

User-Chatbot Interaction: an Acceptancy Evaluation

Stefano Valtolina¹, Ricardo A. Matamoros Aragon² and Riccardo Castelli¹

¹ *Università degli Studi di Milano, via Celoria 18, Milano, Italy*

² *Milan Bicocca University, Piazza dell'Ateneo Nuovo, Milano, Italy*

Abstract

In a growing number of contexts of use, chatbots are becoming an increasingly adopted interaction strategy. This paper explores the idea of using a chatbot in two key areas of interest: Customer service and Personal assistants. Specifically, we are interested in researching the acceptance level of this technology with regard to a user's set of behaviours, which defines the interaction process between users and chatbots. The main aim is to present a model to evaluate the most important factors that impact the acceptance of chatbots by users. To do it we devised an extension of the UTAUT model to assess the quality of the chatbot's communication in terms of completeness, clarity and transparency and to evaluate how much the user can trust a chatbot regarding privacy issues. The results demonstrate good effects on what concerns the acceptance of our chatbots by users.

Keywords

User-Chatbot interaction, Conversational interfaces, UTAUT Model

1. Introduction

Different application fields in recent years have adopted several types of algorithms based on Artificial Intelligence (AI) models to facilitate the actions a user performs day-to-day. This set of algorithms is used to develop Intelligent Assistants (AIs).

A chatbot is a system that uses a subset of AIs algorithms that communicates with users through a mix of vocal, text and visual messages using predefined rules [1]. Typically, a chatbot is supported by AI, which can interact with users in a conversational manner using natural language processing. The conversation aims at simulating a dialogue with users to provide information, give assistance or support them in daily tasks such as setting alarms, making calls, shopping, etc. Currently, chatbots are successfully penetrating various industries, such as finance, retail, medical, tourism, government departments, education, etc. Moreover, chatbots' use is growing in several contexts of use thanks to the commercial success of intelligent personal assistant devices such as Amazon Alexa and Google Assistant. This solution is not a new technology itself, but the rise of reliable linguistic functionality and the addition of intelligence through machine learning has increased its popularity [2].

In [3], the authors investigate the potential usefulness of devising a classification to analyse chatbot interaction design. They propose a set of typology dimensions based on four kinds of interactional styles discussed by IBM's research group² on conversational UX design. This classification aims at distinguishing chatbots according to the intended duration of the relationship with users (short vs long term) and the locus of control for the dialogue (user-driven vs chatbot-driven interaction). The proposed typology leads to the identification of four chatbot purposes: Customer service, Personal assistants, Content curation, and Coaching.

In this paper, the main research question is to investigate the influence of AI-based chatbot suggestions in the minimization of effort in the activities that humans perform in their daily lives and how this

Proceedings of CoPDA2022 - Sixth International Workshop on Cultures of Participation in the Digital Age: AI for Humans or Humans for AI? June 7, 2022, Frascati (RM), Italy

EMAIL: stefano.valtolina@unimi.it (S. Valtolina) r.matamorosaragon@campus.unimib.it (R. A. Matamoros Aragon);

riccardo.castelli3@studenti.unimi.it (R. Castelli)

ORCID: 0000-0003-1949-2992 (S. Valtolina); 0000-0002-1957-2530 (R. A. Matamoros Aragon)



© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

² <http://conversational-ux.mybluemix.net/design/conversational-ux/>

influence is related to factors that moderate human behaviour in adopting AI systems. To this aim, we specifically studied chatbots based on dialogues that are controlled by the user who uses the chatbot as an assistant able to provide her/his with useful suggestions. From a Human-Centred Design (HCD) perspective, we investigated how the use of these chatbots affects the users' empowerment in managing their activities.

According to this objective, in Section 2, we describe the prototypes of chatbots we developed to summarise the main experiences that a user typically faces when interacting with them. These chatbots have been developed with the purpose to reply to a main research question about the experience that users have when they are involved in a dialogue with a virtual agent. This question arises to understand the issues and quality attributes associated with designing chatbots and identify a methodology for predicting future intention to use them by users. Therefore, identifying the factors that influence the intention to use the chatbot is aimed at understanding the human behaviour that determines the acceptance of AI-based systems.

In Section 3, we present a framework to understand users' acceptance behaviour of a chatbot on a conceptual level extending a basic model known as the Unified Theory of Acceptance and Use of Technology (UTAUT) [4]. This extension aims at incorporating new constructors specifically designed to study the quality of the chatbot's communication in terms of completeness, clarity and transparency and to evaluate how much the user can trust a chatbot regarding privacy issues. Finally, in the last Section, we conclude with the implications and future research for this paper.

2. Chatbots Classification

The proposed typology in [3] has been used as the basis for identifying high-level approaches that can lead designers to design effective conversations between users and our chatbot prototypes. This classification aims at identifying four key areas of interest in which the chatbots are used: Customer service, Personal assistants, Content curation, and Coaching.

In detail, this work aims at distinguishing two types of interactions: chatbot-driven dialogue and user-driven dialogue. In our paper, we are interested in studying this last type of chatbot. A chatbot that can support users in their activities by providing them with suggestions to enable them in dealing with work and daily activities without taking control of the dialogue. The design of this conversation is more challenging, both technologically and in terms of the needed breadth and volume of content. The chatbot will need to identify the users' intent, both on the level of the individual messages and overall for the interaction, and also to be able to respond adequately to these intents. As a consequence, some user-driven chatbots can provide interaction sequences that are typically relatively brief or others may enable longer conversational sequences. According to this last difference, we developed two prototypes:

- **Chatbot for Customer Service - Carlo:** As depicted in Figure 1 we developed a chatbot named Carlo that has been specifically equipped with all features for processing personal data and for taking care of the user's privacy. It can obscure users' personal data in the stored log file, request consent to the processing of data, answer users' doubts in terms of the privacy policy and GDPR issues and allow users to delete the conversation she/he had with the chatbot. Carlo has been designed to understand the factors driving the users' intention to use a chatbot for customer service and its ability to help users to tailor banking services that best meet their needs.
- **A Chatbot as Personal Assistant in Education - *WhoTeach-PA*:** With the chatbot *WhoTeach-PA* we developed the idea of using a personal learning assistant as an expert able to advise teachers about the e-learning objects to take into account when they have to create new online courses (Figure 2). The chatbot has been designed to investigate the level of acceptance influencing the intention to use it by teachers. In particular, because teachers may not trust the suggestions provided by a chatbot when they ignore the underlying architecture and AI-based algorithms, we studied how a trustworthy assistant can affect the intention to use it. In particular, we investigated how much this acceptability is influenced by recommendations that are provided to the user in a clear, transparent and understandable way.

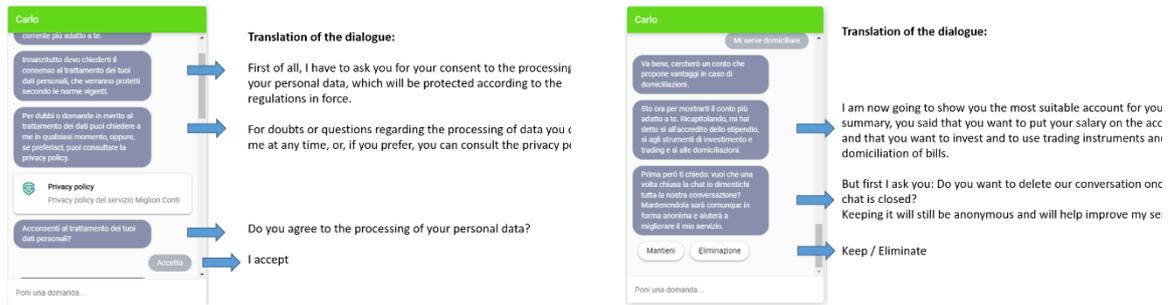


Figure 1: A portion of the chatbot dialogue. On the left, Carlo is requesting the user's consent to process her/his personal data. On the right, it is summarising what the user wants to select the best-suited bank account services. In the end, Carlo asks the user if she/he wants to delete the dialogue. If the dialogue will be kept, it will be anonymized anyway.

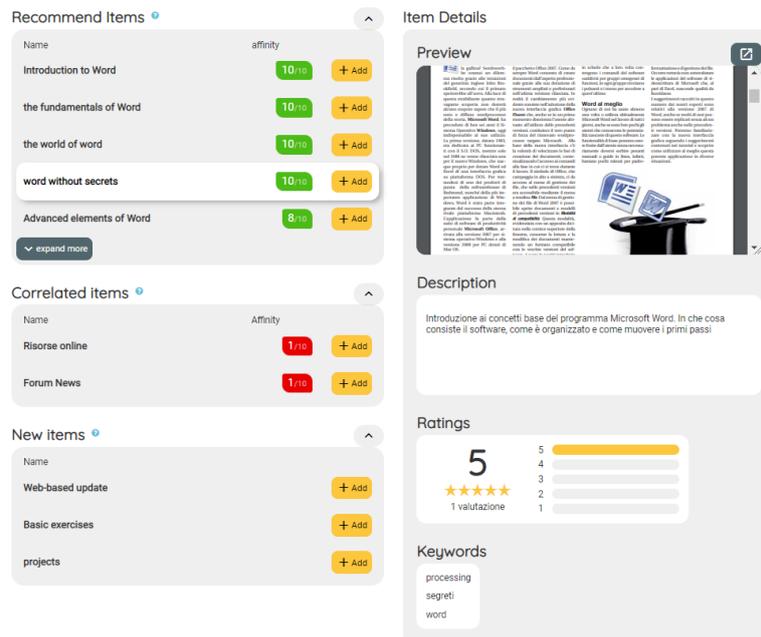


Figure 2: Screenshots of a WhoTeach-PA dialogue during which it is suggesting possible learning objects to use for composing a new course. The suggestions are provided by using different interactive visualisation strategies.

3. Evaluation of the User's Intention to Use a Chatbot

For analysing the user's intention to use a given technology or service, researchers over these decades have proposed a plethora of competing theories and models [4,5]. As shown in a recent literature review TAM has evolved to become a key model in understanding predictors of human behaviour toward potential acceptance or rejection of the technology. The Unified Theory of Acceptance and Use of Technology (UTAUT) is the latest derivative of TAM [5,6]. Since being introduced, the UTAUT model has been tested extensively in various fields and promises to be a great tool for analysing users' acceptance of specific technology [7, 8].

For this reason, to evaluate the user's intention to use a chatbot, in this paper we propose an extension of the UTAUT model. The final measured construct in the UTAUT model is Behavioural Intention to Use (BI) defined as "a measure of the strength of one's intention to perform a specified behaviour" [9]. It is influenced by four main constructors: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), and Facilitating Conditions (FC). To study chatbots we need to extend the UTAUT model emphasising motivations that can affect the intention to use such technology

incorporating new constructors such as communicability/quality of service, perceived trust/privacy management, and perceived transparency.

Digital assistants are software tools that are communicative by their nature since users achieve their goals by exchanging messages with the system [10]. We specifically studied this factor by analysing the interaction between the user and the chatbot Carlo. Carlo acts in the customer service field and so communicability translates into a factor that measures the quality of service.

A second constructor aims at measuring the user's trust in a chatbot. This constructor has been inserted to evaluate how the user's intention to use a chatbot can be affected by how much she/he can trust the provided information and the consequent risk of accepting the processing of personal data proposed by it. Carlo's credibility is based on the users' evaluation and it is quantified according to how they consider it believable, competent, and trustworthy [11]. This constructor has been associated with a related factor concerning privacy management. Specifically, this factor aims at investigating the effects of the dialogues according to how users worry that their personal information may be used illegally.

Finally, the last constructor strictly associated with the communicability factor is the perceived transparency. This factor studied by using WhoTeach-PA aims at investigating how the lack of transparency in chatbot communication may lead users to give up on using it. To assure a high level of transparency, we need a chatbot that can explain the motivations and reasons behind its suggestions. Figure 3 presents the traditional and new UTAUT constructors and related structural relationships we used to study the acceptance and use of a chatbot for supporting users' interactions.

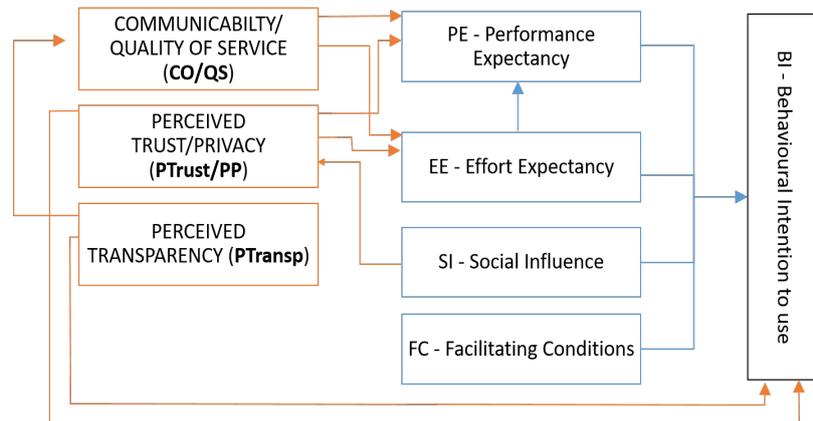


Figure 3: Extended UTAUT model that incorporates new constructors: communicability/quality of service (CO/QS), perceived trust/privacy management (Ptrust/PP), and perceived transparency (PTransp).

3.1. Objectives and Settings of the Experiments

To evaluate the relationships presented in Figure 3, we conducted a set of experiments involving our chatbots. Carlo was designed to suggest how to personalise the bank account services best suited to the user's needs through a dialogue that can inspire trust in accordance with the user's privacy. A total of 18 users aged from 18 to 80 were recruited to test the environment. Each user, after signing an informed consent document, was involved in a scenario test to evaluate her/his intention to use the chatbot. Males and females were represented with a distribution of: 61% males and 39% females; and their technical competencies in using instant messaging tools or other types of chatbots were equally dispersed. Their understanding of privacy issues and GDPR policy was considered sufficient by around 70% of them, but more than 80% of testers considered the privacy issues very important in the processing of personal data.

For what concerns WhoTeach-PA, we used it for understanding the factors affecting the teachers' intention to use a personal assistant to create courses. A total of 14 teachers aged from 20 to 60 were recruited to test the environment. Males and females were represented with a distribution of: 71.4% female and 28,6% males; and most teachers (72%) had between 1 and 5 years of experience in creating digital courses.

3.2. Data analysis and results

In all experiments, the constructors were tested to ensure an adequate level of scale reliability using Cronbach’s Alpha and Average Variance Extracted (AVE). All constructors in this study are greater than 0.7, which satisfies the rule of thumb of validity [12].

To validate the hypotheses at the base of each relationship presented in Figure 3 we carried out a set of SEM (Structural Equation Modelling) path coefficients’ estimates and test statistics related to all UTAUT constructors for the two chatbots: Carlo and WhoTeach-PA. The results, described in Tables 1 and 2, offer insights into the factors’ influence on the structural level or effect size. All hypotheses were confirmed with a high level of statistical significance of $p < 0.001$.

Table 1 presents estimates of the relations between the new constructs that have been introduced in the UTAUT model. The SEM analysis results have been filtered taking into account only users who are indicated to have good technical competencies in using chatbots (first part of Table 1) and proper expertise in privacy issues (second part of Table 1).

This analysis aims at demonstrating how when people with good technical competencies and a good understanding of privacy issues are involved in the interaction their degree of acceptance of using chatbots is very high. The positive values demonstrate how a chatbot that takes care of the data processing is perceived as a valid solution for tools of customer relationship management. Regarding WhoTeach-PA, the results in table 2 show high values for the constructor “Perceived Transparency” demonstrating how it has a very positive effect on teachers’ behavioural intentions to use a virtual assistant for creating a digital course. We obtained similar high values even when we tested the acceptance of WhoTeach-PA filtering the results according to control variables such as the teachers’ age and didactic competence.

Table 1: SEM analysis for Carlo chatbot.

Carlo Chatbot	Estimate	Acceptance	Carlo Chatbot	Estimate	Acceptance
CO/QS->PE	0,479	Accepted	CO/QS->PE	0,1504	Accepted
CO/QS->EE	0,325	Accepted	CO/QS->EE	0,743	Accepted
Ptrust/PP->PE	1,222	Accepted	Ptrust/PP->PE	0,361	Accepted
Ptrust/PP->EE	0,262	Accepted	Ptrust/PP->EE	0,0814	Accepted
Ptrust/PP->BI	0,406	Accepted	Ptrust/PP->BI	0,0757	Accepted
SI->Ptrust/PP	0,00596	Accepted	SI->Ptrust/PP	0,0876	Accepted

Table 2: SEM analysis for WhoTeach-PA.

WhoTeach Chatbot	Estimate	Acceptance
Ptransp->CO/QS	1	Accepted
Ptransp->BI	1,24	Accepted

4. Conclusion

In this paper, we have presented two chatbots, Carlo and WhoTeach-PA, specifically designed for investigating chatbot-user communication in the customer service, and personal assistant areas of interest.

The final goal of our research is to investigate the degree of acceptance of chatbot technology when it is used to empower users to carry out their work and daily activities.

The answer to this question stems from the study carried out with Carlo and WhoTeach-PA, through an extension of the UTAUT model to assess whether the user will be able to accept the new technologies and the user’s ability to deal with them. As discussed in the paper, the results of our tests demonstrate good effects on what concerns the acceptance of our chatbots. Specifically, higher results are achieved when the chatbot interacts with people who have proper competency in using such technology and good knowledge about trust and privacy issues. The new constructs introduced also provide the basis for future studies of factors influencing human behaviour and the intention to adopt AI-based systems.

Nevertheless, we are aware of some limitations that affect our study. The main issue concerns the sample size of participants in several of our tests. Recruiting a few users does not allow us to present a complete statistical confirmation and validation of the reliability of the collected data. For this reason, further research aims at extending the study by involving more users with a wider context of use in each of the key areas of interest in which the chatbots can be used.

5. References

- [1] Jadeja, M., & Varia, N. (2017). Perspectives for evaluating conversational AI. arXiv preprint arXiv:1709.04734.
- [2] Tsvetkova, M., García-Gavilanes, R., Floridi, L., & Yasseri, T. (2016). Even Good Bots Fight. arXiv preprint arXiv:1609.04285.
- [3] Følstad, A., Skjuve, M., & Brandtzaeg, P. B. (2018, October). Different chatbots for different purposes: towards a typology of chatbots to understand interaction design. In *International Conference on Internet Science* (pp. 145-156). Springer, Cham.
- [4] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
- [5] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. User acceptance of information technology: Toward a unified view. *MIS quarterly*, pages 425–478, 2003.
- [6] Anne Kohnke, Matthew L Cole, and Richard Bush. Incorporating UTAUT predictors for understanding home care patients' and clinician's acceptance of healthcare telemedicine equipment. *Journal of technology management & innovation*, 9(2):29–41, 2014.
- [7] Hoque, R., & Sorwar, G. (2017). Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. *International journal of medical informatics*, 101, 75-84.
- [8] Kaili Dou, Ping Yu, Ning Deng, Fang Liu, YingPing Guan, Zhenye Li, Yumeng Ji, Ningkai Du, Xudong Lu, and Huilong Duan. Patients' acceptance of smartphone health technology for chronic disease management: a theoretical model and empirical test. *JMIR mHealth and uHealth*, 5(12):e7886, 2017.
- [9] Fred D Davis, Richard P Bagozzi, and Paul RWarshaw. Extrinsic and intrinsic motivation to use computers in the workplace 1. *Journal of applied social psychology*, 22(14):1111–1132, 1992.
- [10] Sundar, S. S., Bellur, S., Oh, J., Jia, H., & Kim, H. S. (2016). Theoretical importance of contingency in human-computer interaction: effects of message interactivity on user engagement. *Communication Research*, 43(5), 595-625
- [11] Raj Arora, Charles Stoner, and Alisha Arora. Using framing and credibility to incorporate exercise and fitness in individuals' lifestyle. *Journal of Consumer Marketing*, 2006.
- [12] Chia-Ying Li. Persuasive messages on information system acceptance: A theoretical extension of elaboration likelihood model and social influence theory. *Computers in Human Behavior*, 29(1):264–275, 2013.