aspect and the smart IoT context. The following keywords and phrases were used in our search:

("conversational agent" OR "conversational AI" OR "intelligent assistant" OR "conversational assistant" OR "virtual assistant" OR "intelligent agent" OR "chatbot" OR "chatterbot" OR "chatterbox" OR "socialbot" OR "digital assistant" OR "conversational UI" OR "conversational interface" OR "conversation system" OR "conversational system" OR "dialogue system" OR "dialog system" OR "vocal interaction" OR "natural language interaction" OR "natural language processing") AND ("ambient computing" OR "smart spaces" OR "IoT" OR "Internet Of Things" OR "IoT environments" OR "automations" OR "smart environment" OR "smart home" OR "IoT service mashup" OR "intelligent environment" OR "intelligent spaces" OR "intelligent ambient")

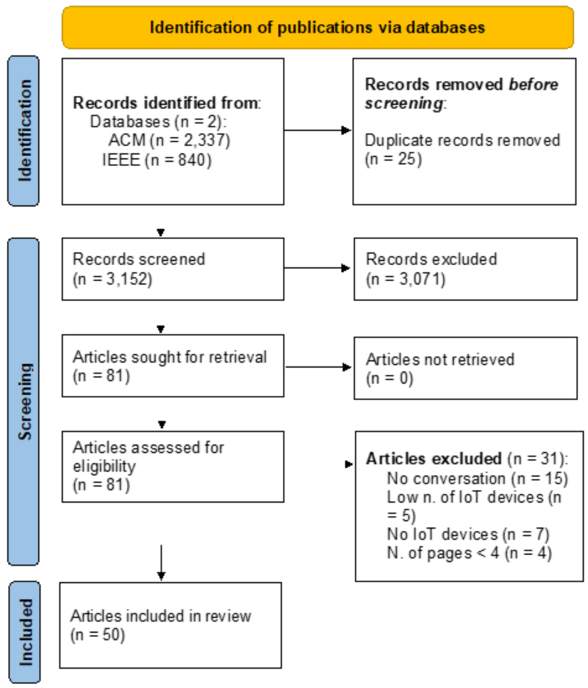

Figure 1: Process followed for paper selection.

2.3 Inclusion and Exclusion criteria

In order to obtain relevant articles for the review, we established a set of selection criteria. These were used to determine whether a paper should be included in the analysis. Specifically, the following selection criteria were applied: 1) the papers must be written in English; 2) the papers must address the application of conversational systems to IoT environments, excluding studies that focus solely on interaction with a single object; 3) the interactive system must utilize natural language within a conversational approach, thus we did not consider applications that exclusively use predefined commands (for instance, certain Telegram chatbots that employ commands such as '/start', '/createItem', or '/stop' to communicate); and 4) the papers must be at least four pages in length, in order to have sufficient content to analyse.

2.4 Papers selection

With the use of the keywords outlined in the previous section, we identified a total of 3,177 articles within the selected digital libraries. The ACM Digital Library contributed 2,337 of these articles, while the remaining 840 articles were obtained from the Institute of Electrical and Electronics Engineers (IEEE) library. References to the articles were collected in BibTeX format and processed using an open-source reference manager called JabRef. The information regarding the selection process is summarized below, using the PRISMA flow diagram (Figure 1).

An initial screening phase (title + abstract) on the 3177 articles from the two digital libraries led to the exclusion of 25 duplicated papers (through JabRef filter), and of 3071 papers that did not cover the topic of our interest (for example including only NLP or IoT methods and algorithms, systematic reviews about only chatbots or IoT, chatbot not related to IoT or IoT not related to chatbots). We then applied the exclusion criteria detailed in the preceding section on the resulting 81 papers, determining the final relevance and inclusion of each paper in our analysis.

As a result, a further 31 articles were excluded, largely due to their lack of conversational aspects, consideration of insufficient or no Internet of Things (IoT) devices, and inadequate length. After applying these exclusion criteria, we arrived at a refined list of articles. The search phase concluded with the set of the remaining 50 articles (see the Appendix for the papers list).

3 LITERATURE ANALYSIS

From a temporal viewpoint (Table 1), an overview of the articles retrieved shows some "early tentative" in the first decade of the 2000s, starting from 2005 when the concept of IoT was put forward and Natural Language Processing technologies were not developed as today. It seems that interest in this topic began to be more relevant starting in 2017 (the year when Alexa and Google Home were released worldwide), peaking in 2018 and 2021.
A preview of the key aspects related to the research questions is summarised in Figure 2. The boxes refer to the research questions. They include a list representing the categories derived from the analysis of the papers' content, each element is associated with the number of contributions for the corresponding category. In the box on UX evaluation methods the papers have been assigned to multiple categories since usually the studies adopted more than one evaluation metric. The full list of evaluation metrics can be found in Table 2.

3.1 Methods and Tools for conversational agents

Conversational agents can be categorized into Task-Oriented and Non-Task Oriented types [17, 39], with the former focused on specific goals and the latter engaging in open-ended conversations with users, and they can be further subcategorized based on their architecture, such as rule-based, corpus-based, frame-based and dialogue state-based, each employing different methods for generating responses and managing interactions [39].

Given such a premise, and since most work does not explain in-depth (and in some cases at all) the architecture exploited, we choose to label as "rule-based" all the implementations that do not use machine learning (ML) or deep learning (DL) in the process of intent classification and entity extraction. Often the available solutions can be defined as architectural hybrids, applying frame-based architecture augmented with some dialogue-state components, as they use machine learning and deep learning techniques to identify intents and entities but base the dialogue management and the response generation on predefined rules and patterns. On the other hand, the solutions implemented in Rasa (an open-source framework to build conversational systems) adopt a dialogue policy that uses machine learning to predict the next most accurate dialogue action (send a response, wait for other messages, ask for clarifications, . . .).

Among the 50 conversational agents considered in the survey, 18 have been considered rule-based, such as [1, 6, 7, 10, 14]. These

Table 1: Distribution of selected articles per year

Year	2005	2015	2016	2017	2018	2019	2020	2021	2022
N. of articles	1	1	1	7	13	7	4	10	6

chatbots use pre-defined rules (e.g., using regular expressions) to classify the intent and respond to user inputs. Such works do not rely on machine learning algorithms, but in most cases (see in the following sections) use third-party services such as Google Speech API, which exploit machine learning to perform speech-to-text and text-to-speech. Other analysed solutions use ML-based approaches for entity extraction and intent classification implementing their custom architecture or using frameworks. For example, [15, 38, 49, 62] use ML-based algorithms for entity extraction and intent classification, while others [59, 71, 73] use Deep Learning approaches (such as RNNs, LSTM and Transformers) to reach the same goal. One solution [22] adopts a reinforcement learning algorithm (using the Q-learning method) to allow the chatbot to make new associations between unseen commands and the actions to be performed. Several papers, on the other hand, report the use of frameworks such as Rasa [67], Dialogflow [19, 27, 48], IBM Watson [43], Amazon AVS [28] to manage both the recognition of intents and entities, but also for the conversational flow and ease of integration with instant messaging platforms or virtual assistant (e.g., Telegram, Facebook Messenger, Line, Alexa, Google Assistant). Among the ML-based agents (32 of 50 articles), the most popular framework is Dialogflow, used in nine systems, followed by Rasa and Amazon AVS in four systems respectively. Other frameworks such as IBM Watson, Microsoft Louis, WIT.ai, Google Assistant SDK, Amazon Skill Kit and Mycroft are rarely used (six times in total). Finally, nine systems are based on custom implementations using techniques such as Support Vector Machines (SVMs), Recurrent Neural Networks (RNNs), and Transformers such as BERT. Overall, such data suggest that the use of ML-based frameworks and techniques is the most common choice in the development of conversational systems. This is likely due to the ability of ML to handle complex and varied inputs, as well as the availability of pre-trained models (e.g., BERT) that can be fine-tuned for specific use cases. Rule-based systems, on the other hand, are less commonly used due to their limited flexibility and the difficulty of maintaining and updating the large number of rules required for natural language processing. In terms of specific frameworks and techniques, Dialogflow and Rasa are the most popular among the systems surveyed, likely due to their – relative – ease of use and the availability of pre-built components and integrations. Dialogflow is proprietary, while Rasa is open source. Custom implementations using techniques such as SVMs, RNNs, and Transformers are also relatively popular, indicating that several developers are willing to invest the time and effort required to build a tailored solution for their specific needs. In many of the papers analysed the authors pay little attention to the description of the implementation of conversational agents and, in some cases, this aspect is completely ignored.

Among the considered articles, only three share the implementation code [15, 45, 59], while 23 present an implementation description that can be used to reproduce the work to some extent. Thus, in addition to stating, for instance, which framework or algorithm was used for input classification, there is also a description of how the various intents and entities were organised (e.g., by giving examples of the training phrases and the NLP analysis pipeline), the chatbot functionalities, the management of the conversational flow and, in specific cases, the management of breakdowns (analysed in section 3.5). For instance, in some contributions [19, 27, 28] the authors dedicate a significant part of the work to describing the implementation of intents and the related entities, providing examples of training phrases, chatbot functionalities, how the conversation is handled and how the conversational system is integrated into third-party or customised instant messaging applications. The remaining 26 articles do not present enough information to reproduce the work partially or entirely. The support for creating automations

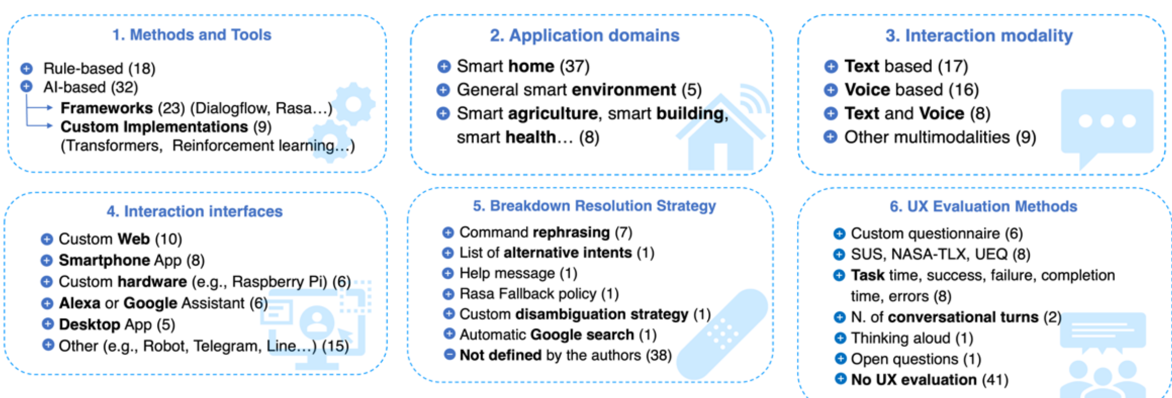

Figure 2: Summary of the key aspects of the research questions and the corresponding literature review contributions.

in a trigger-action format, also defined as "customization rules" or "routines," has been discussed in several works, such as [15, 19, 20, 27, 45, 67]. These papers propose different approaches for creating automations with varying types of triggers and actions. For instance, in [15] users can create automations that perform one or more actions when a time-related trigger occurs (e.g., "every day at 6 am get the latest weather forecast and send it via email to Bob"). In other cases [19, 20] the system suggests existing IFTTT rules to the user based on an abstract description of the desired behaviour. The authors divide a rule into two components, the "what" component indicating the desired automated action and the "when" component specifying the context for execution (e.g., "I would like to secure my places when I leave them"). The system can also identify rules that cannot be implemented with the user's connected devices. In one case [27] the chatbot allows the creation of rules consisting of several triggers and actions (e.g., "If the lights are on when leaving the house, turn them off and send me a notification"), where triggers can refer to events and conditions related to sensors and smart objects and are concatenated using the logical AND/OR operators, while actions are executed sequentially. Kim et al. [45] propose a system that can identify and allow users to modify existing rules in their context according to specific user's goals. This is done in a two-step process: localizing the relevant automations currently deployed in the smart home and then modifying them (by appending a new component, replacing existing components, or updating parameters) according to the user input. The chatbot in [67] can compose rules consisting of a trigger and an action, considering several users called "actors." An example of such a rule is "I want the bedroom lights to turn off when me and my husband get in bed at night", where "me" and "my husband" are identified as actors. Finally, Barricelli et al. [9] developed an Alexa skill that allows the creation of Alexa routines through a multimodal approach, combining the use of voice and touch (on Amazon devices with a screen) to guide the user in selecting and configuring triggers and actions.

3.1.1 Other AI Technologies Involved. In addition to the technologies used for the realisation of conversational systems, various AI technologies were employed in the reviewed studies to enhance and/or add functionality to such systems, including speech recognition, conflict resolution reasoning, and various approaches for face and emotion recognition. Regarding better support for user interaction, speech recognition technologies, such as the Google API [6, 18, 23, 57, 59], Microsoft Bing Speech API [2], Web Speech API [27] and Android Speech Recognition [13] for speech-to-text and text-to-speech, were used in several studies. These technologies allow users to communicate with the system using voice commands, enabling a more convenient and intuitive user experience, and enhancing the accessibility of the application.

Deep Learning technologies, such as Vokaturi for emotion classification using voice [13] and convolutional neural networks for emotion analysis [66], were utilized to enable a home automation system to understand and respond to the emotional states of users. This can be useful in ambient assisted living environments, where the system can provide appropriate assistance or support based on the user's emotional state. Machine learning techniques, including algorithms for device classification [16] and user pattern recognition [54, 71], were applied in several studies to improve the functionality and security of home automation systems. For example, a machine learning-based algorithm took a stream of packets sent by a device and classified the device based on the contents of the packets to enhance IoT cybersecurity [16]. Moreover, one contribution [38] used the Frequent Pattern Growth Algorithm to mine user activities starting from sensors data to optimize the users' commands based on previous interactions. Case-based reasoning, a problem-solving approach that utilizes solutions from previously solved problems, was used in [59] to resolve conflicts between commands given by different people in a multi-user smart home environment. Finally, various approaches for face recognition, such as the Local Binary Pattern Histogram (LBPH) algorithm [52] and Deep Learning models, were employed in several studies to enable a home automation system to identify and respond to individual users. For example, [38] and [68] use OpenCV to authorize user access and control. The use of such technologies can be useful for further personalizing the users' experiences through learning their preferences, and for security purposes such as authenticating users before allowing them to execute certain commands. Overall, the use of these technologies has the potential to enhance the functionality, usability, and security of home automation systems, and to provide improved efficiency and personalized experiences for users.

3.2 Application Fields

The combination of IoT and AI has the potential to transform many human activities in several domains (retail, industry, home, . . .) by improving safety, productivity and user experience. Despite the wide range of possible applications in several fields, the papers identified in the systematic review mainly focus on the development of solutions oriented to smart homes, for example, to improve comfort, provide assistance to older adults or impaired users, or monitor energy consumption. Fewer applications address aspects concerning healthcare, smart agriculture or, more in general, smart environments, such as offices or buildings. Of the 50 articles analysed, 37 propose written or spoken language to control and monitor home appliances and sensors. For example, one system [1] uses a chatbot accessed through Facebook Messenger to control a variety of smart home devices, such as sensors for noise and gas, a door sensor, and a relay to control lights. Moreover, some contributions consider the interaction with a physical robot. Among the papers in the smart home domain, five of them [2, 13, 28, 56, 60] are specifically designed to provide support for older adults or people with disabilities. For example, one paper [13] discusses the development of a human-robot-smart environment interaction interface for ambient assisted living, which was able to control lights, curtains, a radio, an air conditioner, and temperature. In another study, [56] the interaction with a robot is exploited focusing on helping older adults or movement-impaired people in everyday tasks, including the monitoring of vital parameters. Moreover, in [66] the authors propose a smart mirror that can act as a virtual assistant making it possible to query web services (e.g., weather, news) or to execute actions on home appliances. Only two papers discuss healthcare applications. One study [11] discusses the development of an IoT-AI powered healthcare kit, which was able to measure blood pressure, temperature, oxygenation, and heart rate. In such a solution, this information can only be retrieved from a classical web interface, while a chatbot is used to make diagnoses based on the description of the user's symptoms. The other one [35] describes a chatbot to obtain information about the user's heart rate and provides the possibility to book a doctor's appointment or set reminders for taking medications. Five papers have been classified as "smart environments" since they do not focus on particular domains but propose conversational interfaces to control, in general, IoT devices. For example, one contribution [34] proposes a virtual assistant for student laboratories that can answer general questions and, in addition, can control the status of laboratory instruments. While [73] proposes a system to perform multiple operations contained in one complex natural language command (in Chinese) for three main domains (agriculture, industry and smart home). The remaining articles present applications in different fields such as a chatbot application for supporting smart urban agriculture through the measurement of soil moister, overall humidity and temperature, and programming or remote controlling irrigation [30], or a chatbot to command a 3D printer [50], including the upload of the 3D model to print, the status and progress of the printing and to guide the user through the whole process.

3.3 Interaction Device and Modality

The interaction modality and the devices with which conversational agents can be accessed play an important role from the point of view of both accessibility and usability. Text-only interaction modes favour privacy, the possibility of keeping track of the conversation (and possibly carrying it on at different times), and made possible using graphic elements (e.g., choice buttons) to speed up and simplify the interaction. Voice modes emphasize the possibility of "hands-free" interaction, which is primarily useful for impaired with limited movement possibilities, and more generally for natural and more immediate interaction. Among the articles analysed, the distribution of the interaction modality is almost equally divided between vocal (16 articles) and textual (17 articles) interaction. Most of the solutions using only voice rely on devices such as Alexa [29, 35], Google Assistant [23, 35, 45, 56] or custom hardware implementations [34, 38, 60, 62, 66] (e.g., using Raspberry Pi with microphone). In one work [13] the authors use a humanoid robot (a Pepper one) as the interface for receiving commands and estimating the user's emotions through voice or facial expressions. It is worth noting that solutions that use Google Assistant can be used either from stationary devices such as Google Home or from any Android-enabled device.

Regarding text-based agents, custom web platforms [6, 11, 20, 67] and commercial messaging applications, like Facebook Messenger [1, 33], Slack [40], Telegram [50, 55], Line [30, 41, 68] and WeChat [57], allow an interaction independent from the device (smartphone or desktop), since all they require is an internet connection to access the application or the web site. Moreover, integration with these commercial applications requires less effort than developing customised interfaces. Thus, the integration with well-known applications, perhaps already used for everyday messaging, makes access and use of the chatbot straightforward. Finally, custom smartphone [73] and desktop applications [32, 37, 61] seem to be related to early prototypes (e.g., [61]) or to integrate the chatbot into solutions with other functionalities (e.g., visualisations of energy consumption [73]). The remaining papers present multimodal solutions, such as text and voice (8), voice and gestures (3) and voice and touch (3). Text and voice solutions maintain the positive aspects of the text-based ones, the vocal interaction is constrained to the click of a button to activate and deactivate the microphone, thus requiring the use of the hands to be used. Among the contributions analysed, most are accessible through custom web interfaces [7, 19, 27, 43, 59] or smartphone applications [15, 54]. Salvi et al. [65] preferred to split the two modalities, making the conversational agent accessible both from Google Assistant and the Telegram application. Concerning works that integrate voice and gesture, two of them combine the voice command with a hand movement to perform actions on home devices. Anbarasan et al. [2] use a Kinect and their solution captures voice and gestures, using the combination of both to execute the command on a device; while [28] also uses a Kinect but the voice commands and the gestures are managed separately, thus is not possible to use them in conjunction. Finally, [42] exploits a wristband to capture gestures around the environment and a Bluetooth-enabled wireless earpiece for getting voice commands, combining both to execute commands (e.g., "turn on that light"). The possibility to simultaneously use touch and voice is exploited in an Alexa skill [9] allowing the user to switch between the two modalities while interacting with an Amazon device that presents a display, while the other works [14, 22] exploit the two modalities separately, so the user must choose whether to interact using either voice or touch. One paper [18] describes interaction via three different modalities: voice, touch and BCI (Brain Computer Interface), designed for users with mobility limitations for controlling smart home devices. This solution uses a mobile Android application with voice recognition and a dialogue system, with the additional possibility to alternatively use a NeuroSky MindWave mobile headset as input. Finally, two papers [64, 71] do not describe the mode of interaction with the developed systems since they present a generic application of natural language to smart environment control.

3.4 Conversational Breakdowns

The effective handling of conversational breakdowns is an essential aspect to consider in conversational agents. Breakdowns occur when a conversational agent fails to understand user inputs, leading to frustration, loss of credibility, and dropping the conversation [21, 50, 51]. This is especially crucial in task-oriented chatbots where the user has a specific goal in mind. Therefore, repair strategies are necessary to recover from breakdowns and continue the conversation. There are several methodologies to repair breakdowns, including presenting alternative options, highlighting keywords, and rephrasing [5]. Out of the 50 articles reviewed, only 12 describe how breakdowns are handled. The breakdown repair in most of these articles involves rephrasing the command to make it clearer [18, 19, 20, 28, 59, 67]. In [27], since it allows the creation of complex rules using a single input (which may include more than one trigger and one action), the chatbot asks the user to rephrase only the part of the input that was not understood, while showing to the user which part has been correctly classified. In the case of simple commands (e.g., entering a single trigger or action), rephrasing the entire input is requested. When a breakdown occurs, a solution [41] shows the user a help message containing possible chatbot commands, while [54] opens a Google Search page with the text of the input as a query. Campagna et al. [15] provides a list of possible matches with available intents, and in the case of none being valid, the system asks for rephrasing. Then, if the user selects one of the suggested intents, the chatbot will add the input text that generated the breakdown to the training phrases for that intent. Kim et al. [45] uses a "disambiguation strategy" by asking the user additional questions in case of incomplete or ambiguous input, while the authors do not describe the chatbot behaviour when the input does not match any intent. Since the chatbot presented in [45] is dedicated to the modification of trigger-action rules, the disambiguation strategy is applied when it is unclear whether a trigger or action is to be added or modified, or when it is unclear which internal parameter (to the trigger or action) is to be changed. Follow-up messages are then sent to identify the user's intent uniquely. Oumard et al. [60] use the Rasa Fallback policy, which consists of a two-stage breakdown resolution[1]: when a message is classified with low confidence, the user is asked to confirm the most probable intent. If the users do not confirm the intent, they are asked to rephrase the message. Then, if the new message is classified with low confidence, the chatbot asks again for confirmation. Finally, if the user rebuts again, a breakdown message is sent, and the conversation state is reset.

Both quantitative and qualitative analyses of the reviewed articles demonstrate the need for more research effort to address conversational breakdown resolution. Referring to the discussion on repair strategies for conversational breakdowns [5], the twelve articles that address breakdown issues do not use particularly efficient techniques. Simply asking to rephrase the input can be considered simple and quick as it immediately highlights a lack of understanding on the part of the chatbot but does not provide particular help to the user to recover from the error. Instead, the methodology reported in [15] would appear to be more efficient, as the chance of choosing between different options makes the chatbot's possibilities immediately clear and less input is required by the user at the expense, however, of less natural interaction. Furthermore, the resulting possibility of re-training the chatbot is useful in preventing future breakdowns. In summary, the presence of conversational breakdowns in chatbots creates significant obstacles to achieving an optimal user experience, especially in task-oriented chatbots. After the literature analysis, it becomes apparent that the implementation of effective repair strategies is crucial to allow users to overcome such breakdowns and continue with the conversation. Although most of the articles propose rephrasing as a solution, there is a need for further research in devising and applying more efficient techniques. Research suggests [5] that presenting alternative options, emphasizing keywords, and providing disambiguation strategies could potentially improve the chatbot's ability to handle breakdowns proficiently. A relevant contribution [24] has investigated learning mechanisms to minimize conversational breakdowns in human-agent interaction in the manufacturing industry. It compares three scenarios where, after a successful repair, the learning burden is assigned to the agent (it must adapt to the user's terminology), the users (they must adapt to the agent's terminology), or both (the agent adapts to the user's terminology, but the user should use agent's standard terms to reduce the possibility of breakdown). Participants (N=26) showed a preference for distributing the learning responsibility with the conversational agent, with a likelihood of 61.3%. However, participants expressed that allowing the user to manage the learning burden, rather than sharing it with the agent, is a more efficient approach. Therefore, additional research should tackle the design and evaluation of more successful strategies for repairing conversational breakdowns in a variety of scenarios in IoT settings.

[1]https://rasa.com/docs/rasa/fallback-handoff/

3.5 User Experience Evaluation Methods

One key point has been to analyse how users have been involved in assessing their experience with the proposed conversational interfaces for controlling IoT ecosystems. Among the articles selected for this review, only 11 present an evaluation of the usability and overall experience of the conversational systems. The remaining studies only evaluate the systems in terms of computational performance, using classic Machine Learning evaluation metrics such as accuracy, F1 score and loss. The evaluation methods used in such eleven studies include Likert scales, measures of task success and failure, NASA-TLX and SUS evaluations, which are commonly used to assess the usability of the systems. The NASA-TLX and SUS evaluation methods were used in a study [2] that assessed the usability, accuracy and workload of the system for older adults and compared it to Google Home and Amazon Echo. One study [60] also applied the UEQ questionnaire, which measures the user's overall experience with the system.

One example study using the Likert scale [16] evaluated the use of a voice assistant for cybersecurity tasks. Participants were asked to rate the difficulty of the tasks after performing them with both traditional methods and the voice assistant. Task success and failure measures and thinking aloud were used in [67] which recorded the number of errors and help requests during tasks completion, while in [42] users were asked to vocalize their thoughts on the interaction after completing each of the three proposed tasks. Sometimes more specific evaluation criteria have been used such as custom questionnaires, or specific metrics such as the number of conversational turns. One study [19] used metrics for evaluating Perceived Effectiveness and Fun (PEF), alongside the total number of messages, the number of desired automations expressed by the user and the number of satisfying automations identified by the associated recommendation system. Another example [59] used a custom questionnaire with Likert scales to evaluate the virtual assistant's speech recognition, conflict resolution, interaction with the virtual assistant, and user-friendliness. Custom questionnaires were also used in some studies, such as [34], which asked specific questions about the user's experience with the virtual assistant (e.g., "Did you enjoy the overall experience?" or "Would you enjoy my services on a daily basis?"). The number of conversational turns and task time were measured in [27, 45] where participants performed some tasks using two different approaches (they compared form-based interfaces and a conversational one in creating and modifying automations specified in terms of trigger-action rules). Finally, [9] used a between-subject protocol with two groups for comparing the system proposed along with the standard solution, measuring the success and error rate, the execution time and the number of errors. Then the users were asked to fill out a SUS and a UEQ questionnaire.

Table 2: User test details

Paper	Procedure (n. tasks)	Evaluation methods	N. user	Users age	Users type	Users gender
[16]	Predefined Tasks (1)	Likert scale, task time	2	No info	1 cyber security expert; 1 IoT device owner	2 m
[34]	Predefined Tasks (8)	Custom questionnaire, task success rate, task failure rate	15	#1: 10 (18-24) #2: 3 (25-28) #3: 1 (40), 1 (50)	#1: student non-tech related; #2: researcher; #3: professors;	6 m, 9 f
[59]	Randomly generated scenario	Custom questionnaire, Likert scale	13	No info	No info	No info
[2]	List of commands to try	NASA-TLX, SUS	20	From 65 to 80	Older adults	8 m, 12 f
[27]	Predefined Tasks (4)	Custom questionnaire, Likert scale, n. conversational turn, task error rate	10	From 20 to 30 (avg. 27)	2 with programming exp., 8 without exp.	8 m, 2 f
[42]	Predefined Tasks (3)	Custom questionnaire	10	From 21 to 30	3 with smart homes exp., 7 without exp.	6 m, 4 f
[45]	Predefined tasks (3)	Custom interview, NASA-TLX, n. conversational turns, SUS, task time	20	From 20 to 42 (avg. 27)	10 with IoT platform exp., 17 with chatbot experience	15 m, 5 f
[60]	Predefined Tasks (20)	Custom questionnaire, Likert scale, UEQ questionnaire	7	From 19 to 53 (avg. 38)	No info	5 m, 2 f
[67]	Predefined Tasks (-)	Completion time, n. help requests, n. errors made, thinking aloud	5	From 25 to 40	No info	No info
[19]	15 minutes free task	n. messages, n. of need expressed by the user, n. of automations identified, open questions, Perceive Effectiveness and Fun (PEF)	8	From 24 to 30 (avg. 26)	All students with computer science backgrounds	5 m, 3 f
[9]	Predefined Tasks (5)	Between-subject protocol with two groups (1 and 2), task success rate, execution time, number of errors, SUS, UEQ.	20	Group 1: avg. 32.5 Group 2: avg. 38.8	19 without programming experience	10m, 10f

User study sample sizes ranged from two [16] to twenty [2, 9, 45] participants, with an average of 11.81 participants. There were two studies with 15 users [34, 59], and two with ten users [27, 42], more details are in Table 2. Still, regarding the users involved, eight out of eleven studies reported a user age range that varied from a minimum of 18 years old to 53 ([16] does not present any information about the users' age, while [59] does not report any user information at all). In one case [2] users' age ranged from 65 to 80 since the study focuses on solutions for older adults. Another specific category of users was considered in [60] where seven blind users were involved. Moreover, seven out of eleven studies specified the sex distribution among participants with an overall percentage of 41.96% (47 users) females and 58.03% (65 users) males. In general, the numbers suggest that little weight has been given to the evaluation of the user's experience during interaction with conversational systems in IoT settings, while more attention has been given to the computational performance of the implemented systems. Furthermore, it can be noted that there is no single shared evaluation methodology among researchers in this field who often use general-purpose evaluation methods, such as NASA-TLX, SUS, and thinking aloud. Alongside these "classical" methods, specific metrics for evaluating information related to purely conversational aspects have been used, such as in [27] and [45], which tracked the number of conversational turns (for task-oriented chatbots, a lower number of turns may imply a higher efficiency) or in [19] the number of messages.

The variety of evaluation methods applied, ranging from custom metrics to general methods (such as SUS) highlights the lack of specific methodologies for evaluating conversational systems. In this perspective, the Subjective Assessment of Speech System Interfaces (SASSI) [31] can be used to assess user perception of some aspects of conversational interfaces using the vocal modality. The SASSI questionnaire presents generic items such as "The system is accurate" or "The system is easy to use", and specific conversational items such as "It is clear how to speak to the system" or "I sometimes wondered if I was using the right word". More specifically for conversational interfaces, the BOT Usability Scale (BUS-15) [12] has been recently put forward, emphasising aspects such as the

perceived quality of the conversation (e.g., "I find that the chatbot understands what I want and helps me achieve my goal") and the perceived quality of the chatbot functions (e.g., "The chatbot was able to keep track of context"). As its authors themselves point out, the questionnaire needs further testing and validation.
Finally, we would like to underline that none of the eleven articles carried out user tests in a real uncontrolled environment (i.e. in-the-wild) since all tests were performed in controlled and supervised environments by the researchers themselves. Table 2 shows the user test details for the papers discussed above.

4 DISCUSSION AND CONCLUSIONS

The combination of IoT and AI technologies has the potential to transform human activities in various domains by improving comfort, assistance and productivity, enriching future peoples' smart living. Given the recent advantages in natural language processing, in this systematic review, we analysed 50 articles (selected from an initial set of 3177) that focused on the development of chatbots and virtual assistants for controlling intelligent environments in IoT settings.

Employed technologies. We observed that the primary technologies employed to develop chatbots in these systems were Artificial Intelligence-based methods (e.g., Reinforcement Learning, Transformers, SVM), with many works using frameworks, such as Rasa and Dialogflow; while a smaller, but still relevant number of works has implemented rule-based systems. A large portion of the articles do not present a sufficiently in-depth description to reproduce the work proposed, and only three articles publicly share the implementation code. Moreover, current solutions seem limited in terms of the support for flexibly creating automations that involve multiple connected smart objects in a conversational way, since the current solutions mainly focus on modifying existing automations or on creating simple ones (e.g., with only one trigger and one action).

Despite AI solutions being more efficient compared to rule-based ones, they are still limited in terms of flexibility in understanding user input and managing conversational flow in case of "unexpected" user behaviour. The potential unlocked by recent advancements in NLP with LLMs (e.g., GPT-4) opens up a range of new possibilities through prompting techniques, both by reducing developer workload (e.g., there's no longer the need to define intents and entities by creating training datasets from scratch, and there is no need to predict and define possible user conversational paths) and by enhancing the capabilities of conversational agents. These agents exhibit linguistic abilities that enable autonomous and self-sufficient dialogue management, independent of pre-defined pathways.

Evaluation metrics. Overall, from our analysis, it emerges that despite the significant technological evolution in the areas of conversational interfaces and IoT, their integration is still an open issue, and several areas need more research to better exploit the possibilities of conversational interfaces in smart contexts, in particular in terms of user-centred approaches. Indeed, only a small number of studies considered user-based evaluation of the proposed solutions. The evaluation methods for these systems varied, with some studies evaluating usability and user experience (about 20% of the papers) using methods such as Likert scales, custom questionnaires, and the NASA-TLX and SUS, while the remaining works focused only on evaluating the computational performance of the systems using metrics such as accuracy, F1 score, and loss. However, some contributions have used metrics aimed at assessing, specifically, the conversational experience, for instance by considering the number of conversational shifts. Thus, the lack of specific metrics is clear, as is the consequent need to research and identify evaluation methods for conversational agents. In this perspective, the SASSI [31] questionnaire was put forward for the evaluation of speech interfaces, whereas more recent studies such as [12] have started to consider the evaluation of more specific conversational aspects, but further validation is necessary to understand whether they are the optimal solution in this case. In addition, there is a clear lack of studies in real contexts that can provide more information on the actual user experience over time during daily activities (i.e., in-the-wild studies).

Application domains. The majority of studies focused on the smart home domain, which is the one with the most immediate impact on people's lives, improving comfort (e.g., using routines or remotely controlling devices), assisting older or impaired users (e.g., through vocal commands and robots), and monitoring energy consumption in smart homes. Other areas of interest relate to smart health, giving the possibility to monitor vital parameters or perform diagnostics. The use of conversational interfaces in home automation systems has the potential to enhance functionality, usability, and security, and to provide personalised experiences for users. Nevertheless, there is a need for more research on the user experience evaluation of chatbots in real-world scenarios, as studies in this area were usually conducted in controlled environments.

Conversational limitations and areas for improvement. One area for improvement lies in the enhancement of conversational agent capabilities for smart home automation (also called "routines" or "trigger-action rules"). Existing commercial solutions (e.g., Alexa and Google Assistant) do not allow users to create automations using the natural language but only through classic buttons interfaces available in the smartphone applications, while research solutions (some of them presented in section 3.1) primarily revolve around creating or modifying simple automations, but they lack the flexibility to handle complex tasks involving multiple interconnected smart objects. Future research should be dedicated to empowering conversational agents to act as personal assistants, guiding users through the configuration and personalization of smart environments, providing comprehensive support and insights into the possibilities and limitations of sensors and smart objects, and empowering users to make informed decisions and achieve their desired outcomes.

An additional area to explore is system transparency and the ability to provide explanations. This factor becomes particularly important when dealing with agents that enable the creation and execution of automations. Specifically, users should have access to information about the automations and the capability to seek explanations for system behaviour (e.g., addressing basic requests such as "Why did you turn on the thermostat?" or "Why is it so hot now?").

Another potential area for improvement lies in the development of recommendation systems based on acquired sensor data and user preferences. By considering user habits, preferences, and goals, the system can suggest relevant automations that align with each user's unique requirements. This personalised approach can greatly enhance smart environments' usability and overall user experience.

One further relevant research direction should focus on mitigating errors and improving user interaction. Breakdowns or fallbacks, instances where the agent fails to understand user input, can easily lead to user frustration. Innovative strategies can be employed to address this issue. For instance, rather than merely asking users to repeat their input, the agent could automatically rephrase the sentence (e.g., using language models or rule-based algorithms for replacing terms and modifying the syntactic structure) and seek user confirmation. Furthermore, misunderstandings leading to breakdowns can be repurposed into training data. This allows the agent to learn from its mistakes and expand its vocabulary, potentially enhancing the agent's performance and providing a more satisfying user experience over time.

Challenges and possibilities for the future. Lastly, as previously mentioned to some extent, recent advances in Natural Language Processing (NLP) have greatly improved conversational capabilities and language-related tasks. Large Language Models (LLMs), such as GPT-3 and ChatGPT, have been instrumental in achieving those improvements. These models have enabled text generation, translation across various languages, text style rewriting, question answering, and more. With the use of prompt engineering techniques, they can perform zero-shot and few-shot tasks from a small set of examples, without the need of training or fine-tuning a new model [47]. The emergence of LLMs presents a new challenge in managing smart environments through their advanced capabilities of understanding and generating natural language as well as maintaining context during the entire conversation, making it very close to a human and consequently more natural and less frustrating. However, applying these models to real-world problems requires further research since they exhibit probabilistic and not deterministic behaviour so even a small change in the input may lead to the generation of an unexpected or inaccurate output. Initial explorations have started to investigate how combining prompt engineering and high-level function libraries enables ChatGPT to adapt to different robotics tasks, using natural language communication to control the movements of objects, such as a robotic arm or a small drone [70]. Another possibility of using LLM, without granting them full control, could involve the combined use of well-known and validated systems (e.g., Rasa) with dedicated LLM for specific tasks [26]. In this way, it would be possible to develop efficient and reliable systems, reducing the risks associated with LLM while simultaneously enhancing the capabilities of more "conventional" systems. Overall, LLMs are paving the way for numerous possibilities and different developments. However, this progress necessitates further research, not only to adapt these models to the Internet of Things (IoT) context but primarily to ensure their safe and practical use. As these models will have an impact on people interacting with real-world objects, it is essential to address concerns related to privacy, security, and ethical implications.

Limitations. The aim of this systematic literature review is to provide a comprehensive framework for understanding how conversational agents have been employed to control IoT devices in intelligent environments, analysed through an HCI perspective. It can be useful to note that the validity of this study may be influenced and limited by different factors. Specifically, the relevant studies were gathered from digital libraries (as detailed in Section 2.4), which curate the most representative conference papers and journal articles for our research purposes. However, these libraries, while valuable, are not exhaustive. This aspect could potentially affect the comprehensiveness of our research findings.

REFERENCES

[1] Sakib Ahmed, Debashish Paul, Rubaiya Masnun, Minhaz Uddin Ahmed Shanto, and Tanjila Farah. 2020. Smart Home Shield and Automation System Using Facebook Messenger Chatbot. In *2020 IEEE Region 10 Symposium (TENSYMP)*, 1791–1794. DOI:https://doi.org/10.1109/TENSYMP50017.2020.9230716

[2] Anbarasan and Jeannie S.A. Lee. 2018. Speech and Gestures for Smart-Home Control and Interaction for Older Adults. In *Proceedings of the 3rd International Workshop on Multimedia for Personal Health and Health Care* (HealthMedia'18), Association for Computing Machinery, New York, NY, USA, 49–57. DOI:https://doi.org/10.1145/3264996.3265002

[3] Carmelo Ardito, Giuseppe Desolda, Rosa Lanzilotti, Alessio Malizia, and Maristella Matera. 2020. Analysing trade-offs in frameworks for the design of smart environments. *Behaviour & Information Technology* 39, 1 (January 2020), 47–71. DOI:https://doi.org/10.1080/0144929X.2019.1634760

[4] Parvaneh Asghari, Amir Masoud Rahmani, and Hamid Haj Seyyed Javadi. 2019. Internet of Things applications: A systematic review. *Computer Networks* 148, (January 2019), 241–261. DOI:https://doi.org/10.1016/j.comnet.2018.12.008

[5] Zahra Ashktorab, Mohit Jain, Q. Vera Liao, and Justin D. Weisz. 2019. Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, Glasgow Scotland Uk, 1–12. DOI:https://doi.org/10.1145/3290605.3300484

[6] Cyril Joe Baby, Faizan Ayyub Khan, and J. N. Swathi. 2017. Home automation using IoT and a chatbot using natural language processing. In *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, 1–6. DOI:https://doi.org/10.1109/IPACT.2017.8245185

[7] Cyril Joe Baby, Nalin Munshi, Ankit Malik, Kunal Dogra, and R. Rajesh. 2017. Home automation using web application and speech recognition. In *2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS)*, 1–6. DOI:https://doi.org/10.1109/ICMDCS.2017.8211543

[8] Sara Bardi and Claudio Enrico Palazzi. 2022. Smart Hydroponic Greenhouse: Internet of Things and Soilless Farming. In *Conference on Information Technology for Social Good*, ACM, Limassol Cyprus, 212–217. DOI:https://doi.org/10.1145/3524458.3547221

[9] Barbara Rita Barricelli, Daniela Fogli, Letizia Iemmolo, and Angela Locoro. 2022. A Multi-Modal Approach to Creating Routines for Smart Speakers. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (AVI 2022), Association for Computing Machinery, New York, NY, USA. DOI:https://doi.org/10.1145/3531073.3531168

[10] Parabattina Bhagath, Samanvi Parisa, Sasi Dinesh Reddy, and Fareeda Banu. 2021. An Android based Mobile Spoken Dialog System for Telugu language to control Smart appliances. In *2021 IEEE XXVIII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, 1–4. DOI:https://doi.org/10.1109/INTERCON52678.2021.9532783

[11] Shreya Bhutada, Akshata Singh, Kaushiki Upadhyaya, and Purvika Gaikar. 2021. Ru-Urb IoT-AI powered Healthcare Kit. In *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 417–422. DOI:https://doi.org/10.1109/ICICCS51141.2021.9432257

[12] Simone Borsci, Alessio Malizia, Martin Schmettow, Frank van der Velde, Gunay Tariverdiyeva, Divyaa Balaji, and Alan Chamberlain. 2022. The Chatbot Usability Scale: the Design and Pilot of a Usability Scale for Interaction with AI-Based Conversational Agents. *Pers Ubiquit Comput* 26, 1 (February 2022), 95–119. DOI:https://doi.org/10.1007/s00779-021-01582-9

[13] Ha-Duong Bui and Nak Young Chong. 2018. An Integrated Approach to Human-Robot-Smart Environment Interaction Interface for Ambient Assisted Living. In *2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, 32–37. DOI:https://doi.org/10.1109/ARSO.2018.8625821

[14] Julio Cabrera, María Mena, Ana Parra, and Eduardo Pinos. 2016. Intelligent assistant to control home power network. In *2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)*, 1–6. DOI:https://doi.org/10.1109/ROPEC.2016.7830531

[15] Giovanni Campagna, Rakesh Ramesh, Silei Xu, Michael Fischer, and Monica S. Lam. 2017. Almond: The Architecture of an Open, Crowdsourced, Privacy-Preserving, Programmable Virtual Assistant. In *Proceedings of the 26th International Conference on World Wide Web* (WWW '17), International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 341–350. DOI:https://doi.org/10.1145/3038912.3052562

[16] Jeffrey S. Chavis, Malcom Doster, Michelle Feng, Syeda Zeeshan, Samantha Fu, Elizabeth Aguirre, Antonio Davila, Kofi Nyarko, Aaron Kunz, Tracy Herriotts,

Daniel Syed, Lanier Watkins, Anna Buczak, and Aviel Rubin. 2021. A Voice Assistant for IoT Cybersecurity. In *2021 IEEE Integrated STEM Education Conference (ISEC)*, 165–172. DOI:https://doi.org/10.1109/ISEC52395.2021.9764005

[17] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. 2017. A Survey on Dialogue Systems: Recent Advances and New Frontiers. *SIGKDD Explor. Newsl.* 19, 2 (November 2017), 25–35. DOI:https://doi.org/10.1145/3166054.3166058

[18] Miguel Angel Contreras-Castañeda, Juan Antonio Holgado-Terriza, Gonzalo Pomboza-Junez, Patricia Paderewski-Rodríguez, and Francisco Luis Gutiérrez-Vela. 2019. Smart Home: Multimodal Interaction for Control of Home Devices. In *Proceedings of the XX International Conference on Human Computer Interaction*, ACM, Donostia Gipuzkoa Spain, 1–8. DOI:https://doi.org/10.1145/3335595.3335636

[19] Fulvio Corno, Luigi De Russis, and Alberto Monge Roffarello. 2020. HeyTAP: Bridging the Gaps Between Users' Needs and Technology in IF-THEN Rules via Conversation. In *Proceedings of the International Conference on Advanced Visual Interfaces*, ACM, Salerno Italy, 1–9. DOI:https://doi.org/10.1145/3399715.3399905

[20] Fulvio Corno, Luigi De Russis, and Alberto Monge Roffarello. 2021. From Users' Intentions to IF-THEN Rules in the Internet of Things. *ACM Trans. Inf. Syst.* 39, 4 (August 2021). DOI:https://doi.org/10.1145/3447264

[21] Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. "What can i help you with?": infrequent users' experiences of intelligent personal assistants. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, ACM, Vienna Austria, 1–12. DOI:https://doi.org/10.1145/3098279.3098539

[22] Hanif Fakhrurroja, Ahmad Musnansyah, Muhammad Dewan Satriakamal, Bima Kusuma Wardana, Rizal Kusuma Putra, and Dita Pramesti. 2022. Dialogue System based on Reinforcement Learning in Smart Home Application. In *Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications*, ACM, Virtual Event Indonesia, 146–152. DOI:https://doi.org/10.1145/3575882.3575911

[23] Renato Ferrero, Mohammad Ghazi Vakili, Edoardo Giusto, Mauro Guerrera, and Vincenzo Randazzo. 2019. Ubiquitous fridge with natural language interaction. In *2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, 404–409. DOI:https://doi.org/10.1109/RFID-TA.2019.8892025

[24] Mina Foosherian, Samuel Kernan Freire, Evangelos Niforatos, Karl A. Hribernik, and Klaus-Dieter Thoben. 2022. Break, Repair, Learn, Break Less: Investigating User Preferences for Assignment of Divergent Phrasing Learning Burden in Human-Agent Interaction to Minimize Conversational Breakdowns. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*, ACM, Lisbon Portugal, 151–158. DOI:https://doi.org/10.1145/3568444.3568454

[25] Anna Förster and Julian Block. 2022. User Adoption of Smart Home Systems. In *Conference on Information Technology for Social Good*, ACM, Limassol Cyprus, 360–365. DOI:https://doi.org/10.1145/3524458.3547118

[26] Simone Gallo, Alessio Malizia, and Fabio Paternò. Towards a Chatbot for Creating Trigger-Action Rules based on ChatGPT and Rasa. In *Workshops, Work in Progress Demos and Doctoral Consortium at IS-EUD 2023*, CEUR-WS.org, Cagliari. Retrieved from https://ceur-ws.org/Vol-3408/short-s4-01.pdf

[27] Simone Gallo and Fabio Paterno. 2022. A Conversational Agent for Creating Flexible Daily Automation. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces* (AVI 2022), Association for Computing Machinery, New York, NY, USA. DOI:https://doi.org/10.1145/3531073.3531090

[28] Alexandru Florin Gavril, Mihai Trascau, and Irina Mocanu. 2017. Multimodal Interface for Ambient Assisted Living. In *2017 21st International Conference on Control Systems and Computer Science (CSCS)*, 223–230. DOI:https://doi.org/10.1109/CSCS.2017.38

[29] Steven Guamán, Adrián Calvopiña, Pamela Orta, Freddy Tapia, and Sang Guun Yoo. 2018. Device Control System for a Smart Home using Voice Commands: A Practical Case. In *Proceedings of the 2018 10th International Conference on Information Management and Engineering*, ACM, Salford United Kingdom, 86–89. DOI:https://doi.org/10.1145/3285957.3285977

[30] Reza Gunawan, Ichan Taufik, Edi Mulyana, Opik T Kurahman, Muhammad Ali Ramdhani, and Mahmud. 2019. Chatbot Application on Internet Of Things (IoT) to Support Smart Urban Agriculture. In *2019 IEEE 5th International Conference on Wireless and Telematics (ICWT)*, 1–6. DOI:https://doi.org/10.1109/ICWT47785.2019.8978223

[31] Kate S. Hone and Robert Graham. 2000. Towards a tool for the Subjective Assessment of Speech System Interfaces (SASSI). *Nat. Lang. Eng.* 6, 3 & 4 (September 2000), 287–303. DOI:https://doi.org/10.1017/S1351324900002497

[32] Cheng-Chi Huang, Alan Liu, and Pei-Chuan Zhou. 2015. Using Ontology Reasoning in Building a Simple and Effective Dialog System for a Smart Home System. In *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 1508–1513. DOI:https://doi.org/10.1109/SMC.2015.267

[33] Husain Husain, Herlinda Herlinda, Hasriani Hasriani, Ahyuna Ahyuna, Kasmawaru Kasmawaru, and Ahmad Ahmad. 2021. Increasing the Smart Home Automation by using Facebook Messenger Application. In *2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS)*, 1–6. DOI:https://doi.org/10.1109/ICORIS52787.2021.9649442

[34] Giancarlo Iannizzotto, Lucia Lo Bello, Andrea Nucita, and Giorgio Mario Grasso. 2018. A Vision and Speech Enabled, Customizable, Virtual Assistant for Smart Environments. In *2018 11th International Conference on Human System Interaction (HSI)*, 50–56. DOI:https://doi.org/10.1109/HSI.2018.8431232

[35] Andrej Ilievski, Dimitri Dojchinovski, and Marjan Gusev. 2019. Interactive Voice Assisted Home Healthcare Systems. In *Proceedings of the 9th Balkan Conference on Informatics* (BCI'19), Association for Computing Machinery, New York, NY, USA. DOI:https://doi.org/10.1145/3351556.3351572

[36] Marco Jahn, Marc Jentsch, Christian R. Prause, Ferry Pramudianto, Amro Al-Akkad, and Rene Reiners. 2010. The Energy Aware Smart Home. In *2010 5th International Conference on Future Information Technology*, IEEE, Busan, Korea (South), 1–8. DOI:https://doi.org/10.1109/FUTURETECH.2010.5482712

[37] Akhil Jain, Poonam Tanwar, and Stuti Mehra. 2019. Home Automation System using Internet of Things (IOT). In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 300–305. DOI:https://doi.org/10.1109/COMITCon.2019.8862201

[38] Farzeem D. Jivani, Manohar Malvankar, and Radha Shankarmani. 2018. A Voice Controlled Smart Home Solution With a Centralized Management Framework Implemented Using AI and NLP. In *2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)*, 1–5. DOI:https://doi.org/10.1109/ICCTCT.2018.8550972

[39] Daniel Jurafsky and James H Jr Martin. Chatbots & Dialogue Systems. In Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Third Edition draft). Standford Edu. Retrieved from https://web.stanford.edu/$\sim$jurafsky/slp3/

[40] Charbel El Kaed, Andre Ponnouradjane, and Dhaval Shah. 2018. A Semantic Based Multi-Platform IoT Integration Approach from Sensors to Chatbots. In *2018 Global Internet of Things Summit (GIoTS)*, 1–6. DOI:https://doi.org/10.1109/GIOTS.2018.8534520

[41] Yao-Chiang Kan, Hsueh-Chun Lin, Han-Yu Wu, and Junghsi Lee. 2020. LoRa-Based Air Quality Monitoring System Using ChatBot. In *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 1561–1565. Retrieved from https://ieeexplore.ieee.org/document/9306463

[42] Runchang Kang, Anhong Guo, Gierad Laput, Yang Li, and Xiang "Anthony" Chen. 2019. Minuet: Multimodal Interaction with an Internet of Things. In *Symposium on Spatial User Interaction* (SUI '19), Association for Computing Machinery, New York, NY, USA. DOI:https://doi.org/10.1145/3357251.3357581

[43] Maksym Ketsmur, Ant
'onio Teixeira, Nuno Almeida, Samuel Silva, and M
'ario Rodrigues. 2018. Conversational Assistant for an Accessible Smart Home: Proof-of-Concept for Portuguese. In *Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion* (DSAI 2018), Association for Computing Machinery, New York, NY, USA, 55–62. DOI:https://doi.org/10.1145/3218585.3218594

[44] Lachaux Killian, Maitre Julien, Bouchard Kevin, Lussier Maxime, Bottari Carolina, Couture Mélanie, Bier Nathalie, Giroux Sylvain, and Gaboury Sebastien. 2021. Fall Prevention and Detection in Smart Homes Using Monocular Cameras and an Interactive Social Robot. In *Proceedings of the Conference on Information Technology for Social Good*, ACM, Roma Italy, 7–12. DOI:https://doi.org/10.1145/3462203.3475892

[45] Sanghoon Kim and In-Young Ko. 2022. A Conversational Approach for Modifying Service Mashups in IoT Environments. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), Association for Computing Machinery, New York, NY, USA. DOI:https://doi.org/10.1145/3491102.3517655

[46] Barbara Kitchenham and Stuart Charters. 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering. 2, (January 2007).

[47] Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2023. Large Language Models are Zero-Shot Reasoners. Retrieved March 14, 2023 from http://arxiv.org/abs/2205.11916

[48] Satender Kumar, Bipin Kumar, Kamlesh Sharma, Rohan Raj, and Suresh Kumar. 2021. IoT Based Secured Home Automation System Using NLP. In *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 1–5. DOI:https://doi.org/10.1109/ICAECA52838.2021.9675570

[49] Lee Hoi Leong, Shinsuke Kobayashi, Noboru Koshizuka, and Ken Sakamura. 2005. CASIS: A Context-Aware Speech Interface System. In *Proceedings of the 10th International Conference on Intelligent User Interfaces* (IUI '05), Association for Computing Machinery, New York, NY, USA, 231–238. DOI:https://doi.org/10.1145/1040830.1040880

[50] Shi Liu, Shahrier Erfan Harun, Florian Jasche, and Thomas Ludwig. 2021. Supporting the Onboarding of 3D Printers through Conversational Agents. In *Mensch Und Computer 2021* (MuC '21), Association for Computing Machinery, New York, NY, USA, 494–498. DOI:https://doi.org/10.1145/3473856.3474010

[51] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, San Jose California USA, 5286–5297. DOI:https://doi.org/10.1145/2858036.2858288

[52] Mithil K M, Khatri Bimal Mukesh Kumar, Lalit Sharma, Mohammed Zaki Sayeed Pasha, and Kallinath H D. 2018. An Interactive Voice Controlled Humanoid Smart Home Prototype Using Concepts of Natural Language Processing and Machine Learning. In *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, 1537–1546. DOI:https://doi.org/10.1109/RTEICT42901.2018.9012359

[53] Sinarwati Mohamad Suhaili, Naomie Salim, and Mohamad Nazim Jambli. 2021. Service chatbots: A systematic review. *Expert Systems with Applications* 184, (December 2021), 115461. DOI:https://doi.org/10.1016/j.eswa.2021.115461

[54] Amr El Mougy, Ahmed Khalaf, Hazem El Agaty, Mariam Mazen, Noureldin Saleh, and Mina Samir. 2017. Xenia: Secure and interoperable smart home system with user pattern recognition. In *2017 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)*, 47–52. DOI:https://doi.org/10.1109/IINTEC.2017.8325912

[55] Muhamad Muslih, Somantri, Dedi Supardi, Elpid Multipi, Yusup Maulana Nyaman, Aditya Rismawan, and Gunawansyah. 2018. Developing Smart Workspace Based IOT with Artificial Intelligence Using Telegram Chatbot. In *2018 International Conference on Computing, Engineering, and Design (ICCED)*, 230–234. DOI:https://doi.org/10.1109/ICCED.2018.00052

[56] Mahmoud Nasr, Fakhri Karray, and Yuri Quintana. 2020. Human Machine Interaction Platform for Home Care Support System. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 4210–4215. DOI:https://doi.org/10.1109/SMC42975.2020.9283095

[57] Trung Nguyen, Barth Lakshmanan, Chengjie Lin, Weihua Sheng, Ye Gu, Meiqin Liu, and Senlin Zhang. 2017. A Miniature Smart Home Testbed for Research and Education. In *2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, 1637–1642. DOI:https://doi.org/10.1109/CYBER.2017.8446621

[58] Chinedu Wilfred Okonkwo and Abejide Ade-Ibijola. 2021. Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence* 2, (2021), 100033. DOI:https://doi.org/10.1016/j.caeai.2021.100033

[59] Bauyrzhan Ospan, Nawaz Khan, Juan Augusto, Mario Quinde, and Kenzhegali Nurgaliyev. 2018. Context Aware Virtual Assistant with Case-Based Conflict Resolution in Multi-User Smart Home Environment. In *2018 International Conference on Computing and Network Communications (CoCoNet)*, 36–44. DOI:https://doi.org/10.1109/CoCoNet.2018.8476898

[60] Christina Oumard, Julian Kreimeier, and Timo G
"otzelmann. 2022. Implementation and Evaluation of a Voice User Interface with Offline Speech Processing for People Who Are Blind or Visually Impaired. In *Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments* (PETRA '22), Association for Computing Machinery, New York, NY, USA, 277–285. DOI:https://doi.org/10.1145/3529190.3529197

[61] Biswamoy Pattnaik, Sreya Dey, and B Jaganatha Pandian. 2021. A Secure and Interactive Home Automation System with Machine Learning Based Power Prediction. In *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*, IEEE, Kuala Lumpur, Malaysia, 1–6. DOI:https://doi.org/10.1109/i-PACT52855.2021.9696603

[62] Rohit Raj and Nandini Rai. 2018. Voice Controlled Cyber-Physical System for Smart Home. In *Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking* (Workshops ICDCN '18), Association for Computing Machinery, New York, NY, USA. DOI:https://doi.org/10.1145/3170521.3170550

[63] Amon Rapp, Lorenzo Curti, and Arianna Boldi. 2021. The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies* 151, (July 2021), 102630. DOI:https://doi.org/10.1016/j.ijhcs.2021.102630

[64] Eugenio Rubio-Drosdov, Daniel Díaz-Sánchez, Florina Almenárez, Patricia Arias-Cabarcos, and Andrés Marín. 2017. Seamless human-device interaction in the internet of things. *IEEE Transactions on Consumer Electronics* 63, 4 (November 2017), 490–498. DOI:https://doi.org/10.1109/TCE.2017.015076

[65] Sanket Salvi, V Geetha, and S Sowmya Kamath. 2019. Jamura: A Conversational Smart Home Assistant Built on Telegram and Google Dialogflow. In *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, 1564–1571. DOI:https://doi.org/10.1109/TENCON.2019.8929316

[66] SPSS Sirinayake, Dkak Dasanayake, Tmlu Rodrigo, Wuy Perera, MDJT Hansika Mahaadikara, and Surath Kahandawala. 2021. IOT-Based Intelligent Assistant Mirror For Smart Life & Daily Routine Using Raspberry PI. In *2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)*, 1–6. DOI:https://doi.org/10.1109/ICter53630.2021.9774788

[67] Evropi Stefanidi, Maria Korozi, Asterios Leonidis, and Margherita Antona. 2018. Programming Intelligent Environments in Natural Language: An Extensible Interactive Approach. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference* (PETRA '18), Association for Computing Machinery, New York, NY, USA, 50–57. DOI:https://doi.org/10.1145/3197768.3197776

[68] Chwan-Lu Tseng, Che-Shen Cheng, Yu-Hsien Hsu, and Bing-Hung Yang. 2018. An IoT-Based Home Automation System Using Wi-Fi Wireless Sensor Networks. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2430–2435. DOI:https://doi.org/10.1109/SMC.2018.00417

[69] Lionel Sujay Vailshery. Number of Internet of Things (IoT) connected devices worldwide from 2019 to 2021, with forecasts from 2022 to 2030. *Statista*. Retrieved from https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/#:$\sim$:text$=$Number%20of%20IoT%20connected%20devices,2021,%20with%20forecasts%20to%202030&text$=$The%20number%20of%20Internet%20of,billion%20IoT%20devices%20in%202030.

[70] Sai Vemprala, Rogerio Bonatti, Arthur Bucker, and Ashish Kapoor. 2023. *ChatGPT for Robotics: Design Principles and Model Abilities*. Microsoft. Retrieved from https://www.microsoft.com/en-us/research/publication/chatgpt-for-robotics-design-principles-and-model-abilities/

[71] Tianwei Xing, Luis Garcia, Federico Cerutti, Lance Kaplan, Alun Preece, and Mani Srivastava. 2021. DeepSQA: Understanding Sensor Data via Question Answering. In *Proceedings of the International Conference on Internet-of-Things Design and Implementation* (IoTDI '21), Association for Computing Machinery, New York, NY, USA, 106–118. DOI:https://doi.org/10.1145/3450268.3453529

[72] Lu Xu, Leslie Sanders, Kay Li, and James C L Chow. 2021. Chatbot for Health Care and Oncology Applications Using Artificial Intelligence and Machine Learning: Systematic Review. *JMIR Cancer* 7, 4 (November 2021), e27850. DOI:https://doi.org/10.2196/27850

[73] Zhipeng Xu, Hao Wu, Xu Chen, Yongmei Wang, and Zhenyu Yue. 2022. Building a Natural Language Query and Control Interface for IoT Platforms. *IEEE Access* 10, (2022), 68655–68668. DOI:https://doi.org/10.1109/ACCESS.2022.3186760

A APPENDICES

A.1 List of Selected Papers

Authors	Title	Year	Publication Journal or Conference
Ahmed et al.	Smart Home Shield and Automation System Using Facebook Messenger Chatbot	2020	2020 IEEE Region 10 Symposium (TENSYMP)
Anbarasan and Lee	Speech and Gestures for Smart-Home Control and Interaction for Older Adults	2018	HealthMedia'18: Proceedings of the 3rd International Workshop on Multimedia for Personal Health and Health Care
Baby et al.	Home automation using IoT and a chatbot using natural language processing	2017	2017 Innovations in Power and Advanced Computing Technologies (i-PACT)
Baby et al.	Home automation using web application and speech recognition	2017	2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS)
Barricelli et al.	A Multi-Modal Approach to Creating Routines for Smart Speakers	2022	AVI 2022: Proceedings of the 2022 International Conference on Advanced Visual Interfaces
Bhagath et al.	An Android based Mobile Spoken Dialog System for Telugu language to control Smart appliances	2021	2021 IEEE XXVIII International Conference on Electronics, Electrical Engineering and Computing (INTERCON)
Bhutada et al.	Ru-Urb IoT-AI powered Healthcare Kit	2021	2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)
Bui and Chong	An Integrated Approach to Human-Robot-Smart Environment Interaction Interface for Ambient Assisted Living	2018	2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)
Cabrera et al.	Intelligent assistant to control home power network	2016	2016 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC)
Campagna et al.	Almond the Architecture of an Open, Crowdsourced, Privacy Preserving, Programmable Virtual Assistant	2017	WWW '17: Proceedings of the 26th International Conference on World Wide Web
Chavis et al.	A Voice Assistant for IoT Cybersecurity	2021	2021 IEEE Integrated STEM Education Conference (ISEC)
Contreras et al.	Smart Home Multimodal Interaction for Control of Home Devices	2019	Interacción '19: Proceedings of the XX International Conference on Human Computer Interaction
Corno et al.	From Users' Intentions to IF-THEN Rules in the	2021	ACM Transactions on Information SystemsVolume 39Issue 4
Corno et al.	HeyTAP: Bridging the Gaps Between Users' Needs and Technology in IF-THEN Rules via Conversation	2020	AVI 2020: Proceedings of the International Conference on Advanced Visual Interfaces
Fakhrurroja et al.	Dialogue System based on Reinforcement Learning in Smart Home Application	2022	IC3INA '22: Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications
Ferrero et al.	Ubiquitous fridge with natural language interaction	2019	2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)
Gallo and Paternò	A Conversational Agent for Creating Flexible Daily Automation	2022	AVI 2022: Proceedings of the 2022 International Conference on Advanced Visual Interfaces
Gravril et al.	Multimodal Interface for Ambient Assisted Living	2017	2017 21st International Conference on Control Systems and Computer Science (CSCS)
Guaman et al.	Device Control System for a Smart Home using Voice Commands A Practical Case	2018	ICIME 2018: Proceedings of the 2018 10th International Conference on Information Management and Engineering

Gunawan et al.	Chatbot Application on Internet Of Things IoT to Support Smart Urban Agriculture	2019	2019 IEEE 5th International Conference on Wireless and Telematics (ICWT)
Huang et al.	Using Ontology Reasoning in Building a Simple and Effective Dialog System for a Smart Home System	2015	2015 IEEE International Conference on Systems, Man, and Cybernetics
Husain et al.	Increasing the Smart Home Automation by using Facebook Messenger Application	2021	2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS)
Iannizzotto et al.	A Vision and Speech Enabled Customizable Virtual Assistant for Smart Environments	2018	2018 11th International Conference on Human System Interaction (HSI)
Ilievski et al.	Interactive Voice Assisted Home Healthcare Systems	2019	BCI'19: Proceedings of the 9th Balkan Conference on Informatics
Jain et al.	Home Automation System using Internet of Things IOT	2019	2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)
Jivani et al.	A Voice Controlled Smart Home Solution With a Centralized Management Framework Implemented Using AI and NLP	2018	2018 International Conference on Current Trends towards Converging Technologies (ICCTCT)
Kaed et al.	A Semantic Based Multi Platform IoT Integration Approach from Sensors to Chatbots	2018	2018 Global Internet of Things Summit (GIoTS)
Kan et al.	LoRa-Based Air Quality Monitoring System Using ChatBot	2020	2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)
Kang et al.	Minuet Multimodal Interaction with an Internet of Things	2019	SUI '19: Symposium on Spatial User Interaction
Ketsmur et al.	Conversational Assistant for an Accessible Smart Home	2018	DSAI 2018: Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion
Kim and Ko	A Conversational Approach for Modifying Service Mashups in IoT Environments	2022	CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems
Kumar et al.	IoT Based Secured Home Automation System Using NLP	2021	2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)
Leong et al.	CASIS a Context-Aware Speech Interface System	2005	IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces
Liu et al.	Supporting the Onboarding of 3D Printers through Conversational Agent	2021	MuC '21: Proceedings of Mensch und Computer 2021
Mithil et al.	An Interactive Voice Controlled Humanoid Smart Home Prototype Using Concepts of Natural Language Processing and Machine Learning	2018	2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)
Mougy et al.	Xenia Secure and interoperable smart home system with user pattern recognition	2017	2017 International Conference on Internet of Things, Embedded Systems and Communications (IINTEC)
Muslih et al.	Developing Smart Workspace Based IOT with Artificial Intelligence Using Telegram Chatbot	2018	2018 International Conference on Computing, Engineering, and Design (ICCED)
Nasr et al.	Human Machine Interaction Platform for Home Care Support System	2020	2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
Nguyen et al.	A Miniature Smart Home Testbed for Research and Education	2017	2017 IEEE 7th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)
Ospan et al.	Context Aware Virtual Assistant with Case Based Conflict Resolution in Multi User Smart Home Environment	2018	2018 International Conference on Computing and Network Communications (CoCoNet)
Oumard et al.	Implementation and Evaluation of a Voice User Interface with Offline Speech Processing for People who are Blind or Visually Impaired	2022	PETRA '22: Proceedings of the 15th International Conference on PErvasive Technologies Related to Assistive Environments
Pattnaik et al.	A Secure and Interactive Home Automation System with Machine Learning Based Power Prediction	2021	2021 Innovations in Power and Advanced Computing Technologies (i-PACT)

Raj and Rai	Voice controlled cyber-physical system for smart home	2018	Workshops ICDCN ’18: Proceedings of the Workshop Program of the 19th International Conference on Distributed Computing and Networking
Rubio et al.	Seamless human-device interaction in the internet of things	2017	IEEE Transactions on Consumer Electronics (Volume: 63, Issue: 4, November 2017)
Salvi et al.	Jamura A Conversational Smart Home Assistant Built on Telegram and Google Dialogflow	2019	TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)
Sirinayake et al.	IOT-Based Intelligent Assistant Mirror For Smart Life amp Daily Routine Using Raspberry PI	2021	2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)
Stefanidi et al.	Programming Intelligent Environments in Natural Language An Extensible Interactive Approach	2018	PETRA ’18: Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference
Tseng et al.	An IoT-Based Home Automation System Using Wi-Fi Wireless Sensor Networks	2018	2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)