



Conversational Agents for Insurance Companies: From Theory to Practice

Falko Koetter¹, Matthias Blohm¹(✉), Jens Drawehn¹, Monika Kochanowski¹,
Joscha Goetzer², Daniel Graziotin², and Stefan Wagner²

¹ Fraunhofer Institute for Industrial Engineering,
Nobelstr. 12, 70569 Stuttgart, Germany

{falko.koetter,matthias.blohm,jens.drawehn,
monika.kochanowski}@iao.fraunhofer.de

² University of Stuttgart, Universitätsstr. 38, 70569 Stuttgart, Germany
joscha.goetzer@gmail.com, {daniel.graziotin,
Stefan.Wagner}@iste.uni-stuttgart.de

Abstract. Advances in artificial intelligence have renewed interest in conversational agents. Additionally to software developers, today all kinds of employees show interest in new technologies and their possible applications for customers. German insurance companies generally are interested in improving their customer service and digitizing their business processes. In this work we investigate the potential use of conversational agents in insurance companies theoretically by determining which classes of agents exist which are of interest to insurance companies, finding relevant use cases and requirements. We add two practical parts: First we develop a showcase prototype for an exemplary insurance scenario in claim management. Additionally in a second step, we create a prototype focusing on customer service in a chatbot hackathon, fostering innovation in interdisciplinary teams. In this work, we describe the results of both prototypes in detail. We evaluate both chatbots defining criteria for both settings in detail and compare the results and draw conclusions for the maturity of chatbot technology for practical use, describing the opportunities and challenges companies, especially small and medium enterprises, face.

Keywords: Conversational agents · Intelligent user interfaces · Hackathon · NLP chatbot · Insurance

1 Introduction

With the digital transformation changing usage patterns and consumer expectations, many industries need to adapt to new realities. The insurance sector is next in line to grapple with the risks and opportunities of emerging technologies, in particular *Artificial Intelligence* [31]. Additionally, innovation methods like design thinking and open innovation are on the rise. In unsecure market times innovation is crucial, and all organizations and also traditional companies need to keep up to date by using new technologies for innovative business processes [27].

Fraunhofer IAO as an applied research institution supports digital transformation processes in an ongoing project with multiple insurance companies [35]. The goal of this project is to scout new technologies, investigate them, rate their relevance and evaluate them (e.g. in a model trial or by implementing a prototype). While insurance has traditionally been an industry with very low customer engagement, insurers now face a young generation of consumers with changing attitudes regarding insurance products and services [32]. Another goal of the project is the establishment of innovation methods within the companies and enable them to develop new products and services themselves.

Traditionally, customer engagement uses channels like mail, telephone and local agents. In 2016, chatbots emerged as a new trend [16], making it a topic of interest for Fraunhofer IAO and insurance companies. With the rise of the smartphone, many insurers started offering apps, but success was limited [33], which may stem from app fatigue [38]. App use has plateaued, as users have too many apps and are reluctant to add more [13]. In contrast, conversational agents require no separate installation, as they are accessible via messaging apps, which are likely to be already installed on a user's smartphone. Conversational agents are an alternative to improve customer support and digitize processes like claim handling or managing customer data.

The objective of this work is to describe the creation of conversational agents in theory and practice and show the outcomes of both views. We facilitate the creation of conversational agents by defining the traits of an agent more clearly using a (1) classification framework, which is based on current literature and research topics, and systematically analyzing (2) use cases and requirements in an industry, shown in the example insurance scenario. We frame two application scenarios with this theoretical foundation. Prototype 1 is a claim-handling scenario, which shows technological progress for a conversational agent. In this extended version of our former paper [23], we present prototype 2. This new prototype has been created for the scenario of customer service and cross selling. It is created in the setting of a chatbot hackathon event that Fraunhofer IAO organized in 2018. The goal is to gain more insights about conversational agent creation while examining the practicability of chatbot implementation for small insurance scenarios. Furthermore, we enriched the evaluation chapter of both prototypes and compare the results of both activities. We derive possible applications, knowledge about challenges and success factors as learnings from both activities. We apply this knowledge in a new project for supporting small and medium enterprises in adoption of new technologies.

2 Related Work

In this section we investigate work in the area of conversational agents, dialog management, and research applications in insurance. In extension to the previous paper [23], we add theory on hackathons at the end of the section.

[26] offer detailed explanations about background and history of conversational interfaces as well as techniques to build and evaluate own agent applications. Another literature review about chatbots was provided by [4], where common approaches and

design choices are summarized followed by a case study about the functioning of IBM's chatbot Watson, which became famous for winning the popular quiz game *Jeopardy!* against humans.

Many chatbot applications have already been built nowadays with the goal to solve actual problems. One example is PriBot, a conversational agent, which can be asked questions about an application's privacy policy, because users tended to skip reading the often long and difficult to understand privacy notices. Also, the chatbot accepts queries of the user which aim to change his privacy settings or app permissions [17].

In the past there have already been several studies with the goal to evaluate how a conversational agent should behave for being considered as human-like as possible. In one of them, conducted by [22], fourteen participants were asked to talk to an existing chatbot and to collect key points of convincing and unconvincing characteristics. It turned out that the bot's ability to hold a theme over a longer dialog made it more realistic. On the other hand, not being able to answer to a user's questions was regarded as an unsatisfying characteristic of the artificial conversational partner [22].

In another experiment, which was done by [40], eight users had to talk to two different kinds of chatbots, one behaving more human-like and one behaving more robotic. In this context, they had to fulfill certain tasks like ordering an insurance policy or demanding an insurance certification. All of the participants instinctively started to chat by using natural human language. In cases in which the bot did not respond to their queries in a satisfying way, the users' sentences continuously got shorter until they ended up with writing key words only. Thus, according to the results of this survey, conversational agents preferably should be created human-like, because users seem to be more comfortable when feeling like talking to another human being, especially in cases in which the concerns are crucial topics like their insurance policies [40].

Dialog management strategies (DM) define the conversational behaviors of a system in response to user message and system state [26].

In industry applications, DM often consists of a handcrafted set of rules and heuristics, which are tightly coupled to the application domain [26] and improved iteratively. One problem with handcrafted approaches to DM is that it is challenging to anticipate every possible user input and react appropriately, making development resource-intensive and error-prone. But if few or no recordings of conversations are available, these *rule-oriented* strategies may be the only option.

As opposed to the rule-oriented strategies, data-oriented architectures work by using machine learning algorithms that are trained with samples of dialogs in order to reproduce the interactions that are observed in the training data. These statistical or heuristic approaches to DM can be classified into three main categories: Dialog modeling based on *reinforcement learning*, *corpus-based* statistical dialog management, and *example-based* dialog management (simply extracting rules from data instead of manually coding them) [26,41]. [41] highlights neural networks, Hidden-Markov Models, and Partially Observable Markov Decision Processes as possible implementation technologies.

The following are common strategies for rule-based dialog management:

- Finite-state-based DM uses a finite state machine with handcrafted rules, and performs well for highly structured, system-directed tasks [26].

- Frame-based DM follows no predefined dialog path, but instead allows to gather pieces of information in a frame structure and no specific order. This is done by adding an additional entity-value slot for every piece of information to be collected and by annotating the intents in which they might occur. Using frames, a less restricted, user-directed conversation flow is possible, as data is captured as it comes to the mind of the user [37].
- Information State Update represents the information known at a given state in a dialog and updates the internal model each time a participant performs a *dialog move*, (e.g. asking or answering). The state includes information about the mental states of the participants (beliefs, desires, intentions, etc.) and about the dialog (utterances, shared information, etc.) in abstract representations. Using so-called update moves, applicable moves are chosen based on the state [43].
- Agent-based DM uses an agent that fulfills conversation goals by dynamically using plans for tasks like intent detection and answer generation. The agent has a set of beliefs and goals as well as an information base which is updated throughout the conversation. Within this information framework the agent continuously prioritizes goals and autonomously selects plans that maximize the likelihood of goal fulfillment [29].

[6] describes how multiple DM approaches can be combined to use the best strategy for specific circumstances.

A virtual insurance conversational agent is described by [46], utilizing *TEATIME*, an architecture for agent-based DM. *TEATIME* uses emotional state as a driver for actions, e.g. when the bot is perceived unhelpful, that emotion leads the bot to apologize. The shown example bot is a proof of concept for *TEATIME* capable of answering questions regarding insurance and react to customer emotions, but does not implement a full business process.

[25] describe a text-based healthcare chatbot that acts as a companion for weightloss but also connects a patient with healthcare professionals. The chat interface supports non-textual inputs like scales and pictorials to gather patient feedback. Study results showed a high engagement with the chatbot as a peer and a higher percentage of automated conversation the longer the chatbot is used.

Overall, these examples show potential for conversational agents in the insurance area, but lack support for complete business processes.

Considering *hackathons* previous research has been done on (examples include [27] and [3]). Important for hackathons are *goals* of a hackathon as well as *success factors*. Hackathons are problem-focused computer programming events in which teams of programmers and other stakeholders prototype a software solution within a limited timeframe [3]. Hackathons usually are characterized by three features: (1) intensive collaborative work experience (2) solution of a concrete problem with a demonstrable solution (3) and a short time span. Depending on the focus and target group several specific formats are possible, like internal or external or application or technology specific hackathons [3]. Much work apart from the work cited here on hackathons has been published. To the best of our knowledge, an internal but company-spanning conversational agent hackathon in the insurance industry has not been described yet. We will compare the resulting prototypes based on the same technology of the hackathon with the

prototype developed within a traditional project setting for deriving potentials and success factors for conversational agent creation. Furthermore, we will compare technological progress of the resulting prototypes.

3 Theory on Conversational Agents and Insurance Industry

3.1 Application Scenarios and Types of Agents

The idea of conversational agents that are able to communicate with human beings is not new: In 1966, Joseph Weizenbaum introduced *Eliza*, a virtual psychotherapist, which was able to respond to user queries using natural language and which could be considered as the first *chatbot* [45]. However, Eliza used quite simple structures by just picking up keywords and asking more questions, not serving a purpose itself. Nowadays, the idea of speaking machines has experienced a revival with the emergence of new technologies, especially in the area of artificial intelligence. Novel machine learning algorithms allow developers to create software agents in a much more sophisticated way and in many cases they already outperform previous statistical NLP methods [26]. Additionally, the importance of messaging apps such as WhatsApp or Telegram has increased over the last years. In 2015, the total number of people using these messaging services outran the total number of active users in social networks for the first time. Today, each of these app has about between 200 million and 1.5 billion users [19]. Currently the topic voice is on the rise - not only Gartner considers the breakthrough of voice applications in the next years.

Conversational agents can be basically employed in these settings:

- **Customer Service.** In 2016 [16] the topic of customer service chatbots lead to a great variety with a wide range of terminology.
- **Recruitment.** Recruitment chatbots become more popular, also the insurance company Allianz launched a recruitment bot recently.
- **Marketing.** Chatbots can be used for giving a company an innovative and up-to-date view without really serving a business process.
- **Internal Support.** Before chatbots became so popular, many companies already used chatbots for internal purposes. One example is IBM with its’ “Whatis Bot”¹, which answered questions by instant messaging about acronyms already many years ago. The requirements for internal chatbots tend to be lower than for external ones, as customers usually have the choice of a communication channel or provider.

This paper focuses on *customer service* only for demonstration purposes and simple explainability. However, in the insurance project mentioned beforehand, the second category of internal support by NLP chatbots or voice systems has gathered even more interest.

For being able to draw a big picture of the current trends in the area of conversational agents, we divide them into the following four common categories:

¹ https://www.academia.edu/35150361/IBM_whatIs.

- **(Virtual, Intelligent, Cognitive, Digital, Personal) assistants (VPAs)**. Agents fulfilling tasks intelligently based on spoken or written user input and with the help of data bases and personalized user preferences [7] (e.g. Apple’s Siri or Amazon’s Alexa [8]).
- **Specialized Digital Assistants (SDAs)**. Focused on a specific domain of expertise, goal-oriented behavior [8]. SDAs can be used in customer service as well as for internal support tasks.
- **Embodied Conversational Agents (ECAs)**. Visually animated agents, e.g. in form of avatars or robots [34], where speech is combined with gestures and facial expressions.
- **Chatterbots**. Bots with focus on small talk and realistic conversations, not task-oriented, e.g. Cleverbot [5].

Figure 1 shows the results of evaluating these four classes in terms of different characteristics such as *realism* or *task orientation* based on own literature research. Chatterbots provide a high degree of entertainment since they try to imitate the behavior of human beings while chatting, but there is no specific goal to be reached within the scope of these conversations. In contrast, general assistants like Siri or Alexa are usually called by voice in order to fulfill a specific task. Specialized assistants concentrate even more on achieving a specific goal, which often comes at the expense of realism and user amusement because their ability to respond to not goal-oriented conversational inputs like small talk is mostly limited. The best feeling of companionship can be experienced by talking to an embodied agent, since the reactions of these bots are closest to human-like behavior.

Taking a look at the insurance project, it was decided to create prototypes for customer service in the type of *specialized digital assistants*. In the next paragraph, the processes in the insurance domain which might be chosen for this implementation are described. As shown in Fig. 1, it has shown that although the goal was to create a specialized digital assistant, humans have their own goals in prototype creation. Adding small talk in a limited scope affected the prototype creation and led to a more realistic and human-like user experience and more entertainment for the prototype in the hackathon as well.

3.2 Insurance Processes and Requirements for Prototypes

Insurance is an important industry sector in Germany, with 560 companies that manage about 460 million policies [39]. However, the insurance sector is under a high cost pressure, which shows in a declining employee count and low margins [42]. The insurance market is saturated and has transitioned from a growth market to a displacement market [1]. For the greater part, German insurance companies have used conservative strategies, caused by risk aversion, long-lived products, hierarchical structures, and profitable capital markets [47]. As these conditions change, so must insurance companies. One effort is the insurance project [35] with the goal of innovation and new technologies performed by Fraunhofer IAO since several years as described in Sect. 1. The two touch points of interest in the insurance industry are *selling a product* and *the claims process*. A study found that consumers interact less with insurers than with any other

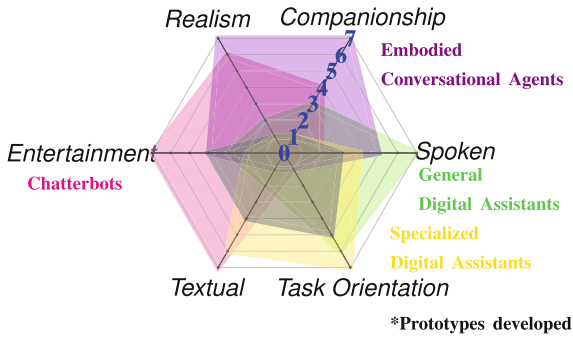


Fig. 1. Classification of conversational agents with their characteristics (based on own presentation from [23]). Values between 0 and 7 indicate how strong a characteristic applies for the given type of agent. Additionally to the classification, the prototype implementations are shown by the black box.

industry [30]. This is the reason why although chatbots become more popular in other use cases like recruitment, the focus of this paper is the application in customer service.

Many insurance companies have heterogeneous IT infrastructures incorporating legacy systems (sometimes from two or more companies as the result of a merger) [44]. These grown architectures pose challenges when implementing new data-driven or AI solutions, due to issues like data quality, availability and privacy. Nonetheless, the high amount of available data and complex processes make insurance a prime candidate for machine learning and data mining. The adoption of AI in the insurance sector is in early stages, but accelerating, as insurance companies strive to improve service and remain competitive [31].

Conversational agents are one AI technology at the verge of adoption. In 2017, ARAG launched a travel insurance chatbot, quickly followed by bots from other insurance companies [15]. Examples are a chatbot on moped insurance by wgv² and a chatbot on car insurance by Allianz³.

To identify areas of possible chatbot support, we surveyed the core business processes of insurance companies as described in [1] and [18]. Three core areas of insurance companies are customer-facing: *marketing/sales*, *contract management* and *claim management*. Figure 2 shows the main identified processes related to this area.

We identified all these processes as possible use cases for conversational agent support, in particular support by SDAs. As two prototypes are planned, the criteria are analyzed for both settings. The chosen scenario for prototype 1 is a special case of the damage claim process: *The user has a damaged smartphone or tablet and wants to make an insurance claim*. The scenario for prototype 2 in the hackathon is: *The user has received an annual bill. Answer frequently asked questions concerning the annual bill for car insurance and combine with change of personal data and cross selling activities* (see also Fig. 2).

² <https://www.wgv.de/versicherungen/kfz/moped/>.

³ <https://www.facebook.com/AllianzCarlo/>.

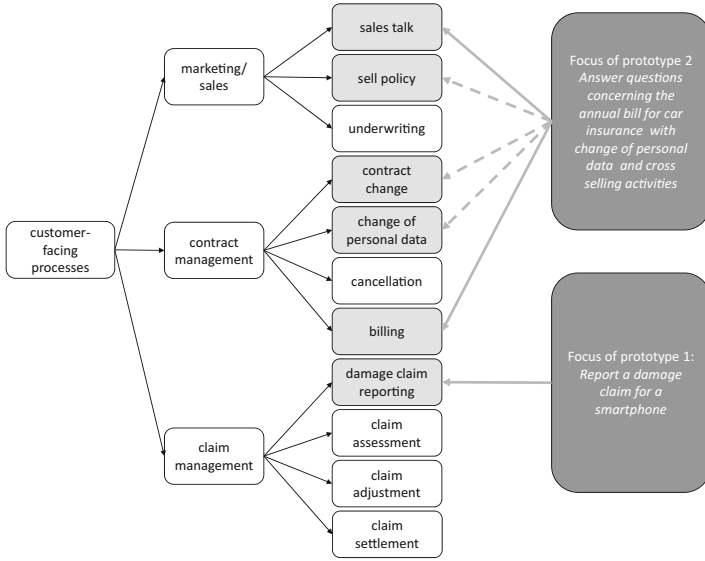


Fig. 2. Customer-facing insurance processes (original in [23] based on [1] and [18]) with additional information on the prototypes as described in this paper shown in grey.

Furthermore, we investigated general requirements for conversational agents in these processes:

Availability and Ease-of-use. Conversational agents are an alternative to both conventional customer support (e.g. phone, mail) as well as conventional applications (e.g. apps and websites). Compared to these conventional solutions, chatbots offer more availability than human agents and have less barriers of use than conventional applications, requiring neither an installation nor the ability to learn a new user interface, as conventional messaging services are used [10]. This includes the requirement of understanding and answering to human language, which applies to both prototypes developed.

Guided Information Flow. Compared to websites, which offer users a large amount of information they must filter and prioritize themselves, conversational agents offer information gradually and only after the intent of the user is known. Thus, the search space is narrowed at the beginning of the conversation without the user needing to be aware of all existing options. This is done for both prototypes by narrowing the scope.

Smartphone Integration. Using messaging services, conversational agents can integrate with other smartphone capabilities, e.g. making a picture, sending a calendar event, setting a reminder or calling a phone number. This applies for both prototypes.

Customer Call Reduction. Customer service functions can be measured by reduction of customer calls and average handling time [16]. SDAs can help here by automating conversations, handling standard customer requests and performing parts of conversations (e.g. authentication). This is relevant for projects, but out of scope for the prototype. However, questions about the annual bill arise very frequently.

Human Handover. Customers often use social media channels to escalate an issue in the expectation of a *human* response, instead of an *automated* one. A conversational agent thus must be able to differentiate between standard use cases it can handle and more complicated issues, which need to be handed over to human agents [28]. One possible approach is to use sentiment detection, so customer who are already stressed are not further aggravated by a bot [16]. Being out of scope for the prototype, this has only be investigated for some technology providers that have different levels of experience with this question.

Digitize Claim Handling. Damage claim handling in insurance companies is a complex process involving multiple departments and stakeholders [24]. Claim handling processes are more and more digitized within the insurance companies [18], but paper still dominates communication with claimants, workshops and experts. [12] defines maturity levels of insurance processes, defining *virtual handling* as a process where claims are assessed fully digitally based on digital data from the claimant (e.g. a video, a filled digital form), and *touchless handling* as a fully digital process with no human intervention on the insurance side. SDAs help moving towards these maturity levels by providing a guided way to make a claim digitally and communicate with the claimant (e.g. in case additional data is needed). Prototype 1 covers this area.

Conversational Commerce. Is the use of Conversational Agents for marketing and sales related purposes [11]. Conversational Agents can perform multiple tasks using a single interface. Examples are using opportunities to sell additional products (*cross-sell*) or better versions of the product the customer already has (*up-sell*) by chiming in with personalized product recommendations in the most appropriate situations. One example would be to note that a person's last name has changed during an address update customer service case and offer appropriate products if the customer has just married. Prototype 2 covers this area.

Internationalization. Is an important topic for large international insurance companies. However, most frameworks for implementing conversational agents are available in more than one language. To the best of our knowledge, the applied conversational agents in German insurance today are optimized only for one language. So this topic is future work in respect to both prototypes, but will become more important in the future.

Compliance. to privacy (GDPR) is usually guaranteed by the login mechanisms on the insurance sites, therefore the topic is out of scope for our research prototype. For broader scenarios not requiring identification on the insurance site and the usage of the data for non-costumers, this is an area of ongoing research on compliant technical solutions or workarounds.

4 Practice in Two Prototypes in Insurance Industry

4.1 Technical Requirements and Framework Options for the Prototypes

For dialog design within prototype 1, experimenting with machine learning algorithms was the preferred implementation strategy. For this purpose, discussions with insurance companies were held to assess the feasibility of receiving existing dialogs with

customers, for example for online chats, phone logs or similar. However, such logs generally seem to be not available at German insurers, as the industry has self-regulated to only store data needed for claim processing [14]. As a research institute represents a third party not directly involved in claims processing, data protection laws forbid sharing of data this way without steps to secure personal data. During our talks we have identified a need for automated or assisted anonymization of written texts as a precondition for most customer-facing machine learning use cases, at least when operating in Europe [20]. However, these issues go beyond the scope of our current project, but provide many opportunities for future research.

To still build a demonstrator in face of these challenges as outlined in [2], dialogs for both prototypes were manually designed without using real-life customer conversations and fine-tuned by user testing with fictional issues. As this approach entails higher manual effort for dialog design, a narrower scenario was chosen for both prototypes to still allow for the full realization of a customer-facing process.

Based on the work presented in the last sections and our talks with insurance companies, we arrived at the following non-functional requirements that the chatbot prototype 1 ideally should fulfill:

- **Interoperability:** The agent should be able to keep track of the conversational context over several message steps and messengers.
- **Portability:** The agent can be run on different devices and platforms (e.g. Facebook Messenger, Telegram). Therefore it should use a unified, platform-independent messaging format.
- **Extensibility:** The agent should provide a high level of abstraction that allows designers to add new conversational content without having to deal with complicated data structures or code.

For natural language understanding, we compared four possible frameworks (Microsoft’s LUIS, Google’s Dialogflow, Facebook’s wit.ai and IBM’s Watson) regarding important criteria for prototype implementation in a first step just for prototype 1. The comparison was extended for the frameworks moni.ai and Kauz.net for prototype 2 in the hackathon. All six frameworks support textual input and output, this was amongst others a basic requirement, but not all support complex conversation flows for advanced use cases. A comparison table for these criteria is shown in Table 1. As a result of the comparison, Google Dialogflow was chosen as a basic framework for prototype 1 based on the fulfillment of all requirements of prototype 1, one of which was the free availability. For prototype 2 and the hackathon, Google Dialogflow and IBM Watson Assistant were chosen, as an important factor next to the available user interface to enable non-programmers to work with the software was the fact that the providers agreed to accompany the hackathon event by sending experts for local support.

4.2 Prototype 1: Claim Management with Technological Extensions

Prototype 1 fulfills the following scenario: *The user has a damaged smartphone or tablet and wants to make an insurance claim.* The goal here is to focus on technology and build a demonstratable prototype in a ‘traditional’ project setting. We describe

Table 1. Comparison of Microsoft’s LUIS, Google’s Dialogflow, Facebook’s wit.ai, and IBM’s Watson (from the requirements for the technical prototype (1) as in the original paper [23], based on [9]) and extended for prototype 2 by additional hackathon requirements and two new providers nameley moni.ai and Kauz.net (n.c. stands for Not Considered anymore or not yet for the prototype 1 or 2 as of early in 2018).

Requirement		LUIS	Dialogflow	Wit.ai	Watson	moni.ai	Kauz.net
Overall (1,2)	Textual in-/output	yes	yes	yes	yes	yes	yes
Overall (1,2)	German language	yes	yes	in Beta	yes	yes	yes
Technical (1)	Python bindings	no	yes	yes	yes	n.c.	n.c.
Technical (1)	Free service	no	yes	yes	partly ^a	n.c.	n.c.
Technical (1)	Remember state	yes	yes	yes	yes	n.c.	n.c.
Technical (1)	Service bound	yes	yes	yes	yes	n.c.	n.c.
Technical (1)	Simple training	partly	yes	yes	yes	n.c.	n.c.
Hackathon (2)	Complex conversation flows	no	yes	n.c.	yes	n.c.	n.c.
Hackathon (2)	Provider support	n.c.	yes	n.c.	yes	no	yes
Hackathon (2)	User Interface	yes	yes	n.c.	yes	yes	no

^a10 000 free messages per month

the results technically in the following. Figure 6 shows the main components of the prototype and their operating sequence when processing a user message. To provide extensibility prototype architecture strictly separates service integration, internal logic and domain logic.

The user can interact with the bot over different communication channels which are integrated with different *bot API clients*. To integrate a different messaging service, a new bot API client needs to be written. The remainder of the prototype can be reused. See Fig. 3 for an example of the prototype on different communication channels.

Once a user has written a message, a lookup of *user context* is performed to determine if a conversation with that user is already in progress. User context is stored in a database so no state is kept within external messaging services. Afterwards, a *typing* notification is given to the user, indicating the bot has received the message and is working on it. This prevents multiple messages by a user who thinks the bot is not responsive.

In the next step, the message has to be understood by the bot. In case of a voice message, it is transcribed to text using a *Google* speech recognition web service. Dialogflow is used for intent identification, which determines the function of a message and based on that a set of possible parameters [26]. For example, the intent of the message “the display of my smartphone broke” may have the intent `phone_broken` with the parameter `damage_type` as `display_damage`, while the parameter `phone_type` is not given. Together, this information given by Dialogflow is a `MessageUnderstanding`.



Fig. 3. Top: The system is mirrored on both the Facebook and Telegram Messengers. Bottom left: Additional view with customer data input and intelligent recognition of words like *yesterday*. Bottom right: Dialog excerpt of the prototype, showing the possibility to clarify the phone model via multiple-choice input. Extended version of figures in [23].

As soon as the message is understood, the user context is updated. Afterwards, a response needs to be generated. This process, which was labeled with *Plan and Realize Response* in Fig. 6, is shown in detail in Fig. 7.

In the prototype, an agent-based strategy was chosen in order to combine the capabilities of the frame-based entities and parameters in Dialogflow with a custom dialog controller based on predefined rules in a finite state machine. This machine allows to define rules that trigger *handlers* and *state transitions* when a specific intent or entity-parameter combination is encountered. That way, both intent and frame processing happen in the same logically encapsulated unit, enabling better maintainability and extensibility. The rules are instances of a set of `*Handler` classes such as an `IntentHandler` for the aforementioned intent and parameter matching, supplemented by other handlers, e.g. an `AffirmationHandler`, which consolidates different intents that all express a confirmation along the lines of “yes”, “okay”, “good” and “correct”, as well as a `NegationHandler`, a `MediaHandler` and an `EmojiSentimentHandler` (to analyze positive, neutral, or negative sentiment of a message with emojis). Each implements their own `matches (Message Understanding)` method.

The following types of rules (handlers) are used within the dialog state machine:

1. *Stateless* handlers are checked independently of the current state. For example, a `RegexHandler` rule determines whether the formality of the address towards the user should be changed (German differentiates the informal “du” and the formal “Sie”)
2. *Dialog States* map each possible state to a list of handlers that are applicable in that state. For instance, when the user has given an answer and the system asks for *explicit confirmation* in a state `USER_CONFIRMING_ANSWER`, then an `AffirmationHandler` and a `NegationHandler` capture “yes” and “no” answers.
3. *Fallback* handlers are checked if none of the applicable state handlers have yielded a match for an incoming `MessageUnderstanding`. These fallbacks include static, predefined responses with lowest priority (e.g. small talk), as well as handlers to repair the conversation by bringing the user back on track or changing the topic.

At first, the system had only allowed a single state to be declared at the same time in the router. However, this had quickly proven to be insufficient as users are likely to want to respond or refer not only to the most recent message, but also to previous ones in the chat. With only a single contemporaneous state, the user’s next utterance is always interpreted only in that state. In order to make this model resilient, every state would need to incorporate every utterance that the user is likely to say in that context. As this is not feasible, the prototype has state handlers that allow layering transitions on top of each other, allowing multiple simultaneous states which may advance individually.

To avoid an explosion of active states, the system has *state lifetimes*: new states returned by callbacks may have a lifetime that determines the number of dialog moves this state is valid for. On receiving a new message, the planning agent decreases the lifetimes of all current dialog states by one, except for the case of utter non-understanding (“fallback” intent). If a state has exceeded its lifetime, it is removed from the priority queue of current dialog states.

Figure 7 contains details about how the system creates responses to user queries. Based on the applicable rule, the conversational agent performs chat actions (e.g. sending a message), which are generated from response templates, taking into account dialog state, intent parameters, and information like a user's name, mood and preferred level of formality.

RuleHandlers, states and other dialog specific implementations are encapsulated, so a new type of dialog can be implemented without needing to change the other parts of the system.

Generated chat actions are stored in the user context and performed for the user's specific messenger using the *bot API*. As the user context has been updated, the next message by the user continues the conversation.

The prototype explains its functionality and offers limited small talk. As soon as the user wants to make a damage claim, a predetermined questionnaire is used about type of damage, damaged phone, phone number, IMEI, damage time, damage event details, etc. Interpretation results of answers have to be confirmed by the user. For specific questions domain specific actions for clarification are implemented (see bottom right in Fig. 3). In a real-life application, claim management systems would be integrated to automatically trigger subsequent processes.

4.3 Prototype 2: Customer Service and Cross-selling in Hackathon

For receiving more insights about the practicability of introducing chatbots to the insurance domain and for gaining experience with the usage of conversational frameworks, Fraunhofer IAO organized a four-day hackathon with five German insurance companies participating [21]. This results in prototype 2 for the given task: *Create a chatbot using IBM Watson Assistant or Google Dialogflow for answering questions about the annual bill of car insurance and leveraging cross-selling opportunities.*

In the scope of the event four minimal products were created by four interdisciplinary teams of IT specialists, sales experts and other employees of the insurance companies. Doing so, in contrast to posing the challenge to external developers, our insurance partners were directly involved and could profit from the lessons learned from this *internal hackathon* [36]. One impression from the resulting video is shown in Fig. 4 and also described in a blog article [21].

Following four prototypes can be characterized as the teams worked independently:

Prototype A Voice Focus. One more technically oriented team started by adding voice technology to the chatbot for output purposes. Analogously, voice input could be used - although findings in the insurance project show that the input direction is more difficult to handle than the output direction. It showed that the focus is very entertaining in presentation and that the presentation especially of voice technology has to be performed carefully. In addition, the chatbot has been made more human-like by adding personal opinions on sports.

Prototype B Multimedia Focus. Team B integrated several resources for better multimedia presentation, like images, videos, and the like. This already started with using an QR-code for accessing the prototype. The idea of using sophisticated multimedia content for explaining the annual bill like clickable graphics with videos has

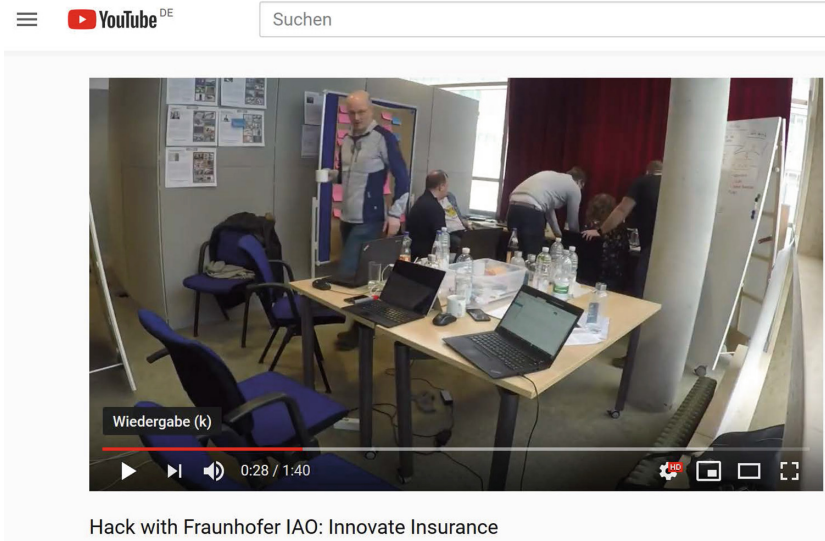


Fig. 4. Hackathon impressions (www.youtube.com/watch?v=yHRLYJ.olZ8, [21]).

impressed the jury. The team focused therefore on customer experience and fine-tuned their interaction patterns by introducing delay in the response times.

Prototype C Stability and Scope Focus. The C-team focused on building a stable prototype for the complete task, achieving a large coverage of topics. The team spent most time in designing entities and intents as well as dialogue flow. This led to a comprehensive design and the most resilient result. The team was successful in maintaining background knowledge in a database and integrate it into the conversation flow.

Prototype D Customer Identification Focus. The team D focused on solving the customer identification issue. Using this information, they could give very detailed information on the current contract of the customer and the bill and use customer specific information for guiding the conversation itself. Another demonstration was the change of customer data. Additionally, some small talk was introduced for entertainment purposes.

Concerning the results, it is worth mentioning that all four groups succeeded in creating a usable product within the given timeframe that was able to handle the required use case of answering questions about annual bills for a small set of predefined queries. However, when letting a chatbot talk to people of other groups who were not involved in its development and thus not aware of the underlying dialog structure, the solutions proved to be error prone since they could not handle these unexpected inputs. This stems from *hardcoding* parts of the scenario due to time constraints. In some cases, more expertise in dialog design would have helped anticipate typical user inputs. All teams worked with the *entities* and *intents* as are defined in most chatbot technologies, adding no programmable extensions (as compared to prototype 1). Therefore, the

dialogue structure is static and creating the chatbot is done by adding intents and entities as well as *if-then* like programming of the dialogue. The teams got only the plain frameworks of the providers and no specific extensions as for example for thorough testing. Working in teams on a chatbot proved to be helpful, but added organizational complexity, since resource access had to be shared and organized. Additionally, it turned out that professional content designers that build appropriate conversation models may be even more important than programmers, at least in case the used technology is enhanced enough. A final survey among the participants showed that they enjoyed working in the chosen hackathon format and could benefit a lot from its results and lessons learned.

5 Evaluation

Both prototypes were evaluated using appropriate methods. For prototype one, a questionnaire-based approach with 14 participants was chosen. For prototype two, an expert commission had to choose and rate all four prototypes based on a very short questionnaire and come up with a point rating in a very short time span, as is typical for hackathons. We will first describe the results of both evaluation processes and then compare the results in the overall conclusion of this paper (see Sect. 6).

5.1 Evaluation of Prototype 1 Claim Management

To evaluate the produced prototype's quality and performance, we conducted a model trial with the goal to report a claim by using the chatbot without having any further instructions available.

Of the 14 participants (who all had some technical background), 35.7% claimed to regularly use chatbots, 57.1% to use them occasionally, and only 7.1% stated that they had never talked to a chatbot before. However, all participants were able to report a claim within a range of about four minutes, resulting in an overall task completion rate of 100%.

Additionally, the users had to rate the quality of their experiences with the conversational agent by filling out a questionnaire. For each question they could assign points between 0 (did not apply at all) and 10 (did apply to the full extent). The most important quality criteria, whose choice was oriented on the work of [34], are listed with their average ratings in Fig. 5 and are discussed in detail.

Ease of Use. With an average of 8 points for *Ease of Use*, the users had no problems with using the bot to solve the task, since none of them gave less than 5 points. However, a variance of 2.46 still indicates a strong gap among the participants' experienced degree of usability.

Appropriate Formality. 8.3 points on average for *Appropriate Formality* indicate that the participants were comfortable with the formal and informal language the bot talked to them. Nonetheless, this criteria was also rated with points of only one and two. One of these users stated that he felt worried about permanently being called by his first name after he told it. Therefore, development of a more fine-grained detection mechanism for formal and informal language should be considered in future versions of the chatbot, since for now we only rely on simple regular expressions.

Natural Interaction. The rating for convincing *Natural Interaction* with 7.9 points may be due to the fact that the conversation was designed in a strongly questionnaire-oriented way, which might have restricted the feeling of having a free user conversation. Therefore, improving the flexibility of the conversational flow and granting more freedom for user-centric dialog control might strengthen the authentic feeling during interaction with the agent.

Response Quality. The satisfaction with given answers to users' domain specific questions was considered quite (but not totally) convincing with 7.6 points. Note that the high number of points might not be justified entirely, because the chatbot's implemented ability to answer questions is still very basic and restricted to concerns of claim handling. But, since the whole conversation is strongly driven by the agent itself, the users probably didn't find the time to ask many questions that went beyond the current limits of understanding. Connecting any kinds of knowledge bases might serve as a first future step towards extending the agent's response qualities.

Personality. The least convincing experience was that chatbot's *Personality*, which was rated with only 5.2 points on average. This is not surprising, since during this work we put comparatively less efforts in strengthening the agent's personal skills as it does not even introduce itself with a name, but instead mainly acts on a professional level, always concentrating on the fulfillment of its task. Facing these facts, a professional copywriter should have no problems developing a more convincing character for the chatbot.

Funny and Interesting. With 7.2 points, talking to the chatbot was experienced as quite *Funny & Interesting*, but still with a lot of room for further improvement. Again, the key here stays to loosen the strict procedure of forcing the user to finish the process and to allow more room for smalltalk and off topic contents.

Entertainment. The agent's *Entertainment* capabilities, which are at 7.7 points on average, could be upgraded by extending the conversational contents with additional enjoyable features not related to the questionnaire. At the moment, the chatbot is only able to tell some jokes from the insurance domain, but does not provide a holistic concept for customer entertainment.

No Deception Feeling. The agent's *lack of deceptiveness*, i.e. the degree to which users know it is not human, which at 9.6 points show that the bot's statements made its nature clear to users.

5.2 Evaluation of Prototype 2 as in the Hackathon

Typically, the evaluation in hackathons is done by a very short demo and questions from the audience by an expert committee. After four days, the resulting prototypes were examined by a jury considering the following predefined criteria. These differ strongly from the criteria in the first evaluation due to the time-constrained focused question. Most of the questions tackle the aforementioned *response quality*, with a second thought on *natural interaction*. They clarify what is actually a focus of the hackathon and what is not, putting emphasis on getting done. The *formality* is just a subcriterion if the language style is adapted. All further going criteria like personality, fun, entertainment

etc. have explicitly not been stated, but it is interesting that all groups put a strong focus on this during their presentation, trying to stand out from the field and enjoying to add human-like behavior.

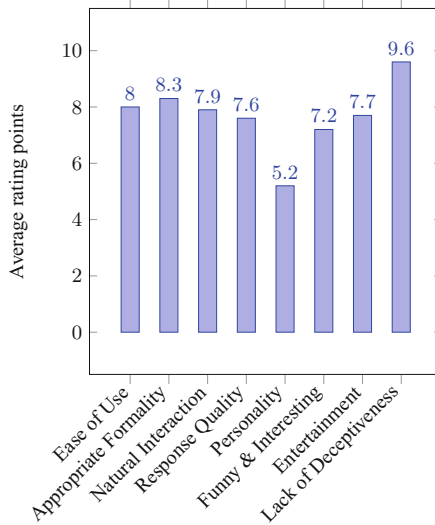


Fig. 5. Survey results for prototype 1: average user experience ratings (fourteen participants, 0..10 points) from [23].

Language Support

- Is the chatbot able to recognize the user input?
- Are the outputs of the chatbot adequate and understandable?
- Is the language style used by the bot adequate and consistent?
- Is the language style used by the bot adapted individually to user properties?

Flexibility

- Is the chatbot able to correctly recognize input even in unusual phrases?
- Is the chatbot able to respond appropriately to unexpected input?

Scope of Functions

- How well has the scenario *Annual fee bill* been covered by the prototype?
- How well are extensions and transitions implemented leading to other topics such as cross selling?

Presentation

- Is the presentation convincing? Are there any differences to the other groups?
- Did the team manage to explain the chatbot in a timely manner?

The expert committee had difficulties finding a winner in the given time span, as all of the four prototypes fulfilled certain aspects that were identified as interesting to the jury. Altogether the participants and experts learned a lot about chatbot design and technologies and were satisfied with the results. Both technologies led to good results in the hackathon, but most of the participants felt they needed provider support and expert knowledge on dialogue design for creating a real product for customers. More overall conclusions follow in the next section.

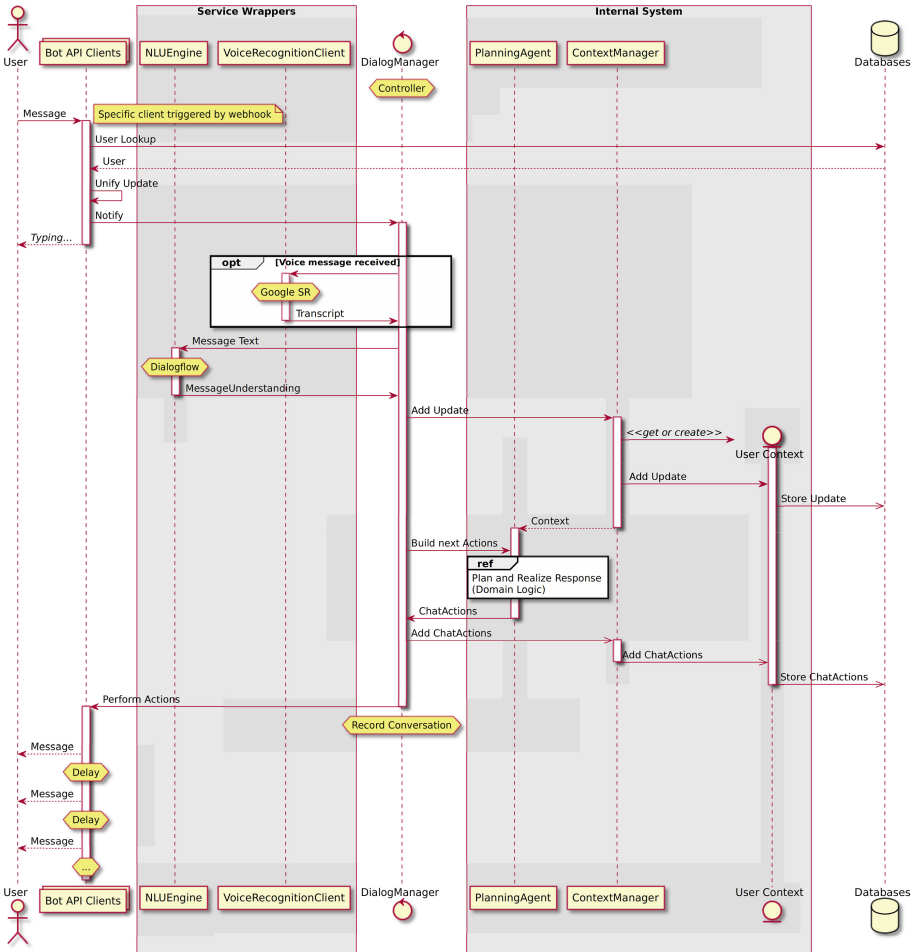


Fig. 6. Sequence diagram of the conversational agent prototype from [23].

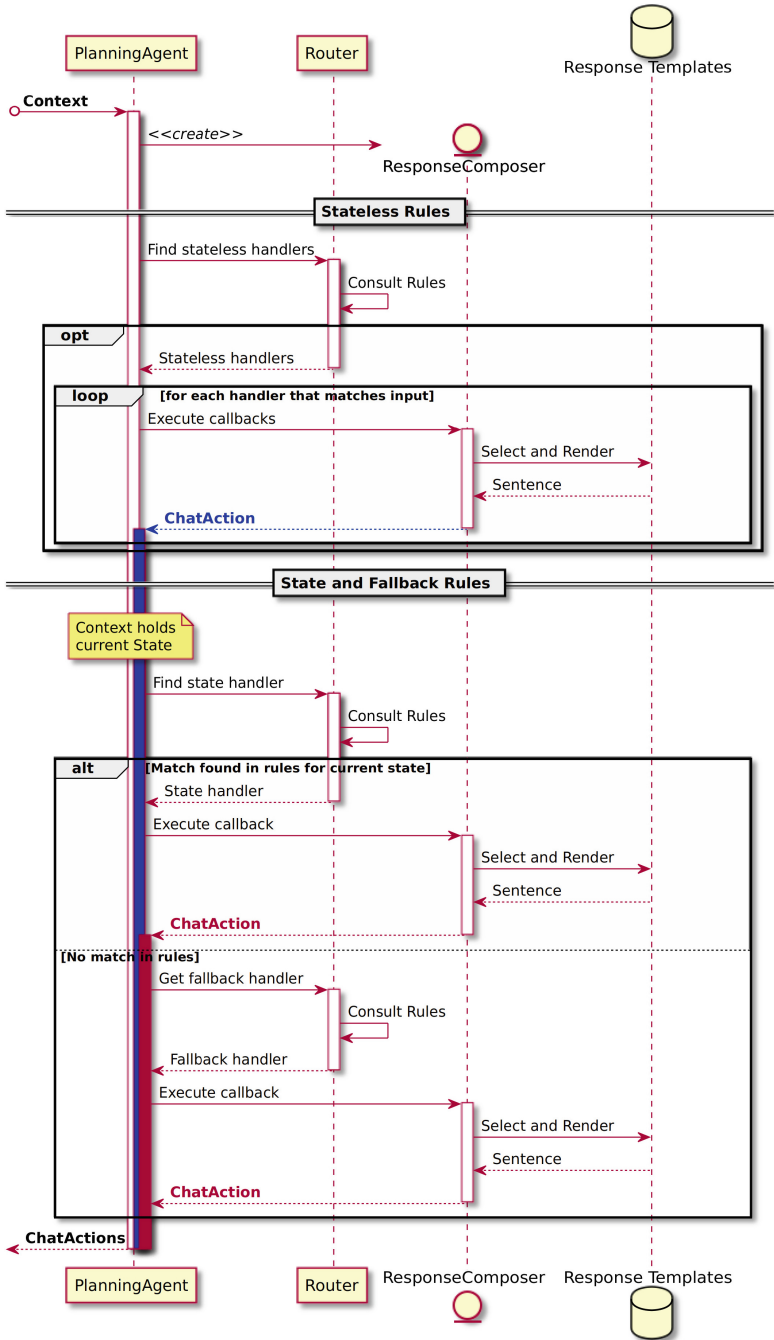


Fig. 7. Detailed sequence diagram of the response generation in the conversational agent prototype from [23].

6 Conclusions and Outlook

In this work we have shown how conversational agents can be applied for different use cases in theory and practice. We showed how our classification of conversation agents applies for the prototypes generated for our two scenarios: (1) claim management process support in traditional project setting as well as (2) customer service and cross selling in an interdisciplinary hackathon. The potential processes to employ chatbots have been shown in general for the insurance companies and focused on customer service processes. One key result of our former paper [23] containing prototype 1 is a system of multiple conversational states enabling more flexible conversations. We extended the evaluation with real users and additionally showed the prototype to various customer groups, small businesses and more insurance companies. Altogether, the prototype is able to handle the scenario satisfactory. One possible improvement is the point of *realism*, for example by more human-like behavior in a consistent persona and better determination of the desired degree of formality. The newly performed and described activity is the creation of prototype 2 for customer service consisting of actually four prototypes in an interdisciplinary hackathon. It has shown that the results here differ strongly from prototype 1 due to different goals, different time span and different skills. No extensions to the entity-intent concept were performed, but several innovative ideas have been included like multimedia integration, voice integration, several entertaining aspects and especially persona design. Prototype 1 did not have a name in the beginning, whereas all teams came up with innovative names for prototype 2 at the beginning of the hackathon.

Altogether, after two years of chatbot experience, we can summarize the potentials for the conversational agent technology:

Maturity of Technology. Technology matures and is more often perceived in all day activities, most people know chatbots and how they can be used. Many people have already tried out a chatbot.

Service Enhancement. Agents can be used for better availability (24/7/365) and to reduce the workload of the service staff.

Tools. Tools are available especially for good English language support. More languages and features are added as time passes. New technological frameworks are available, the existing ones are improved.

Simple Tasks. Easy application for simple tasks and simple prototype creation is possible in a short time span. Transfer from prototype to live system is still more difficult.

Applicability. Many application scenarios are possible.

We have identified following challenges for new conversational agents and especially for transitioning from demo to prototype to live service:

Testing. Is time-consuming and error-prone

Domain Language. Has to be usually hardcoded or added manually supported by machine learning in an optimal case

Handovers. Designing handovers is a challenging tasks that not all frameworks fulfill perfectly

Maintenance. Of the chatbot and further development is a challenging tasks and the process has to be defined

Self-learning. Is not available in the expected scope as lots of people have very high and unrealistic expectations to machine learning

High Expectations. To chatbot technology in general which might not be fulfilled in the beginning

Security and Integration. Issues as with most technologies

As a result of our two prototypes, the evaluations and the hackathon participant surveys, we came up with the following success factors that we believe need to be respected when planning to introduce conversational agents to companies:

Clear Scope Definition. Use cases and functionality of the conversational agent should be predefined as detailed as possible.

Customer-oriented development. Tests with intended audience and changing test participants to prevent them getting used to the dialog structure.

Careful Improvements and Testing. Sufficient time and care should be invested in testing and improving the agent. A nonfunctional or only partly functional bot deployed to the public too early might cause a negative reception that cannot be corrected with future improvements.

Perform Regression Tests. Especially for self-learning agents it is crucial to ensure that the bot does not “unlearn” skills that once worked successfully.

Facilitate Maintenance. Provide high-level (graphical) dialog customization options for the employees of the related department for supporting easy extension and improvement of the agent.

Choice of Technology Provider. Technology providers should be compared and chosen according to the company environment and its conditions. One partner should be selected for longer cooperation.

The model trial covering two prototypes has shown that conversational agents are ready for productive use. However, the effort in creating and maintaining a conversational agent is not to be underestimated. While a successful conversation with a chatbot provides a satisfying customer experience, errors and gaps in dialog flow let user satisfaction drop rather quickly. While users do not expect a human like conversation and phrase their statements accordingly, they expect clearly formulated requests and answers to be readily understood. Currently we are working on supporting small and medium enterprises with evaluation of the technology and potential use cases for their businesses. In future research, we would like to implement a real-life conversational agent as well as perform a real-life evaluation with an insurance partner to quantify the benefits of agent use, e.g. call reduction, success rate, and customer satisfaction as well as support small and medium businesses with agent creation.

Acknowledgements. The work is partially based on work carried out in the project ‘Business Innovation Engineering Center’, which is funded by the Ministry of Economic Affairs, Labour and Housing Baden-Wuerttemberg under the reference number 3-4332.62-IAO/56. The hackathon itself was conducted within the ‘Innovationsnetzwerk Digitalisierung fuer Versicherungen’ [35]. The authors want to thank all participants for their contributions and feedback.

References

1. Aschenbrenner, M., Dicke, R., Karnarski, B., Schweiggert, F.: Informationsverarbeitung in Versicherungsunternehmen. Springer, Heidelberg (2010). <https://doi.org/10.1007/978-3-642-04321-5>
2. Blohm, M., Dukino, C., Kintz, M., Kochanowski, M., Koetter, F., Renner, T.: Towards a privacy compliant cloud architecture for natural language processing platforms. In: Proceedings of the 21st International Conference on Enterprise Information Systems, ICEIS, vol. 1, pp. 454–461. INSTICC. SciTePress (2019). <https://doi.org/10.5220/0007746204540461>
3. Briscoe, G., Mulligan, C.: Digital innovation: the Hackathon phenomenon (2014). <http://www.creativeworkslondon.org.uk/wp-content/uploads/2013/11/Digital-Innovation-The-Hackathon-Phenomenon1.pdf>
4. Cahn, J.: CHATBOT: Architecture, Design, & Development (2017)
5. Carpenter, R.: Cleverbot (2018). <https://www.cleverbot.com/>
6. Chu, S.W., O'Neill, L., Hanna, P., McTear, M.: An approach to multi-strategy dialogue management. In: Ninth European Conference on Speech Communication and Technology, pp. 865–868 (2005)
7. Cooper, R.S., McElroy, J.F., Rolandi, W., Sanders, D., Ulmer, R.M., Peebles, E.: Personal virtual assistant. US Patent 7,415,100, 19 August 2008
8. Dale, R.: Industry watch: the return of the Chatbots. *Nat. Lang. Eng.* **22**(5), 811–817 (2016)
9. Davydova, O.: 25 Chatbot platforms: a comparative table (2017). <https://chatbotsjournal.com/25-chatbot-platforms-a-comparative-table-aeefc932eaff>
10. Derler, R.: Chatbot vs. app vs. website - chatbots magazine, December 2017. <https://chatbotsmagazine.com/chatbot-vs-app-vs-website-en-e0027e46c983>
11. Eeuwens, M.: Mobile conversational commerce: messenger Chatbots as the next interface between businesses and consumers. Master's thesis, University of Twente (2017)
12. Fannin, T., Brower, B.: 2017 future of claims study. Tech. rep. LexisNexis (2017)
13. Gartner: Market trends: mobile app adoption matures as usage mellows (2015). <https://www.gartner.com/newsroom/id/3018618>
14. GDV: Verhaltensregeln für den Umgang mit personenbezogenen Daten durch die deutsche Versicherungswirtschaft (2012). http://www.gdv.de/wp-content/uploads/2013/03/GDV_Code-of-Conduct_Datenschutz_2012.pdf. datum des Aufrufes des Dokumentes: 11.02.2015
15. Gorr, D.: Ein Versicherungsroboter für gewisse Stunden. *Versicherungswirtschaft Heute* (2018). <http://versicherungswirtschaft-heute.de/schlaglicht/ein-versicherungsroboter-fur-gewisse-stunden/>
16. Guzmán, I., Pathania, A.: Chatbots in customer service. Tech. rep. Accenture (2016)
17. Harkous, H., Fawaz, K., Shin, K.G., Aberer, K.: Pribots: conversational privacy with Chatbots. In: Twelfth Symposium on Usable Privacy and Security (SOUPS 2016). USENIX Association, Denver, CO (2016). <https://www.usenix.org/conference/soups2016/workshop-program/wfpn/presentation/harkous>
18. Horch, A., Kintz, M., Koetter, F., Renner, T., Weidmann, M., Ziegler, C.: Projekt openXchange: Servicenetzwerk zur effizienten Abwicklung und Optimierung von Regulierungsprozessen bei Sachschäden. Fraunhofer Verlag, Stuttgart (2012)
19. Inc, S.: Most popular messaging apps (2018). <https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>
20. Kamarinou, D., Millard, C., Singh, J.: Machine learning with personal data. Queen Mary School of Law Legal Studies Research Paper, no. 247 (2016)
21. Kasper, H.: Hack deine versicherungs-innovation: 4 Chatbots in 4 tagen (2018). <https://blog.iao.fraunhofer.de/hack-deine-versicherungs-innovation-4-chatbots-in-4-tagen/>

22. Kirakowski, J., O'Donnell, P., Yiu, A.: Establishing the hallmarks of a convincing chatbot-human dialogue. In: Mautua, I. (ed.) *Human-Computer Interaction*. InTech, Rijeka (2009). <https://doi.org/10.5772/7741>
23. Koetter, F., Blohm, M., Kochanowski, M., Goetzer, J., Graziotin, D., Wagner, S.: Motivations, classification and model trial of conversational agents for insurance companies. In: Rocha, A.P., Steels, L., van den Herik, J. (eds.) *Proceedings of the 11th International Conference on Agents and Artificial Intelligence, ICAART 2019, Volume 1, Prague, Czech Republic, 19–21 February 2019*, pp. 19–30. SciTePress (2019). <https://doi.org/10.5220/0007252100190030>
24. Koetter, F., Weisbecker, A., Renner, T.: Business process optimization in cross-company service networks: architecture and maturity model. In: *2012 Annual SRII Global Conference (SRII)*, pp. 715–724. IEEE (2012)
25. Kowatsch, T., et al.: Text-based healthcare Chatbots supporting patient and health professional teams: preliminary results of a randomized controlled trial on childhood obesity. In: *Persuasive Embodied Agents for Behavior Change (PEACH2017)*. ETH Zurich (2017)
26. McTear, M., Callejas, Z., Griol, D.: *The Conversational Interface*. Springer, Cham (2016). <https://doi.org/10.1007/978-3-319-32967-3>
27. Mohajer Soltani, P., Pessi, K., Ahlin, K., Wernered, I.: Hackathon - a method for digital innovative success: a comparative descriptive study. In: *Proceedings of the 8th European Conference on Information Management and Evaluation, ECIME 2014, September 2014*
28. Newlands, M.: 10 ways AI and Chatbots reduce business risks. <https://www.entrepreneur.com/article/305073>(2017)
29. Nguyen, A., Wobcke, W.: An agent-based approach to dialogue management in personal assistants. In: *Proceedings of the 10th International Conference on Intelligent User Interfaces*, pp. 137–144. ACM (2005)
30. Niddam, M., Barsley, N., Gard, J.C., Cotroneo, U.: Evolution and revolution: how insurers stay relevant in a digital future (2014). <https://www.bcg.com/publications/2014/insurance-technology-strategy-evolution-revolution-how-insurers-stay-relevant-digital-world.aspx>
31. Nordman, E., DeFrain, K., Hall, S.N., Karapiperis, D., Obersteadt, A.: How artificial intelligence is changing the insurance industry (2017)
32. Pohl, V., Kasper, H., Kochanowski, M., Renner, T.: *Zukunftsstudie 2027 #ichinzehnjahren* (2017). <http://s.fhg.de/zukunft2027>. (in german)
33. Power, J.D.: 2017 U.S. auto claims satisfaction study (2017). <http://www.jdpower.com/resource/jd-power-us-auto-claims-satisfaction-study>
34. Radziwill, N., Benton, M.: Evaluating quality of Chatbots and intelligent conversational agents. arXiv preprint [arXiv:1704.04579](https://arxiv.org/abs/1704.04579) (2017)
35. Renner, T., Kochanowski, M.: *Innovationsnetzwerk digitalisierung für versicherungen* (2018). <http://s.fhg.de/innonetz>. (in German)
36. Rosell, B., Kumar, S., Shepherd, J.: Unleashing innovation through internal hackathons, pp. 1–8, May 2014. <https://doi.org/10.1109/InnoTek.2014.6877369>
37. Rudnický, A., Xu, W.: An agenda-based dialog management architecture for spoken language systems. In: *IEEE Automatic Speech Recognition and Understanding Workshop*, vol. 13 (1999)
38. Schippers, B.: App fatigue (2016). <https://techcrunch.com/2016/02/03/app-fatigue/>
39. Schwark, P., Theis, A.: *Statistisches Taschenbuch der Versicherungswirtschaft. Gesamtverband der Deutschen Versicherungswirtschaft (GDV), Berlin* (2014)
40. Sörensen, I.: Expectations on Chatbots among novice users during the onboarding process (2017). <http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-202710>
41. Spierling, U.: Interactive digital storytelling: towards a hybrid conceptual approach. In: *Worlds in Play: International Perspectives on Digital Games Research* (2005)

42. Stange, A., Reich, N.: Die Zukunft der deutschen Assekuranz: chancenreich und doch ungewiss. In: Zimmermann, G. (ed.) *Change Management in Versicherungsunternehmen*, pp. 3–9. Springer, Wiesbaden (2015). https://doi.org/10.1007/978-3-658-05974-3_1
43. Traum, D.R., Larsson, S.: The information state approach to dialogue management. In: van Kuppevelt, J., Smith, R.W. (eds.) *Current and new directions in discourse and dialogue*, pp. 325–353. Springer, Dordrecht (2003). https://doi.org/10.1007/978-94-010-0019-2_15
44. Weindelt, B.: Digital transformation of industries. Tech. rep. World Economic Forum and Accenture (2016)
45. Weizenbaum, J.: Eliza - a computer program for the study of natural language communication between man and machine. *Commun. ACM* **9**(1), 36–45 (1966). <https://doi.org/10.1145/365153.365168>
46. Yacoubi, A., Sabouret, N.: Teatime: a formal model of action tendencies in conversational agents. *ICAART* **2**, 143–153 (2018)
47. Zimmermann, G., Richter, S.-L.: Gründe für die Veränderungsaversion deutscher Versicherungsunternehmen. In: Zimmermann, G. (ed.) *Change Management in Versicherungsunternehmen*, pp. 11–35. Springer, Wiesbaden (2015). https://doi.org/10.1007/978-3-658-05974-3_2