Talking to Everything

Conversational Interfaces and the Internet of Things in an office environment

MALIN AUGUSTSSON
Talking to Everything
Conversational Interfaces and the Internet of Things in an office environment

Malin Augustsson
KTH Royal Institute of Technology
Stockholm, Sweden
malinau@kth.se

ABSTRACT - English
This paper presents Cyberbot, a conversational user interface (CUI) designed to be used for information retrieval and interaction with smart Internet of Things (IoT) devices (smart objects), in an office environment. Built through the Research Through Design method, tested together with a selection of smart objects, the artifact was used to explore how CUIs can help office workers in their everyday working activities and the impact they can have on the usability in interaction with IoT, as opposed to more commonplace types of interfaces. Cyberbot was evaluated through user tests, surveys and semi-structured interviews. Results show that Cyberbot can be a welcome addition to an office and is useful for getting quick overviews of information from company intranet websites or smart objects that gather data from its environment, at the expense of more in-depth and detailed information. In general, Cyberbot can be said to be a usable and effective interface for IoT interaction.

ABSTRAKT - Svenska

AUTHOR KEYWORDS
Conversational user interfaces, chatbots, human-computer interaction, Internet of Things, HCI, IoT, smart office

INTRODUCTION
Over the recent years, the use of conversational user interfaces, here referred to as chatbots, have spread tremendously and are now found in many different parts of society. Some areas of application of these include customer service [1], banking [2], and even healthcare [3]. Chatbots are a simple and user-friendly way to interact with computers through conversation in natural language and they can help
users find information, complete or effectivize tasks as well as just being entertaining and fun.

There are two trends that have helped add fuel to the recent chatbot popularity. One is instant messaging (IM) apps, which are nowadays used by billions of people worldwide [4] and which work well as an easily accessible platform for chatbots. The other is the app model of executing activities, which many find problematic and difficult since it requires the user to download and learn a new app for each activity they wish to perform. Chatbots can provide support for the user without them having to navigate a new app or website or learning how to use a new interface [4].

Another thing that has grown immensely in use and areas of application over the recent years is the Internet of Things (IoT). This concept consists of a network of “smart” objects (things) interconnected and built upon existing internet architectures.

Though it is difficult to measure just how big the Internet of Things is at any given moment, in 2008 the number of interconnected smart objects outnumbered the human population for the first time and some sources predict the number to reach 20 billion by the year 2020, while some dare go even higher and predict the number to be closer to 50 billion [5], [6], [7].

Through various unique addressing schemes and standard communication protocols, these smart objects, or things, are connected to one another, which allows various technologies to work together in pursuit of common goals [8]. One of these goals is automation of everyday tasks. This has successfully been done in homes [9] retail [10], and even whole cities [11].

This, however, means that often the focus is on either the very largest scales (such as cities), or the individual (such as homes), and those are often different from organizational environments such as offices. And as more and more smart objects are introduced to every suitable space, interacting with each item individually through its own interface gets more troublesome. Part of this problem can be solved by having an app or similar to interact with all available objects, but smart objects can be very different from each other and are used for many different purposes, and many of them lack any sort of graphical interface. And as mentioned in [4], a conversational interface can replace the app to avoid both the need to learn new apps and the need to be specific (for example, asking a smart thermostat to heat the room to 22 degrees Celsius rather than “comfortable room temperature”) in one’s interaction with IoT objects. McTear et. al even propose that in the future, as more and more objects are becoming “smart” and connected, it is quite likely that conversational interfaces will end up being the only way to communicate with these [12].

And, seeing as the number of smart objects in the world is expected to grow as much as it is (as mentioned above), so will the number that we will be expected to communicate with in our daily lives. Having a single, unique conversational interface that works for every device might in this situation be preferable to each device having its own interface [12]. Therefore, the interaction with smart objects through chatbots is something that is likely to be very relevant in the near future and will therefore need to be researched.

Against this background, this thesis investigates the questions of how a chatbot can be used in an office environment as well as how and what types of IoT technology are suited to be interacted with through a chatbot in this context. Presented herein are explorations into the software design and development of a unique chatbot and its use in combination with different IoT smart objects in order to automate tasks and retrieve information in an office environment and its comparison to other types of interaction for the same purpose. Through the method of Research Through Design, a prototype is built and through the methods of user tests, surveys, and semi-structured interviews the prototype is tested for its usability along with the user experience of using it to interact with various smart object in an office environment.

In section Design & Implementation of this paper, the architecture of the bot prototype, named Cyberbot, and the details of its features are presented, and in section Evaluation & Results the results of evaluating the bot through user testing and the results thereof are presented. The results then summarized and discussed and conclusions as well as ideas for future work are discussed and proposed in section Conclusions & Further Research.

BACKGROUND

Conversational User Interfaces

Also known as “Chatbots” or “Dialogue systems”, conversational user interfaces (CUIs) let the user interact with the software through regular conversation, using natural language. This can be done through writing, such as a chat, or through speech.

Although “chatbot” seems like a recent buzzword, chat as a medium has actually been in existence for about as long as humans have interacted with
computers. The first-ever chatbot was called Eliza and was developed by Joseph Weizenbaum in 1966. Nowadays, there are chatbots in many places around us, on many important websites such as banks or governmental services, or even as mobile digital assistants in our smartphones and messaging platforms, like Apple’s Siri [17].

A chatbot consists of several modules and processes that are cooperating to reach the main goal of the interaction. Usually, these are divided into Natural Language Processing (see section below), Dialog Management, Content and External Data. A chatbot platform or framework will often provide Dialog Management and Natural Language Processing modules, and let the developer provide content and external data.

The Dialog Management (DM) part of a chatbot focuses on the important task of deciding and keeping track of dialog states and deciding what action or response the system should take in response to a user's input. The DM can be divided into four main parts: Updating the dialog context, providing a context for interpretation, coordinating the other modules as well as deciding what information to convey to the user and when to convey it [12].

The content is everything that the chatbot knows and shows to the user. In a retrieval-based bot, it can include pre-defined replies, or templates for building reply strings. External data is everything that the bot takes from external sources, such as data taken from databases or requested through APIs. Both content and external data vary greatly, depending on what kind of chatbot it is and what it is supposed to do [18].

In regards to best practices in chatbot interaction design, humans tend to react to technology which exhibits human-like characteristics and behavior according to Lankton et. al [19]. State-of-the-art chatbots often follow common design principles and recommendations. These include that messages from the bot should be short, concise, precise and informational. The bot should use the type of language that is deemed suitable for its context, and emojis, used in moderation, are generally a good idea.

Response cards; showing an array of clickable available choices, can be a good way of letting the user interact quickly, without having to type their reply, while also suggesting choices at the same time. However, if response cards should not be overused, as they can easily give the illusion of limited choices, and available choices should always be easy to understand. The user should always have the option to interact through typing if they prefer it. Response cards are generally considered a good choice for getting confirmation from users. Some chatbot platforms also allow gifs, images and videos to be shown to the user, which can be a good way to convey information as well as make for a more vivid, varied and interesting interaction experience [17].

In chatbot creation, trust is often an important factor. This is often defined as a combination of capability (the belief that the chatbot is capable to answer questions correctly and complete tasks properly), goodwill (the belief that the goal of the chatbot’s action is to provide care and benefit to the user) and ethics (the belief that the chatbot will follow set ethics guidelines relating to things such as privacy and data security) [20]. With trust properly established, the users are more likely to be willing to use the chatbot, find it usable and continue using it.

The abilities of a chatbot can place it into one of four groups, based on its knowledge domain (open or closed) and type of output generation (Retrieval-based or Generative), as is shown in Figure 1.

A retrieval-based model will base its responses and dialogue with the user on already pre-existing sample messages and pick the appropriate pre-existing reply for any given situation [21]. It can make use of various machine learning algorithms in order to better learn to understand the user’s intentions and choose appropriate replies. A generative model will instead create new responses befitting the existing context. A common way of doing this is using a sequence-to-sequence (seq2seq) framework based on recurrent neural networks. The framework divides the user input into tokens and reads them one at a time, using an encoder that creates a context vector. The model is trained to come as close as possible to the most correct output sequence and predict a suitable reply, one token at a time [22], [23].

Regarding domains, a closed-domain chatbot is able to converse with the user on subjects within a specific field, or several, but will not be able to understand or converse about anything outside of this. For instance, a closed-domain chatbot trained on the subject of dinosaurs will not be able to converse about country music. Open-domain chatbots have the ability to converse about anything. Out of necessity, these cannot be retrieval-based as it would be impossible to provide pre-existing samples for every possible conversation a user might ever have, as can be seen in Figure 1. Usually, retrieval-based models tend to be more linguistically coherent and grammatically accurate than generative models, given today’s level of technology [24].

Figure 1.
**Figure 1: Matrix of chatbot types based on knowledge domain (Conversations) and type of output generation (Responses) [25]**

**Natural Language Processing**

The field of interaction between computers and natural language as it is spoken and written by humans is known as natural language processing (NLP). This field combines computational linguistics and artificial intelligence [26], and research into NLP has the ambition to collect knowledge on how humans use natural language in order to develop tools and methods enabling machines to do the same. Understanding natural language poses three important challenges: the thought process, the representation of and linguistic meaning behind the input the machine accepts, and the understanding of one’s context, environment and surroundings.

Today, NLP is used in many different fields such as machine translation, text processing, user interfaces, speech recognition and cross language information retrieval [27].

In chatbots, the NLP module is responsible for analyzing user requests and utterances and turning them into structured data (this is called natural language understanding (NLU)).

Most commercial closed-domain chatbots today understand user input, called utterances by mapping them to pre-defined intents, where each intent represents something the user can ask the chatbot to do. The bot will then distinguish data required to fulfill the intent, called slots (sometimes also called entities or objects) from the utterance, or elicit them from the user through questions to help turn the user’s query into structured data that its back-end logic can work with. For example, a conversation might look like this:

User: "I would like to book a dinosaur ride tonight",
Chatbot: "Great choice of transportation. What kind of dinosaur would you like to ride?",
User: "A Triceratops, please"
Chatbot: “Those are very cute. I will book a ride for you tonight, then.”

In this scenario, the NLP module identifies the intent and slots and turns the user’s query into:

```json
{
  "intent": "bookDinosaurRide",
  "slots.typeOfDino": "triceratops",
  "slots.time": "16/04/2019 7:00pm"
}
```

The chatbot’s back-end logic can then call the appropriate functions to book the ride at the correct time and with the correct dinosaur, using the provided data, and then return a confirmation reply telling the user that the ride has been booked.

As the bot is developed the intents are defined, along with enough sample utterances for the bot to be able to map user utterances properly to intents. Sometimes, certain machine learning methods can be used in retrieval-based bots in order to train the bot to better map utterances to intents so that the bot makes less misunderstandings and mistakes and pick more correct replies.

Some chatbots utilize speech as an interaction modality, with or without also providing text-based chat. These utilize automatic speech recognition, which is the process of converting acoustic input to a text string, to recognize what the user has uttered. The next step is spoken language understanding; the process of obtaining a semantic interpretation of a text string, often involving morphological, lexical, semantic, syntactical, discourse and pragmatic knowledge [12].

**Internet of Things**

As previously mentioned, the Internet of Things consists of various smart objects (IoT devices), interconnected and able to communicate with each other over existing internet architectures.

ITU describes it as “a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.” [28] The RFID group describes it as “world- wide network of interconnected objects uniquely addressable, based on standard communication protocols” [28].
The components of an IoT system consists of sensors to gather data about the environment (such as light, sound, motion or temperature), actuators to create output (such as speakers, displays or motors), computation to process data and commands (such as programs or logic), and communication interfaces (which can be wired or wireless) [16]. Today it is common to see IoT as parts of home automation, with items such as smart electrical plugs, smart lightbulbs or smart door locks [9].

While the Internet of Things is a wide and varied concept, in this paper the focus is on smart objects that can directly interact with human users on an individual level through the use of chatbots. The setting for this study is an office environment and therefore the chatbot features and IoT objects are limited to those that are possible and appropriate to implement in such an environment, given the available information the company has and the predicted needs of the test participants, i.e. its employees. Therefore its scope is limited to those features and objects that are suitable in such an environment, as well as those that can be interacted with through Web Application Programming Interfaces (APIs).

RELATED WORKS

There has, over the recent years, been attempts to augment or automate certain tasks with the help of chatbots in various organizational environments, or to combine chatbots and IoT.

In a study published in 2018, Carayannopoulos presents a chatbot developed to be used in organizational environments, more specifically university students, focusing on new first-year students, to find information necessary to aid their transition from secondary school to higher education [4].

The chatbot could help students by finding information on the university’s website and intranet, providing miscellaneous tips and tricks and assisting them in the contact with their teachers or various other university personnel. The study found that this provided good results in this research context. This study, however, concerns a very specific age group which the author describes as “[A generation that] views and uses information and communication technology in vastly different ways compared to its parents.” This is not the case for a chatbot in an office environment, where the age as well as cultural and technological background among its users can sometimes vary greatly.

Another study of college students online help-seeking behavior found that students want to uphold a positive social image of themselves as competent, therefore they are more likely to seek help if there is a possibility to do so anonymously [13]. This can also be true for new employees at a company, who tend to also want to appear as competent as possible. According to some other studies, high quantities of information may be experienced as confusing, frustrating or have other unintended negative effects such as information anxiety and diminished performance [14]. A chatbot can alleviate these problems by streamlining the search process, making sure the user finds only the information they need and help them, through natural conversation, find out exactly what that is.

A Taiwanese study from 2018 explored the combination of conversational interaction and smart objects for health. In the creation of systems the researchers called “Smart Wireless Interactive Healthcare Systems” (SWITCHes). These were used to monitor the user’s diet, exercise and other health-related data such as reading from medical devices, logging caloric uptake and burn as well as, through a conversational interface, give the user advice and suggestions for a healthy diet, such as helping them pick out food and drink items and giving them nutritional facts about their chosen items. The advice and suggestions can be tailored to the particular user’s profile, based on their previous health statistics and weight loss goals [15]. The health chatbot can be used with different messaging platforms than its own and can also be interacted with using either speech or text means, much like Cyberbot can when used with an Amazon Alexa.

The SWITCHes in the study had, at the time of publication, not yet undergone a sufficient amount of user testing. However, the developed product has several similarities to Cyberbot. The SWITCHes can, when used correctly, provide information about certain aspects of the user’s health through a conversational interface, and be used to plan long-term health improvements for the user. It is however not focused on the same aspects of health as Cyberbot, and it can get its information both from the user, from a smart health device of some sort, carried by the user, as well as other sources. Lastly, it uses an app, and the chatbot was only a part of that, rather than its focus, which is not the case for Cyberbot.

In another study, published in 2016 [16], the researchers explored the possibility of using chatbots to provide better interaction with smart objects. The study proposes that modern IoT systems face two types of challenges; technology-centric and human-centric. The latter often involves the failure of users to properly communicate their intentions in a way that is usable to the devices. Humans think in, and use natural language, so while a user might want to, for example, change the temperature in the room using a smart thermostat, they might be looking for what
they would call a comfortable temperature. But the thermostat has no setting called “comfortable”, as that is very much a relative state. Here a chatbot can help by discussing with the user what “comfortable” might mean or base it on contextual data (like the outdoors temperature and its user’s historically preferred temperatures).

Chatbots can also make feature recommendations to the user regarding different actions, or automate common, cyclic tasks such as checking sensors. The study found that the use of chatbots allowed for more precise control over the smart objects. This was however done in a home, which is a different domain than that of Cyberbot, and a number of the devices and interactions in the study were more appropriate in homes but would not have been as appropriate in offices (for instance, thermostats are fine when only one person or a few are using it, but an entire office staff are likely to have very different preferences).

Like Cyberbot, the scope was limited to those smart objects that can be connected to the chatbot through APIs and HTTP-based REST architectures and the study found that this was enough to allow for a lot of varied devices fulfilling varied user needs [16].

METHOD

The chosen method for this project was research via the design and implementation of a prototype, which was then evaluated through user testing.

In Researching Information Systems, Briony Oates calls this method “Design and Creation” in Researching Information Systems [29] and describes this as a method wherein a new IT product, or artifact, is created. These artifacts can be, among other things: constructs, models, methods or instantiations. An artifact is often a combination of two or more of these. Cyberbot, the artifact in this study, is a combination of models and an instantiation; a collection of use case scenarios and a prototype of a functional computer-based system based on these [29].

In the human-computer interaction field, this method is more commonly known as Research Through Design, which is described by Zimmerman et al. [30] as involving “grounding—investigation to gain multiple perspectives on a problem; ideation—generation of many possible different solutions; iteration—cyclical process of refining concept with increasing fidelity; and reflection.” and in more detail as “through an active process of ideating, iterating, and critiquing potential solutions, design researchers continually reframe the problem as they attempt to make the right thing. The final output of this activity is a concrete problem framing and articulation of the preferred state, and a series of artifacts—models, prototypes, products, and documentation of the design process” [30]. Following this method, Cyberbot was grounded through user interviews and research into user needs and requirements as well as best practices for conversational interaction, then ideated through its implementation, whose solutions were then iterated and tested through user trials to find an acceptable state of the prototype and reflection was done through the documentation of the knowledge gathered during the process.

The design and prototyping of Cyberbot follows the principles behind agile development. These methods favor development through iterative processes, rather than a more plan-driven process such as the waterfall method [31]. Work was done in sprints, each lasting two weeks and focusing on a specific part of the development of Cyberbot, such as user research, design and implementation of specific features, user testing or the gathering of necessary information for continued work.

DESIGN & IMPLEMENTATION

The first step in the development of Cyberbot was the identification of the necessary user requirements. Potential users from the target audience were interviewed. These were all employees of the company, working in different roles such as IT consultant, manager and HR specialist. The interviews were held separately, in a meeting room, and consisted of questions about what tasks related to their work they perform on a regular basis using the company’s websites, as well as what information they repeatedly need to look for or provide to others and what tasks they repeatedly need to perform. Also investigated was their familiarity and level of comfort with the websites and the technology they were expected to use. The requirements were drafted based on these, and a pattern emerged showing that the participants often found the intranet website to be messy and difficult to navigate and retrieve the correct information from, especially since there are several different intranet websites such as payroll, database of knowledge and frequently asked questions, and sometimes it was difficult to know which one to use. The participants hoped that a chatbot could help them more easily find the correct information when they needed it and help automate and effectivize tasks in their everyday working life.

The Architecture of Cyberbot

The Cyberbot prototype was created using a chatbot framework. A framework usually provides a basic system for dialog management and natural language processing, and sometimes a sample of intent, utterances and synonyms. The developer then adds
the content; required intents, utterances and backend logic for auxiliary operations (such as communication with smart objects). Several different frameworks were regarded and eventually Amazon Lex was chosen due to its flexibility, ease of integration with other services and devices, such as Slack and Messenger, as well as the wide range of compatible services offered by the same company (such as Amazon Alexa, CloudWatch, which was used during development, and Sumerian).

Cyberbot consists of different components handling different parts of the user interaction and information flow, as can be seen in Figure 2.

- Visual interface. There are several variations on this due to the bot being deployed in different channels. There are traditional chatbot channels such as for Slack and Messenger, and a web-based graphical interface developed for the bot in React. Voice chat through Amazon Alexa and a 3D model character in Amazon Sumerian were also used.

- Dialog Management system. This involves keeping track of dialog states, eliciting the slots (prompting the user for input) when necessary and handling the errors that arise when the bot does not understand the user's input.

- Natural language processing (see sections Conversational user interfaces and Natural Language Processing in Background). This handles the understanding of the user's utterances, connecting them to intents and slots, extracting of the meaningful information. The Amazon Lex bot framework uses pattern-based recognition to process user input by comparing it to sample input stored in Amazon’s databases.

- Back-end logic and computation. In Lex, this is handled through Lambda functions. A Lambda function is connected to one or more intents and when that intent is invoked, the code in it connected to fulfilling the intent is executed. This can involve querying databases, making calculations based on user input or making requests to various APIs to retrieve or deposit data. The Lambda functions handled the creation of suitable text output for presenting the retrieved information or the confirmation of a completed task.

- Physical and IoT devices. This involves all the devices interacted with through the chatbot. For the prototype, this means the health tags, the smart plug and (in the future) the occupation state detectors.

The prototype was given 56 intents, some with connected Lambda functions and some without. The intents ranged from simple ones which could recognize, for example, a greeting, and pick a random greeting as response from an array or response, to more complex ones which, once invoked, would start longer conversations with users, eliciting several slots and connecting to Lambda functions requesting external data from (or posting user input data to) APIs in order to fulfill the user’s requests. Each intent had 3-5 sample utterances to make it easy for the bot to map similar user utterances to it. The personality of Cyberbot was modeled after the typical, friendly interaction one might have with a colleague. Cyberbot appears warm and empathetic but still very professional, with the occasional fun quip or joke.

The developed Cyberbot prototype was given a number of features. These can be divided into four main feature groups:

**Information Retrieval**

Some of Cyberbot's features relate to the task of retrieving various information for the user. These features include answering questions frequently posed to managers and human resources (HR) staff at the company, as well as helping the user search for information on the internal company website. The difficulties of information searching on the website was brought up in the pre-development user interviews by several employees and Cyberbot provides an alternative to asking humans or searching on your own.

**Communication**

Communication features included retrieving the user’s calendar events, finding optimal meeting times based on attendees’ schedules and available
rooms in the office as well as creating new events and meetings in the calendar. These features were implemented using Microsoft Graph’s Outlook Calendar API. There were also features relating to sending emails with predefined texts to certain colleagues or managers on the user’s behalf.

**IoT features**

The IoT features to be integrated were chosen according to the plan to integrate the chatbot with one common and widely used kind of IoT object, one new and more unusual that is not commonly seen in offices but solves problems that often affect office workers, as well as at least one that was created especially for offices.

The common and widely used IoT object chosen was a smart electrical plug. This allows users to start and shut off electrical appliances remotely by restricting electricity flow. Objects meant to automate everyday tasks for increased comfort of users, such as turning on and off electrical appliances, changing settings (temperature, humidity, channel for instance) are common, although usually meant to be used in homes. The smart plug communicates with the chatbot through Wi-Fi.

The new and more unusual IoT object chosen for the study was a smart health tag. Stress is a common problem for a lot of people today [32], and a big source of stress is often work, both one’s workload and the work environment itself. Many companies today, including the company in this study, attempt to take measures to detect and diminish stress among the employees in their offices, and many new technologies have been developed to help with this [33]. As such, a health tag that measures biometrics related to stress (such as respiration, heart rate or even sweat) can help the individual employee track their stress levels and therefore take steps to reduce them. Cyberbot fetches the data the tag requests from the provider’s database through an API and provides the user with a quick overview of their stress levels, as well as some general health and stress management advice.

The IoT objects created especially for offices were a digital timetable showing the arrivals and departures of commuter- and metro trains at the nearby station, as well as an occupation state detector for office spaces and utilities. The digital timetable collected its traffic data from an API of local public transport. The occupation state detector comes in two parts: one that uses movement detectors in meeting rooms to estimate whether or not the room is currently occupied, and one that shows the occupation state of a ping-pong table, commonly used for recreation by the employees.

The detector for the ping-pong table used scales to weigh the box that held the bats, calculating the amount of bats based on predefined data about the weight of each bat and the box itself. If the box was deemed to hold an amount of bats that was less than a full box, then that amount are number of bats that are currently in use, i.e. the number of occupants of the ping-pong table.

The calculation algorithm recalculated its base weight (the weight of a full box of bats) regularly, meaning that if a new bat was added, a bat was not put back in the box after a longer period, the full-box amount of bats would be changed to the current amount, to avoid showing the table as occupied when it was not.

Both of the occupation state detectors were based on Particle Photon [34], a wi-fi development kit allowing the detectors to connect to the cloud. The live weight and motion data could then be accessed through a REST API and its occupation state calculated, as can be seen in Figure 3.

![Figure 3: Occupation state detectors information flow](image)

**Esoteric features**

Esoteric features were the ones not related to fulfilling any specific practical purpose related to work or IoT, but rather of being entertaining, interesting or relieving boredom. These features included asking the chatbot for public holidays, weather, public transport information (the same as was provided by the previously mentioned digital timetable) or nearby lunch places, as well as telling jokes or facts about interesting topics.

**EVALUATION & RESULTS**

The evaluation of Cyberbot’s usability and the assessment of its value in terms of greater employee efficiency was mainly done through user testing,
surveys, interviews and qualitative as well as quantitative data analysis. The participants were all office workers from two different offices belonging to the same company and were selected based either on availability or previously expressed interest in the Cyberbot project and its IoT features. They all had academic backgrounds, and most of them were between 25 and 60 years of age, with an even gender division.

Information retrieval features – Evaluation

The user test of Cyberbot’s information retrieval and communication features as well as general usability included 10 participants, both men and women, aged 31 to 61, who were given a series of different scenarios that were all common occurrences for people at the office, such as “You as a consultant have just found out that you will have kids and are not sure about how maternity and paternity leave works at this company. See if Cyberbot can help.”, “You have never had to print anything at the office before. Now, you are in a rush and need to print a few documents and are wondering how to do it. Instead of bothering your co-workers you are looking for that information on the intranet but find nothing. See if Cyberbot can assist you.” and “You are sitting at the office and are feeling bored. It is soon time for lunch. You usually bring some food to work, but today you forgot. Ask Cyberbot about suggestions for lunch. And since you are heading out you are wondering about the temperature outside. You can also try to ask Cyberbot about something funny - maybe what the meaning of life is, or something else, like a joke.”

The users were asked to imagine themselves in these scenarios and interact with Cyberbot as they would in those situations. The tasks they used Cyberbot to complete were related to finding information from the company’s websites or the internet, as well as relieving boredom through the esoteric features such as jokes or small talk. After the test, they were subjected to semi-structured interviews with questions about how they normally do these tasks, how user-friendly they felt the bot was, how much they trusted it and whether they would use the bot for similar tasks again, given the chance.

Information retrieval features – Results

The study yielded mixed results. The majority of participants, 8 out of 10, were positive or very positive to the bot, claiming they would gladly use it or something similar for these purposes again, and that it was equal or better than chatting with a human or looking up information yourself. The rest were either ambivalent, or sometimes negatively inclined. For the most part, they appreciated it for its ease of use and as a way to get some quick overview answers, at the cost of more in-depth and detailed information.

As an answer to a question about what they preferred out of chatting with a bot, chatting with a human and looking up the information themselves, one participant said “[...] as a matter of fact, I think it’s good to check in with a chatbot first. I hate looking for information on normal portals, as that returns like 10 documents. But if I’d had a really great relationship with someone who works with exactly this, maybe I would’ve asked them first. But I think [the chatbot] is great. Or I don’t know, perhaps I would have chatted first because it’s so easy, then asked, and at last looked.”

It can also be noted that several participants complained about how intranet websites at companies are always difficult or impossible to search. They really appreciated Cyberbot for that specific purpose, and that the younger participants generally felt much more comfortable with, and were often more positive to, the chatbot than the older participants were.

IoT features

The parts of the evaluation regarding its use in combination with already implemented smart objects, were done in part through user testing and interviews and in part through a survey. Surveys can consist of questions both for qualitative and quantitative data, such as multiple-choice questions and text fields, and can therefore be a useful method of gathering user data when both qualitative and quantitative analysis of data is desired [29].

Occupation state detectors and digital timetable - Evaluation

Due to time and resource constraints, and the occupation state detectors being at a different office than the other tested smart objects, the integration of the occupation state detectors with Cyberbot was not finished at such a level that it could be directly tested. Therefore the evaluation consisted of a survey with questions asking the employees of this office about their experience with the digital timetable and occupation state detectors, which feature they used the most, how often they had been used during the past year and whether they would interact with it through a chatbot, given the chance, as opposed to the map-based graphical interface it currently has on the company’s intranet website, where a map of the office is displayed and occupied rooms and utilities have a red colour while unoccupied are green. They were also asked for additional input and ideas regarding the integration of Cyberbot and the occupation state detectors, the answers being both
quantitatively and qualitatively analyzed. The survey collected no demographic information, as this would have made it longer and therefore decreased participants willingness to reply. The employees in the office had a general demographic of 30–45 years of age and around two thirds were male. The survey was answered by 21 employees.

**Occupation state detectors and digital timetable – Results**

The majority of the users, 71%, who answered the survey claimed to have used the meeting room occupation state detector or the digital timetable the most, with 29% of users claimed to have used the ping-pong table occupation state detector the most. Regarding frequency of use, more than half (57%, rounded numbers) had used the IoT objects 1-5 times, while 29% had used them more than 10 times or around every week. 14% had never used them at all. When asked whether they would interact with them through a chatbot instead of the current map-based view, if available, almost half (43%) of users replied “Yes”, with the rest reported being unsure or needing more information to make a proper choice. 0% of users chose “No” as an answer. One user reported that while it was a nice idea, they would prefer it if a chatbot focused on helping office employees with more information retrieval-related tasks, and mentioned some of the tasks that Cyberbot, in its current state, was already capable of. This is interesting, seeing as the users at this office had never used or heard of Cyberbot previously. The frequency of use was unexpectedly low. Almost three quarters of users had only used the objects 1-5 times or not at all during the past year. However, the rest had used them many times and the person who had been responsible for keeping them up and running noted that he had often been approached during downtimes by coworkers who wanted the objects back up since they were eager to use them. This would suggest that while many people were not interested in using the objects, the ones who were, were very interested and used them a lot.

**Smart plug – Evaluation**

The evaluation of the smart plug was done by connecting to a television in an often-used meeting room. It was left there for the duration of two weeks, (10 working days) along with a message about what how it should be used to turn on and shut off the television using Cyberbot, and a notepad for feedback and a note informing users that feedback could also be sent via email. For this purpose, the chatbot was connected to an Amazon Alexa device, giving it a voice interface as well as a text-based one for added time efficiency for the users.

After the 10 days the feedback from the emails and notebook was collected and a small group of two users, one male and one female, aged around 30, who had previously showed interest in trying Cyberbot for this purpose were asked for additional written feedback about their experiences (positive and negative) of using it and the smart plug for this, as opposed to starting and shutting off the television manually.

**Smart plug – Results**

The results showed that while users generally found the project interesting and appreciated the possibility of interacting with the meeting room’s television this way, they were often unused to and intimidated by the new, unusual technology, and this led to some unwillingness to use it, especially in the context of a meeting in the office where they were with a group of people and everything needed to proceed smoothly and in time as to not create a bad experience for everyone involved in the meeting. It was noted that the older users felt this more than the younger, who were in general much more enthusiastic about using it. One of the younger users said “We often have meetings in that room and always use the TV. The part that is ALWAYS difficult is finding the remote, and then pressing the correct button. I have made it a habit of asking it to turn on the TV before I even sit down [...] it saves me on average 5 seconds and a lot of annoyance per meeting.”

**Health tag - Evaluation**

The health tag came with an app, as is common for this type of device, which can show various biometric facts about the user. Cyberbot instead read the data from the tag manufacturer’s database, compared it to historical data about a specific user and returned a reply about the user’s general stress levels, along with some short advice for stress reduction, such as “Nothing out of the ordinary, your stress levels seem normal. But slowing down and having a break can be a good idea if you are starting to feel stressed.”

The user could also ask Cyberbot for general health and stress advice, a function similar to one in the app. Cyberbot had one intent that would return general stress level info as mentioned above, and one that would return a piece of stress management advice, randomly chosen from a set of predefined replies, such as “If you feel very stressed right now, try taking a quick break and meditating a little or reading a book.”

The health tag was evaluated through a user test where the new artifact to be tested (Cyberbot) was used with the tag and compared to the default interaction method, the tag manufacturer’s app. The test lasted for 8 working days and involved three participants wearing the health tag during work and using the app to get updates about one’s stress levels.
and advice as well as querying Cyberbot for stress level updates and advice.

No test participant had any prior experience with health tags, health-related IoT or the tag manufacturer’s app, and had only a normal level of casual experience with using chatbots. After the test period, semi-structured user interviews were held, with qualitative, open-ended questions about the general feel of the usability of app vs. chatbot interaction with the tag, how useful the tag was in a work context, how they felt about using a chatbot for this information and whether it was something that they would prefer over interaction through an app.

**Health tag – Results**

The interview results were qualitatively analyzed and showed that participants were a bit confused by the concept of the health tag itself since they had not used anything similar, but found it easy to use Cyberbot for the most part, even though it was sometimes difficult to know how to best communicate with it to reach the correct intents and get the desired results. In general they found that the tag made them more aware of their own stress and health patterns, and they enjoyed Cyberbot’s personality, and found the interaction more personal and fun than the app, as well as simpler and easier than starting the app and digging around in it looking for what they needed to know when they just wanted a quick overview, which they wanted most of the time. However, this came at the expense of more detailed and varied information and, which the app did better.

“Most of the time you just want to quickly check your [stress] levels and that’s easier done with the bot,” said a participant. “But sometimes you want to see more in-depth information and details and visuals and things, and the app offers much more of that.”

**Summary of results**

In summary, the testing of the bot itself showed that users were generally positive about using it both for information retrieval and for IoT interaction. A pattern emerged showing that they often found the chatbot to be useful for getting small bits of information, or see a quick overview, because that was often what they were primarily looking for. “Effective” was a word that cropped up often during comments and interviews in regard to the difference in interaction. However, sometimes they wanted more detailed and in-depth information, which then made them prefer the default modes of interaction instead.

Cyberbot was also perceived as more interesting and fun to interact with than the defaults, but some users noticed when it didn’t seem to understand the utterances or sent very flat and impersonal replies, which made them slightly averse to it.

Regarding the digital timetable and occupation state detectors, some users were skeptical or averse to them, some were really enthusiastic, with the majority being mostly ambiguous. The users were in general quite positive to using them with a chatbot, believing that it was likely to be a good way to interact. Another pattern that was visible throughout the evaluations was that some users saw both the bot and the smart objects as new technology which they were unused to. For some, this increased their willingness to try them, but for others (especially older participants), it instead decreased it.

**DISCUSSION**

A few concerns arose during implementation and evaluation. First of all, how valid are the results? Regarding the internal validity, that is, the validity of the results within the researched context, the test participants were a small group, which can call the internal validity into question. However, the test and interview data generally followed the same, clear patterns. The participants generally had similar feelings and opinions about Cyberbot and the devices part of its IoT features, which makes the data unambiguous and the results clear. This makes it reasonable to draw conclusions from the data that these same patterns would also be present throughout the entire population in this particular research context, therefore the internal validity of the results could be said to be fulfilled.

Regarding external validity; the possibility to generalize results from this study to other contexts and situations, the answer is slightly less clear. The study focuses on technology in office environments, but the company in the study is part of the technology sector and many of the participants in the study are either engineers or people who otherwise work closely with new technology and can therefore be expected to be more familiar with IoT and chatbots (a few participants even compared Cyberbot to other chatbots they had used) as well as being introduced to new technology they have not previously seen or used than the average person. This is, however, not true for everyone who works in an office environment. However, digitalization and introduction of new technologies has been happening across many workplaces over the past few decades and in general, office workers across many sectors and industries have been able to adapt to the changes [35]. It is realistic to assume that they will continue to do so.

Furthermore, none of the IoT objects in this study are specific to the needs of people in the technology
sector, neither are chatbots, and nothing in the routines, methods and regular workflow of the company in the study or its employees that participated in the study set them apart from the normal companies and office workers. Taking all this into account, the results can be said to hold external validity.

There are a few potential societal and environmental implications that can arise, regarding technology like this. The health tags log data about health, which can, depending on what type of health data is collected, be potentially personal, sensitive data. This can be a privacy issue, both in the collection of the data itself, and in the eventuality that a company would decide to keep the logs and use them to gain insight into the health and habits of individual employees and use it as basis for when deciding who to hire and whom to let go. On the other hand, keeping the logs could also let companies use statistical methods to find patterns in the employee population, predict when stress is going to occur and therefore figure out what is causing it and prevent it. Similar attempts to predict and prevent stress in office workers have been attempted before [33]. Chatbots themselves can often log data about the conversations it has with users, and the Amazon Lex framework actively does this, as stated in their FAQ and privacy policy. This can pose a privacy issue in case the conversations relate to private matters. However, the design of Cyberbot took this into account, and the bot, in its current state, has no intents that require the user to disclose private information, so this should not be an issue in this case.

Another potential societal issue is that some of Cyberbot’s features are directly imitating tasks that are normally done by a human (such as asking questions to HR or management). Yet another potential issue is automation of jobs, that robots, AIs, chatbots and the like will take over jobs currently done by humans, have been on the rise for the past few years and is something many people are worried about [36]. In the study in this thesis, however, some test participants from the HR department were very positive to Cyberbot doing these tasks, believing it to help them free up time to do more complex tasks and be more productive. Considering this, it is highly unlikely that Cyberbot in its current state will steal anyone’s job.

Regarding environmental implications, Cyberbot could be further developed to bring some positive change, seeing as IoT objects can be smart enough to not use more energy than necessary, and chatbots can give advice and suggestions for energy-saving and environmentally friendly behavior. Some attempts at using IoT for sustainability in office settings have been done before, such as the #SmartMe energy system [37].

**TECHNICAL CHALLENGES**

Some technical challenges Cyberbot and future similar office chatbots might pose include how the bots and the IoT objects should be acquired. A recent Gartner report outlines this as “buy, outsource or build” [38]. Building your own platform, and your own smart objects, lets companies create exactly what they need, suited for their particular situation. The current availability of products like the Particle Photon, which the occupation state detectors were based on, makes it easier and cheaper than before to build your own smart objects. However, it requires in-house skills, and often a lot of time. The occupation state detectors and digital timetable in this study were mostly built directly at the company, requiring in-house skills that not all companies have, and time they might also lack. Buying a platform and developing on top of that is often enough for the most situations, especially as a chatbot with machine learning takes a lot of time to build and train. The platform for Cyberbot itself, Amazon Lex, was bought, and Cyberbot developed on top of it. Outsourcing, letting another company build the bot or the smart objects for you, means less control over the process and result. The smart plug and health tags in the study were built fully by other companies and added to Cyberbot as-is. It would likely have been a better choice to build or at least add to the existing smart objects rather than buying, as they were not originally meant for office environments.

Authentication posed a problem during implementation. Some of the information retrieval features required access to the user’s calendar or contacts. The implementation of Cyberbot used Bearer-type tokens from Azure AD [39]. These are strings requested through an API call with the necessary credentials of the application, the user then logs in, and a token is received. This has a set expiration time and is attached to any API calls that need user authentication. Since many existing systems of authentication assumes that the application either exists as a desktop or web application, and a chatbot often exists as only a small part of a web application, this can cause problems in case the chatbot is developed separately from the web application and added to it at a later stage. Authentication for smart objects can also pose problems since devices are often limited in terms of input interaction possibilities. Azure AD uses something for this “Device Code Grant Flow”, where the user is provided with a code and expected to use it to log in and authenticate in a separate web browser.

This, however, requires more work for the user than simply opening a new browser window the way they
would do in a regular web-based application, and is therefore not optimal for authenticating users through chatbots. New methods for authentication when using chatbots need to be found or developed for a comfortable user experience.

The company in the study used Outlook for all their e-mail and calendar needs, and Microsoft provides a generally easily accessible API for that through Microsoft Graph, with possibilities for requesting lots of different information that might be very useful for a chatbot, such as schedule availability for meeting attendees. This might, however, not be the case for every email provider. A feature for e-mail and calendar retrieval might therefore not always be possible, which can make these features difficult to implement for chatbots for some companies.

CONCLUSION & FURTHER RESEARCH

The study presented in this thesis contributes to an understanding of chatbots in office environments as well as chatbots as a user interface for interaction with smart IoT objects. It shows that users in this context, that is, office workers, find chatbots a more effective way of both retrieving information relevant to their work tasks and interacting with. People in offices (and similar organizational environments) often have similar information needs and similar questions that they need answering, and answering frequent questions is an area that chatbots excel in and having them do that frees up the office workers who normally do it to spend their time on other tasks. Smart objects, are, while not usually made for offices at present, something that can help office workers be more effective, save time, and even be more aware of and diminish their stress levels, which is something which could be very beneficial to both individual employees as well as to the companies. The problems that often arise in the case of using and getting used to several new kinds of IoT, such as app fatigue (growing tired of using apps and having to learn new apps), confusion or an overload of information, can be solved in part with the introduction of a chatbot for smart object interaction, if the goal is to provide the user with easy quick-overview access to data collected by smart objects, which the study showed is often what users want.

Further research on the subject could involve more different types of smart objects than the small sample presented in the study. There are a multitude that could be researched, such as smart thermostats or smart light bulbs (not tested here because they would require making changes to the interior of the office, something which was not possible in the office in this study), or occupation state detectors for other purposes than the ones presented here. There are also other types of IoT for health that could be tested, such as Eccrine Sweatronics, which measures sweat. Testing the combination of chatbots and IoT in other organizational environments and researching the differences between those and the office, also seems like a logical step forward. Since the study uncovered that there was a slight difference in older and younger office workers’ willingness to try IoT and chatbots, it would be interesting to do research on what other factors also affect this. It would also be interesting to do a study with a generative chatbot instead of a retrieval-based, and see how that would affect the interaction and user experience; would the conversation be perceived as more natural or human-like? Or would the generated replies mostly be confusing or annoying?

An area that, given the results, seems likely to be the next step in IoT for offices is Ambient Intelligence. McTear et. al describes this as “The way devices in IoT work together to assist people in their activities of daily living within smart environments” [12]. This could, for example, mean that smart lights, thermostats or air conditioning devices in a room utilizing motion sensors like the one in the meeting room occupation state detectors could modify the lighting or temperature automatically according to how many people are currently in the room. This could both make for added comfort for people as well as a more environmentally friendly, energy-conservative office. McTear et al also notes, a chatbot will most likely be the best way of communicating with these smart objects [12].

The future sure seems bright when conversational interfaces and IoT are in it, especially if they work together for the benefit of office workers everywhere.

ACKNOWLEDGEMENTS

I’d like to thank the Sveavägen and Kista offices of Cybercom, which was the company in the study, my school; the KTH department of Media Technology and Interaction Design at EECS, Magnus Brodén and Gabriel Skantze who were my supervisors, and Sofia Sundström and Irfan Khalid who always supported me when needed, as well as all test participants and survey respondents, and all the Cybercom S20 spring 2019 thesis students, especially my fellow members of Team Cyberbot.

REFERENCES

Lecture Notes in Computer Science, vol 11193. Springer, Cham


[16] Kar, R., Haldar, R. (2016), "Applying Chatbots to the Internet of Things:


leverage-nlp-machine-learning-for-you-chatbot-5


[27] Chowdhury, G. G. (2005), "Natural Language Processing", in Annual Review of Information Science and Technology


[34] Particle (2019), Particle Photon, accessed May 2019 from https://docs.particle.io/photon


[38] Gartner (2019), Five Questions for a Successful AI Project


[40] Eccrine Systems Inc. accessed May 2019 from https://www.eccrinesystems.com/solution