Towards a Taxonomy of Platforms for Conversational Agent Design

Stephan Diederich
Georg-August-Universität Göttingen

Alfred Benedikt Brendel
Technische Universität Dresden

Lutz Kolbe
Georg-August-Universität Göttingen

Diederich, S., Brendel, A.B., & Kolbe, L.M. (2019). Towards a Taxonomy of Platforms for Conversational Agent Design. In Proceedings of Internationale Tagung Wirtschaftsinformatik, Siegen, Germany.

Chair of Information Management

Prof. Dr. Lutz M. Kolbe
Platz der Göttinger Sieben 5
37073 Göttingen – Germany
www.uni-goettingen.de/im

Smart Mobility Research Group (SMRG)

Dr. Alfred Benedikt Brendel
Humboldtallee 3
37073 Göttingen – Germany
www.uni-goettingen.de/smrg

Towards a Taxonomy of Platforms for Conversational Agent Design

Stephan Diederich[1], Alfred Benedikt Brendel[1], Lutz M. Kolbe[1]

[1] University of Goettingen, Chair of Information Management, Goettingen, Germany
stephan.diederich@stud.uni-goettingen.de,
{abrendel, lkolbe}@uni-goettingen.de

Abstract. Software that interacts with its users through natural language, so-called conversational agents (CAs), is permeating our lives with improving capabilities driven by advances in machine learning and natural language processing. For organizations, CAs have the potential to innovate and automate a variety of tasks and processes, for example in customer service or marketing and sales, yet successful design remains a major challenge. Over the last few years, a variety of platforms that offer different approaches and functionality for designing CAs have emerged. In this paper, we analyze 51 CA platforms to develop a taxonomy and empirically identify archetypes of platforms by means of a cluster analysis. Based on our analysis, we propose an extended taxonomy with eleven dimensions and three archetypes that contribute to existing work on CA design and can guide practitioners in the design of CA for their organizations.

Keywords: Conversational agent, chatbot, design science, taxonomy, cluster analysis

1 Introduction

As artificial intelligence, particularly machine learning, increasingly permeates and impacts our daily private and professional lives, it drives a new wave of technological change and unprecedented automation of cognitive tasks [1]. One phenomenon in this wave are continuously improving conversational agents (CAs) which benefit from expanding functionalities and the diffusion of powerful and connected (mobile) devices. The presence of CAs is more and more increasing, such as in the form of Apple's Siri, Amazon's Alexa or in-car assistants. Basic CAs conduct information search for us, send messages or enter meetings in a calendar. Similarly, more and more companies use CAs for different purposes, such as automation and innovation in customer service or marketing and sales [2–7]. CAs can be distinguished from other software by their ability to interact with users based on natural language. This language can be spoken, as for example in the case of Amazon's Alexa, or written, often referred to as chatbots. In recent years, CA capabilities significantly expanded from simple rule-based systems to seemingly intelligent assistants [5, 8, 9] as a result of advances in machine learning and natural language processing.

14th International Conference on Wirtschaftsinformatik,
February 24-27, 2019, Siegen, Germany

In research, CAs attracted increasing interest in the last few years with different foci, such as information disclosure of users [10, 11], human performance improvement [12] or user authenticity perception [8]. In parallel with increased research interest in the IS community, organizations have started to experiment with and introduce CAs, often in the context of larger artificial intelligence initiatives [4, 8, 13]. However, many CAs fell behind expectations and often disappeared due to flaws related to their design, thus successful design remains a complex challenge in practice where various aspects need to be addressed [5, 14, 15].

With the popularity of CAs in both research and practice, a variety of enterprise CA platforms has emerged, supporting the design of CA with different functionality [16]. This includes both offerings of established technological players, such as Google's DialogFlow, as well as start-ups specialized in CAs such as ManyChat. While several studies can inform CA design through principles of form and function [5, 17, 18], the platforms that are used to actually designing CAs, providing both possibilities as well as constraints for the implementation, have not been studied in the past to the best of our knowledge. In order to gain a better understanding of these novel platforms, we first study along which dimensions CA platforms can be categorized (RQ1). Building on these dimensions and empirical data, we then aim to identify archetypes of platforms and their distinctive characteristics (RQ2). To address these research questions, we first develop a taxonomy of CA platforms, both conceptually from a literature review and empirically through the iterative classification of platforms. We then perform a cluster analysis to identify archetypes and gain a better understanding of commonalities and differences between the platforms.

We continue by describing the research background on CAs and presenting our research approach, i.e. taxonomy development followed by a cluster analysis. Finally, we present and discuss our results, particularly the developed taxonomy and identified archetypes, and close by suggesting directions for future work on CAs.

2 Research Background

The basic idea of a CA is to interact with users using natural language just like in a human-to-human conversation [19] and exchange information through verbal communication about a common topic [20]. This idea dates back decades to the 1960s when the first CA, called ELIZA, was developed by Joseph Weizenbaum [21]. Since then, a variety of CAs emerged (and often disappeared) that used simple pattern matching to provide a set of responses to the users [5, 22]. With recent technological advances, particularly in the fields of machine learning and natural language processing, as well as the diffusion of powerful, connected devices, the capabilities and potential of CAs increased significantly and they moved from rule-based systems to seemingly intelligent agents [22, 23]. Due to this development, CAs regained momentum in research and practice in the past few years and a variety of new CA offerings emerged.

In order to organize this variety that is available today, Gnewuch et al. [5] provide a simple taxonomy that consists of two dimensions including primary mode of communication and context (see Table 1). As natural language can be written or spoken [24], the mode of communication indicates the primary way in which users interact with a CA. For example, Apple's virtual assistant Siri is accessed using voice commands whereas Spotify's messenger bot works using digital text messages. CAs with text-based input are often referred to as chatbots in research as well as practice [2, 25, 26], while CAs with speech-based input are described as virtual or digital assistants [25, 27]. Because voice input can be quite easily transferred to written input in most cases, the boundaries between the mode of communication are often blurred as bots offer both spoken and written language as input. For example, a customer can request a ride with Lyft both via chat, e.g. Facebook Messenger or Slack, and by voice command, for example with Amazon Echo [28].

The second dimension, context, indicates whether the CA serves a specific domain such as a task or business function, or can interact on any topic with its users [5, 29]. General-purpose CAs like text-based Cleverbot [30] and Mitsuku [31] can have a conversation about any topic and continuously learn as they interact with users. For speech-based, general-purpose CAs the most prominent examples are from private life, such as Siri or Google Assistant.

Table 1. Classification of CA according to Gnewuch et al. [5]

		Context	
		General-purpose	**Domain-specific**
Communi-cation mode	**Text-based***	ELIZA, Cleverbot, Chatterbot, Mitsuku, …	Enterprise-class CAs, IKEA's Anna, Starbucks Chatbot, …
	Speech-based**	Apple's Siri, Amazon's Alexa, Google Assistant, …	SPECIES [29], in-car assistants, speech-based service agents, …

*Text-based: Chatbot, chatterbot, dialogue system, etc.
**Speech-based: (Virtual) personal assistant, digital companion, smart agent, etc.

Domain-specific CAs include a wide variety of CAs, for example in a professional context for internal and external purposes, such as customer service [4, 8], IT service desk tasks, product marketing [3], and e-commerce [14]. Further exemplary domains from private life include museums [32, 33] and healthcare [34].

In order to design a CA, a variety of development platforms exists to model a bot's behavior and to deploy them, for example on Facebook or by embedding the CA in the company website. Such platforms are characterized by an extensible technological foundation, i.e. the natural language processing and machine learning capabilities, created by a platform owner, on top of which developer can build platform-augmenting applications [35], such as conversational agents for a specific domain and organization.

The development platforms offer different ranges of functionality regarding aspects such as the bot's implementation, continuous training, analytics or hosting. With regard to the implementation for example, the platform Chatfuel [36] offers to quickly model a bot's behavior within a few minutes using a web interface while Twyla [37] uses supervised learning to automatically learn from existing data, such as customer service conversations or product catalogues. Concerning analytics, the functionality of CA platforms ranges from basic analysis (e.g. number and length of conversations) to advanced approaches, such as automatic sentiment and topic detection. Overall, a large number of enterprise platforms exists that allows building and introducing both text- and speech-based CA for general-purpose or specific domains.

3 Research Approach

In order to determine the distinct characteristics of CA platforms (RQ1) and to empirically identify archetypes (RQ2), we develop a taxonomy and perform a cluster analysis after classifying the respective platforms. The role of taxonomies is well recognized in information systems (IS) as they provide structure and organize knowledge in a field [38–41]. Within IS research, a multitude of taxonomies has been developed, covering for example business models of FinTechs [42], (mobile) health IT [43, 44] or cybercrime [45]. In particular in a diverse, emerging research area, taxonomies can provide useful insights into the grouping of objects based on their common characteristics [41].

To create our taxonomy, we follow the method proposed by Nickerson et al. [41] which iteratively develops a taxonomy based both on existing conceptual knowledge as well as empirical observation. This method clearly defines the necessary steps and ending conditions, providing a rigorous and useful approach for the systematic creation of a taxonomy, and to avoid the risk of defining and altering dimensions and characteristics through ad-hoc changes. The Nickerson method has been successfully applied to develop a variety of taxonomies, such as for collaborative applications [46] or carsharing business models [47]. Our complete research approach consists of three phases and is summarized in Table 2.

Table 2. Research approach phases

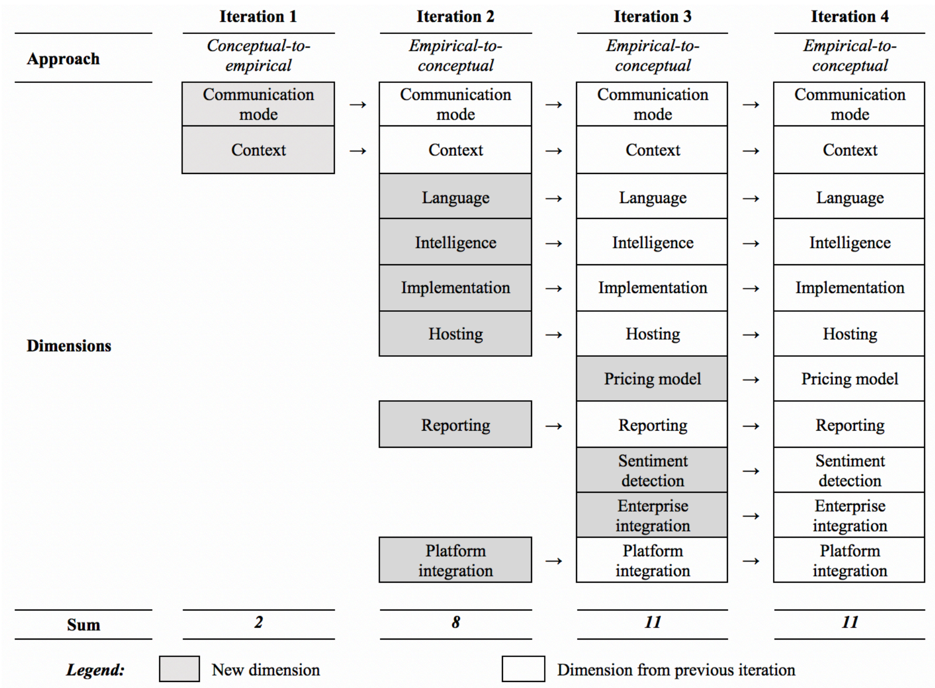
	Phase 1: Create database	**Phase 2: Develop taxonomy**	**Phase 3: Conduct cluster analysis**
Steps	• Search for CA platforms in CrunchBase and on the web • Request additional information where required	• Define meta-characteristic for the taxonomy • Iterate through taxonomy development until ending conditions are met	• Determine useful number of clusters • Specify the companies belonging to each cluster
Method	Lit. review, desk research	Taxonomy development	Clustering algorithms
Source	CA lit., blogs, practice reports, CrunchBase	CA literature, CA platform database	Taxonomy of CA platforms with empirical data
Results	Database with 51 CA platforms	Taxonomy of CA platforms with 11 dimensions	Three identified CA platform archetypes

Phase 1: Set up database: The first research phase aimed at the creation of a database with CA platforms that were operational in May 2018. For this we examined existing literature on CA, searched the world's largest startup database (CrunchBase), a variety of blogs (e.g. https://chatbotsjournal.com), and industry reports (e.g. Oracle [13]). For our search, we used the terms "conversational agent" with the synonyms "chatbot" and "digital assistant" in combination with "design" and "platform". Platforms that were not operational (i.e. actively providing the option to create a CA) were excluded from the database. Missing or incomplete data, particularly on pricing models, was gathered via e-mail requests. At the end of the first research phase, we created a database with 51 platforms for CA design.

Phase 2: Develop taxonomy: The objective of the second phase was to create a taxonomy of CA platforms that contains the most important dimensions along which the platforms differ based on the method described by Nickerson et al. [41]. For our research, we defined CA development platforms as the meta-characteristic for the taxonomy from which all subsequent dimensions follow. Regarding the ending conditions that indicate whether the taxonomy development process is completed, we used the eight objective (such as all objects have been examined and no new dimension or characteristics were added in the last iteration) and five subjective ending conditions (concise, robust, comprehensive, extendible and explanatory) from Nickerson et al. [41]. We started the taxonomy development with a conceptual-to-empirical iteration. In this initial iteration we added two dimensions (CA primary mode of communication, CA context [5, 29]) that were identified in our literature review (see Table 1). The following three iterations were empirical-to-conceptual and added nine dimensions in total, such as pricing model, implementation mode or hosting (see Figure 1). After all platforms in our database were successfully classified and both subjective and objective ending conditions were met, we considered the taxonomy final.

Phase 3: Perform cluster analysis: The objective of the third research phase was the empirical identification of CA platform archetypes (RQ2). For this purpose, we conducted a cluster analysis. Cluster analysis aims at grouping objects where objects in one group are as similar as possible and as dissimilar as possible from objects in other groups [48]. Following the recommendations by Punj and Stewart [49] to first determine the number of clusters and subsequently use an iterative partitioning technique like k-means, we chose a two-stage clustering approach: First, we defined the number of clusters with Ward's method. With this method we agglomeratively clustered (i.e. repeatedly combined the two closest objects into one group until all objects belong to the same group [50]) the CA platforms using SPSS version 25 and squared Euclidean distance. We then reviewed the descriptive data on these iterations, i.e. the coefficient distance, the dendrogram and the scree plot using the elbow rule. These indicated that a three cluster would be most useful. In the second step, we used the chosen number of groups for a k-means clustering procedure. The procedure used three iterations until no significant enhancements were achieved.

Figure 1. Iterations of our taxonomy development



4 Results

In the following, we present our taxonomy for CA platforms (RQ1) and provide examples for platforms to demonstrate their respective characteristics. We then continue with describing the archetypes of platforms we identified in the two-step cluster analysis (RQ2).

4.1 Taxonomy for CA platforms

The resulting taxonomy consists of 11 dimensions with two to four characteristics each (see Table 3). The first two dimensions were found in existing literature [5]. Each platform was assigned one characteristic for each dimension. We omitted dimensions that were the same across all platforms (representation of the CA with an avatar, assigning a name to the CA) as we aim to distinguish them by their main characteristics.

The first dimension, *Communication mode*, refers to the primary way with which a user communicates with a CA and may more broadly described as the user interface, i.e. text-based, speech-based or both [5]. For example, platforms such as ManyChat, pandorabots, or Recime exclusively offer building text-based CAs (referred to as Chatbots) whereas aivo and The Pullstring Platform focus on agents that interact with its users via speech. Furthermore, platforms such as Nuance and IPSoft offer to build and integrate CA that interact via both text and speech. The dimension *context* indicates in which task or business domain a CA built on the respective platform can be used [5]. For example, SurveyBot offers to build specific CAs that conduct surveys and collect their results or Octane AI's CA that provides sales optimization by interactively engaging with users that abandon their digital shopping carts. The dimension *language* refers to the language(s) supported by the CA where platforms offer support for single languages (mostly English, e.g. botmother) or multiple languages (e.g. ChatClub). *Intelligence* indicates whether a CA is primarily based on rules that perform rather simple pattern matching, such as ChatbotsBuilder, or has the ability to self-learn, such as Twyla, enabling the CA to improve over time as it converses with its users.

Table 3. Taxonomy of CA platforms

Dimension	**Characteristics**			
Communication mode	Text-based	Speech-based	Both	
Context	General-purpose	Domain-specific		
Language	Single language	Multi language		
Intelligence	Rule-based	Self-learning		
Implementation	Programming	Modeling	Supervised learn.	Hybrid
Hosting	On-premise	Cloud	Both	
Pricing model	Usage-based	User-based	Instance-based	Free
Reporting	Without reporting	With reporting		
Sentiment detection	Without sentiment	With sentiment		
Enterprise integration	None	API	Pre-build interface(s)	
Platform integration	Single-platform	Cross-platform		

Existing dimension New dimension

The dimension *Implementation* indicates how a bot is built, whether via programming (actually writing code), modeling (modeling typical user conversations in a flow chart), supervised learning (training the CA with labeled conversations), or with the help of a hybrid approach (e.g. modeling in combination with supervised learning). Popular platforms for creating a bot via programming are wit.ai, and Zenbot. With regard to modeling, the most common platforms used to build bots include Massively, ManyChat, and LeadFlip. In contrast to programming and modeling, some platforms such as Twyla rely on training a CA with existing user interactions (supervised learning) while others like Creative Virtual and gupshup use a combination of these implementation approaches. *Hosting* refers to the deployment of CAs where platform offerings range from on-premise (e.g. botpress), public cloud (e.g. ChatterOn or Converse), and both methods combined. *Pricing* refers to the pricing model that is used by the platform. The models we observe in our data include usage-based (i.e. based on number of interactions, such as Microsoft Azure Bot), user-based (i.e. based on number users, such as MobileMonkey), instance-based (i.e. based on number of CA, such as ChatbotsBuilder) and free (such as It's Alive).

Reporting indicates whether a CA platform offers reporting functionality to monitor the CA's interactions and usage, such as number of conversations or unique users (for example provided by reply.ai and Lex). *Sentiment detection* indicates whether a platform allows automatic detection of user sentiment during an interaction. Finally, *Enterprise integration* indicates whether a CA platform offers pre-built interfaces or APIs to let CAs access different enterprise systems such as a CRM for information that is used in a conversation with a user. For example, Microsoft Azure Bot Service can automatically retrieve information from its Dynamics CRM in a user interaction via a standardized interface. Other platforms, for example pandorabots or Rasa, can retrieve data from enterprise systems via API calls.

4.2 Archetypes of CA platforms

The three clusters contain 18 (cluster 1), 19 (cluster 2), and 14 (cluster 3) platforms from our database (Table 4). Each cluster has different centers along the dimensions of the taxonomy developed in this study. As the characteristics within the taxonomy are mutually exclusive and collectively exhaustive, we describe the clusters with a crosstab analysis showing percentages for each characteristic within a cluster (see Figure 2). For example, 22% of all CA platforms in cluster 1 support a single language whereas 78% offer multi language support. In the following, we describe the clusters, highlight their distinctive characteristics, and provide illustrative examples.

Archetype 1 – Multi-language, integrative CA platform with advanced analytical functionality: The first cluster contains platforms that mainly support multiple languages, self-learn over time, and integrate with different enterprise systems, such as CRM software, as well as various platforms, such as social media. All platforms within this cluster offer reporting functionalities and the majority of platforms has built-in sentiment detection. These platforms include the CA offerings of major technology players, such as Oracle Intelligent Bots, Microsoft Azure Bot Service, IBM Watson Assistant or Amazon Lex, and large technology companies that strive to automate tasks particularly in customer service, IT operations as well as product and marketing like IPSoft or Nuance. CA platforms in this cluster support text-based and speech-based communication and include CAs for various purposes. Whereas platforms in cluster 2 and 3 mainly focus on the modeling of typical conversation flows as an implementation approach, platforms in this clusters also offer supervised learning (allowing to train a CA with a set of historical, labelled data) and hybrid approaches (i.e. a combination of modeling and supervised learning). Regarding deployment, many platforms offer cloud or cloud and on-premise hosting and pricing depends on actual usage.

Archetype 2 – General-purpose, cloud-based CA platform with single language and API support: The second cluster includes platforms that focus on CAs for different purposes, support a single language (in most cases English), and are primarily hosted in the cloud. With regard to integration with other enterprise software, these platforms typically offer APIs to program the automatic retrieval of data from existing systems, such as CRM. Examples of platforms in this cluster include pandorabots, Recime and Xenioo. These platforms mostly use modeling as the implementation approach, as in the first cluster. Regarding the analytical functionality, none of the platforms provide sentiment detection, while about two third of the platforms in this cluster offer reporting features. Regarding the integration of CAs with target platforms, the companies within this cluster are split between single-platform (e.g. TalkBot for Facebook) and cross-platform support (e.g. pandorabots).

Figure 2. Cross tab analysis

Dimension	Characteristics	Archetype 1	Archetype 2	Archetype 3
	Number of platforms in cluster	*18*	*19*	*14*
Communication mode	Text-based	33%	68%	100%
	Speech-based	33%	5%	0%
	Both	33%	26%	0%
Context	General-purpose	83%	100%	21%
	Domain-specific	17%	0%	79%
Language	Single language	22%	89%	93%
	Multi language	78%	11%	7%
Intelligence	Rule-based	0%	0%	50%
	Self-learning	100%	100%	50%
Implementation	Programming	6%	11%	0%
	Modeling	50%	84%	100%
	Supervised learning	33%	5%	0%
	Hybrid	11%	0%	0%
Hosting	On-premise	11%	0%	7%
	Cloud	39%	89%	93%
	Both hosting	50%	11%	0%
Pricing model	Usage-based	89%	79%	50%
	User-based	6%	5%	14%
	Instance-based	0%	5%	14%
	Free	6%	11%	21%
Reporting	Without reporting	0%	26%	50%
	With reporting	100%	74%	50%
Sentiment detection	Without sentiment	39%	100%	93%
	With sentiment	61%	0%	7%
Enterprise integration	None	0%	0%	71%
	API	11%	95%	29%
	Pre-build interface(s)	89%	5%	0%
Platform integration	Single-platform	0%	47%	79%
	Cross-platform	100%	53%	21%

Archetype 3 – Text-based, domain-specific CA platform with modeling functionality: The third and final cluster contains platforms that show different distinctive characteristics: First, these platforms exclusively offer text-based CAs, which tend to be chatbots that are used in specific domains and mostly on single platforms. For example, SurveyBot can conduct interactive surveys and collect their results via Facebook Messenger. CA platforms in this cluster typically host their CA in their own clouds and pricing is based on actual usage. With regard to the capability for integration of data from other enterprise software, the majority of platforms in this cluster does not offer an API or pre-built interfaces connecting the CA to existing systems.

Table 4: Clustered platforms

Archetype 1	Archetype 2	Archetype 3
[24]7 AI	BotEngine	ChatbotsBuilder
aivo	botmother	ChatClub
BotCore	Botsify	ChatterOn
botpress	Chatfuel	E.D.D.I.
Creative Virtual	Conversation one	HubSpot / motion.ai
Dialogflow	Converse	It's alive
gupshup	Flow xo	LeadFlip
IBM Watson Assistant	Landbot.io	ManyChat
inbenta	pandorabots	Massively
Interactions	Parlo	MobileMonkey
IPSoft	Rasa	Octane AI
Lex	Recime	rebot.me
Microsoft Azure Bot Service	Sequel	Surveybot
Next IT	Smooch	Zelp
Nuance	TalkBot	
Oracle Intelligent Bots	The PullString Platform	
reply.ai	Wit.ai	
Twyla	Xenioo	
	Zenbot	

5 Discussion

In the following, we discuss the developed taxonomy and identified archetypes against the background of existing research, followed by a description of limitations of this study, and an overview of opportunities for future research.

5.1 Taxonomy and Archetypes

The taxonomy and archetypes from our analysis underline the versatility of CA platforms and indicate three types of platforms. The cross-cluster comparison shows that CA platforms range from high-end offerings (cluster 1), mainly by large technology providers such as IBM or Microsoft that offer a variety of analytical features and options for integration as well as provide CAs both for speech- and for text-based communication, over mid-range general-purpose CA platforms (cluster 2) like pandorabots or Chatfuel that primarily focus on single platforms for deployment and require implementing an API for integration to highly standardized CA platforms (cluster 3) that offer mainly domain-specific CA with a limited set of functionality, such as SurveyBot or MobileMonkey. These archetypes and the underlying taxonomy contribute to theory in different ways. The taxonomy we developed extends the existing, basic classification of CAs according to communication mode and context [5] through the empirical observation of CA design platforms by adding further dimensions. These dimensions describe CAs in greater detail as the existing classification, for example by taking into account the implementation approach, integration capabilities or the intelligence a CA possesses, which in turn provides possibilities and constraints for implementing CAs based on design principles formulated in previous studies [5]. Furthermore, we provide an overview of the state-of-the-art of platforms for conversational agent design through the taxonomy and classified platforms that can be used in future design-oriented research on CAs. For example, studies that investigate empathetic behavior of chatbots in customer service, such as the work by Hu et al. [51], could select a platform that offers built-in sentiment analysis for text-based CA to design their CA. Thus, in the context of design-oriented research, this study contributes to the growing knowledge base on CA [52].

In addition to the aforementioned contributions, our study provides two main insights for practitioners that intend to design CAs. First, the taxonomy can be used to select a vendor for a specific use case, for example by defining the desired characteristics along the 11 dimensions and then choosing a suitable platform. For example, a company that seeks to design a text-based CA with multi language support, on-premise hosting, and built-in analytics functionality could select a platform such as inbenta, Creative Virtual or IBM Watson Assistant. Or, a company that would like their CA to specifically conduct text-based customer surveys on a single platform, Facebook, can use SurveyMonkey for their implementation. The cross-cluster comparison shows that CA platforms range from high-end offerings (archetype 1), mainly by large technology providers such as IBM or Microsoft that offer a variety of analytical features and options for integration as well as provide CAs both for speech- and for text-based communication, over mid-range general-purpose CA platforms (archetype 2) like pandorabots or Chatfuel that primarily focus on single platforms for deployment and require implementing an API for integration to highly standardized CA platforms (archetype 3) that offer mainly domain-specific CAs with a limited set of functionality, such as SurveyBot or MobileMonkey.

Second, the platform database and identified archetypes underline the wide spectrum of CA platforms ranging from basic text-based CAs for single platforms to high-end, adaptive CAs that integrate in existing systems and can communicate with customers both via speech and text. Thus, managers can use the archetypes to strategically decide what type of CA platform they require. Furthermore, the analysis revealed that some platforms address different departments. Whereas multiple platforms can directly be used for design by the department that intends to introduce a CA, such as marketing and sales, as they deliver it based on simple modeling of typical conversation flows and convenient hosting in the cloud, other platforms address and require the IT department to customize, integrate and deploy their solutions.

5.2 Limitations and Opportunities for Future Research

Our study is not free of limitations and offers opportunities for future studies. First, the taxonomy that was developed both from existing CA literature and empirical data (i.e. the platforms in our database) cannot be considered comprehensive in terms of explaining platforms in detail but is helpful for understanding and delineating CA platforms as shown our analysis. As Nickerson et al. [41] highlight a taxonomy can never be perfect but is at best useful to explain the nature of objects under study. We initially demonstrated the usefulness of our taxonomy, but it can benefit from validation and expansion in future studies. A second limitation is that some dimensions might mutually exclude one another. We did not systematically identify these inter-dependencies in our work, yet it would be useful to address this point in the future. The third limitation results from the market dynamics that exist with regard to CA platforms. Present acquisitions, such as Motion.AI acquired by HubSpot, underline that the current CA platform landscape is subject to change which in turn limits the validity of our analysis over time. Similarly, CA platforms might add different functionality over time and provide new interfaces to enterprise software which would reduce the accuracy of our database. However, as the cluster analysis indicated a rather equal distribution of platforms to cluster, we would argue that the three typical CA platforms will still remain applicable even in the light of acquisitions and feature changes.

We suggest two main opportunities for future research: First, the taxonomy created in this paper can be evaluated in the field with organizations that plan to introduce CA for innovation or automation. Incorporating the views from organizations that seek to introduce a CA can be useful to validate and potentially extend the dimensions or characteristics in the taxonomy. Second, engaging with organizations introducing CA can also be helpful to reach a better understanding regarding the reasons for or against selecting specific archetypes as well as with regard to different characteristics. For example, comparing the two implementation approaches modeling of conversation flows with training of a CA based on existing and labeled data (supervised learning) concerning the impact on CA performance is a promising research endeavor not only in the context of CA, but also within the broader spectrum of innovative approaches for task or process automation.

6 Conclusion

In this study, we set out to develop a taxonomy of CA platforms (RQ1) and identify their archetypes (RQ2) in order to better understand the variety of platforms to design natural language agents for organizations. Based on existing CA literature as well as the analysis of 51 platforms, we derived a taxonomy with 11 dimensions which describes CA platform characteristics alongside their implementation and hosting approaches, pricing models, analytical features, and options for enterprise software integration. Afterwards, we empirically identified three archetypes of CA platforms with different ranges of functionality. Our work contributes an overview of the state-of-the-art of platforms for CA design and outlines possibilities and constraints for the implementation of design knowledge on conversational agents. In addition, our results can practically guide CA platform selection through the analysis of platforms based on the taxonomy and outlining aspects to be considered in the design process, such as the need for multi-language support or built-in sentiment analysis.

References

1. Brynjolfsson, E., McAfee, A.: The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. Norton & Company, New York, USA (2016).
2. Chakrabarti, C., Luger, G.F.: Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics. Expert Syst. Appl. 42, 6878–6897 (2015).
3. Meuter, M.L., Ostrom, A.L., Roundtree, R.I., Bitner, M.J.: Self-Service Technologies:Understanding Customer Satisfaction with Technology-Based Service Encounters. J. Mark. 64, 50–64 (2000).
4. Verhagen, T., van Nes, J., Feldberg, F., van Dolen, W.: Virtual customer service agents: Using social presence and personalization to shape online service encounters. J. Comput. Commun. 19, 529–545 (2014).
5. Gnewuch, U., Morana, S., Maedche, A.: Towards Designing Cooperative and Social Conversational Agents for Customer Service. In: Proceedings of the International Conference on Information Systems (ICIS). pp. 1–13. , Seoul, Korea (2017).
6. Jacobs, I., Powers, S., Seguin, B., Lynch, D.: The Top 10 Chatbots For Enterprise Customer Service. Forrester Res. (2017).
7. 80% of businesses want chatbots by 2020, https://www.businessinsider.de/80-of-businesses-want-chatbots-by-2020-2016-12?r=US&IR=T.
8. Wünderlich, N. V., Paluch, S.: A Nice and Friendly Chat With a Bot: User Perceptions of AI-based Service Agents. In: Proceedings of the International Conference on Information Systems (ICIS). pp. 1–11. , Seoul, Korea (2017).
9. Morana, S., Friemel, C., Gnewuch, U., Maedche, A., Pfeiffer, J.: Interaktion mit smarten Systemen – Aktueller Stand und zukünftige Entwicklungen im Bereich der Nutzerassistenz. Wirtschaftsinformatik Manag. 5, 42–51 (2017).
10. Saffarizadeh, K., Boodraj, M., Alashoor, T.M.: Conversational Assistants: Investigating Privacy Concerns, Trust, and Self-Disclosure. In: Proceedings of the International Conference on Information Systems (ICIS). pp. 0–12. , Seoul, Korea (2017).
11. Schroeder, J., Schroeder, M.: Trusting in Machines: How Mode of Interaction Affects Willingness to Share Personal Information with Machines. In: Proceedings of the Hawaii International Conference on System Sciences (HICSS). , Waikoloa , Hawaii, USA (2018).
12. Abul, M., Siddike, K., Spohrer, J., Demirkan, H., Kohda, Y.: People's Interactions with Cognitive Assistants for Enhanced Performances. In: Proceedings of the Hawaii International Conference on System Sciences (HICSS). pp. 1640–1648. , Waikoloa Village, Hawaii, USA (2018).
13. Oracle: Can Virtual Experiences Replace Reality? The future role for humans in delivering customer experience, (2016).
14. Ben Mimoun, M.S., Poncin, I., Garnier, M.: Animated conversational agents and e-consumer productivity: The roles of agents and individual characteristics. Inf. Manag. 54, 545–559 (2017).
15. Ben Mimoun, M.S., Poncin, I., Garnier, M.: Case study-Embodied virtual agents: An analysis on reasons for failure. J. Retail. Consum. Serv. 19, 605–612 (2012).
16. botnerds, http://botnerds.com/types-of-bots/.
17. Gnewuch, U., Morana, S., Adam, M.T.P., Maedche, A.: Faster Is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. In: Proceedings of the European Conference on Information Systems (ECIS). , Portsmouth, United Kingdom (2018).

18. Schuetzler, R.M., Grimes, G.M., Giboney, J.S.: An Investigation of Conversational Agent Relevance, Presence, and Engagement. In: Proceedings of the Americas Conference on Information Systems (AMCIS). pp. 1–10. , New Orleans, USA (2018).
19. McTear, M., Callejas, Z., Griol, D.: The Conversational Interface: Talking to Smart Devices. Springer Publishing Company, Basel, Switzerland (2016).
20. Berg, M.M.: Modelling of Natural Dialogues in the Context of Speech-based Information and Control Systems, http://macau.uni-kiel.de/receive/dissertation_diss_00016739, (2014).
21. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM. 9, 36–45 (1966).
22. Knijnenburg, B.P., Willemsen, M.C.: Inferring Capabilities of Intelligent Agents from Their External Traits. ACM Trans. Interact. Intell. Syst. 6, 1–25 (2016).
23. Berg, M.: NADIA: A Simplified Approach Towards the Development of Natural Dialogue Systems. In: Natural Language Processing and Information Systems. pp. 144–150. Springer International Publishing (2015).
24. Lee, C., Jung, S., Kim, S., Lee, G.G.: Example-based dialog modeling for practical multi-domain dialog system. Speech Commun. 51, 466–484 (2009).
25. Shawar, B.E., Atwell, E.: Chatbots: Are they really useful? In: LDV-Forum. pp. 29–49 (2007).
26. Hill, J., Randolph Ford, W., Farreras, I.G.: Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. Comput. Human Behav. 49, 245–250 (2015).
27. Sarikaya, R.: The technology behind personal digital assistants: An overview of the system architecture and key components. IEEE Signal Process. Mag. 34, 67–81 (2017).
28. Medium: 10 Real Examples How Brands are Using Chatbot for Customer Service, https://medium.com/the-mission/10-real-examples-how-brands-are-using-chatbot-for-customer-service-4fbb5e4617f3.
29. Nunamaker, J.F., Derrick, D.C., Elkins, A.C., Burgoon, J.K., Patton, M.W.: Embodied Conversational Agent-Based Kiosk for Automated Interviewing. J. Manag. Inf. Syst. 28, 17–48 (2011).
30. Cleverbot, https://www.cleverbot.com.
31. Mitsuku, http://www.mitsuku.com.
32. Vassos, S., Malliaraki, E., Dal Falco, F., Di Maggio, J.: Art-Bots: Toward Chat-Based Conversational Experiences in Museums. In: Proceedings of the 9th International Conference on Interactive Digital Storytelling. pp. 433–437. , Los Angeles, USA (2016).
33. Kopp, S., Gesellensetter, L., Krämer, N.C., Wachsmuth, I.: A conversational agent as museum guide - Design and evaluation of a real-world application. In: Proceedings of Intelligent Virtual Agents: 5th International Working Conference. pp. 329–343. , Kos, Greece (2005).
34. Miner, A., Chow, A., Adler, S., Zaitsev, I., Tero, P., Darcy, A., Paepcke, A.: Conversational Agents and Mental Health: Theory-Informed Assessment of Language and Affect. In: Proceedings of the International Conference on Human Agent Interaction, Singapore (2016).
35. Tiwana, A.: Platform Desertion by App Developers. J. Manag. Inf. Syst. 32, 40–77 (2015).
36. Chatfuel, https://chatfuel.com.
37. Twyla, https://www.twylahelps.com.
38. Glass, R.L., Vessey, I.: Contemporary Application-Domain Taxonomies. IEEE Softw. 12, 63–76 (1995).
39. McKnight, D.H., Chervany, N.L.: What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. Int. J. Electron. Commer. 6, 35–59 (2001).

40. Williams, K., Chatterjee, S., Rossi, M.: Design of emerging digital services: A taxonomy. Eur. J. Inf. Syst. 17, 505–517 (2008).
41. Nickerson, R.C., Varshney, U., Muntermann, J.: A method for taxonomy development and its application in information systems. Eur. J. Inf. Syst. 22, 336–359 (2013).
42. Eickhoff, M., Muntermann, J., Weinrich, T.: What do FinTechs actually do ? A Taxonomy of FinTech Business Models. In: Proceedings of the International Conference on Information Systems (ICIS). pp. 1–19. , Seoul, Korea (2017).
43. Varsheny, U., Nickerson, R.C., Muntermann, J.: Taxonomy Development in Health-IT. In: Proceedings of the Americas Conference on Information Systems (AMCIS). pp. 1–10. , Chicago, USA (2013).
44. Yang, A., Varshney, U.: A Taxonomy for Mobile Health Implementation and Evaluation. In: Proceedings of the International Conference on Information Systems (ICIS). pp. 1–10. , Dublin, Ireland (2016).
45. Land, L., Smith, S., Pang, V.: Building a Taxonomy for Cybercrimes. In: Proceedings of the Pacific Asia Conference on Information Systems (PACIS). pp. 34–38. , Jeju, Korea (2013).
46. Nickerson, R.C.: A Taxonomy of Collaborative Applications. In: Proceedings of the Americas Conference on Information Systems (AMCIS). pp. 560–562. , Indianapolis, USA (1997).
47. Remane, G., Hanelt, A., Nickerson, R.C., Tesch, J.F., Kolbe, L.M.: A Taxonomy of Carsharing Business Models. Proc. Int. Conf. Inf. Syst. 1–19 (2016).
48. Kaufman, L., Rousseeuw, P.J.: Introduction. In: Finding Groups in Data: An Introduction to Cluster Analysis. pp. 1–67 (1990).
49. Punj, G., Stewart, D.W.: Cluster Analysis in Marketing Research: Review and Suggestions for Application. J. Mark. Res. 20, 134–148 (1983).
50. Landau, S., Everitt, B.: A handbook of statistical analyses using SPSS. (2004).
51. Hu, T., Xu, A., Liu, Z., You, Q., Guo, Y., Sinha, V., Luo, J., Akkiraju, R.: Touch Your Heart: A Tone-aware Chatbot for Customer Care on Social Media. In: Proceedings of the ACM CHI Conference on Human Factors in Computing Systems. pp. 1–12. , Montréal, Canada (2018).
52. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design Science in Information Systems Research. Manag. Inf. Syst. Q. 28, 75–105 (2004).