Toward Privacy Implications in Developing and Deploying Chatbots: Analysis of Chatbot Development Platforms

Diva Smriti
Drexel University
Philadelphia, PA
ds3659@drexel.edu

Rahil Rathod
Drexel University
Philadelphia, PA
rr922@drexel.edu

Jina Huh-Yoo
Drexel University
Philadelphia, PA
jh3767@drexel.edu

ABSTRACT

The CHI community increasingly develops and evaluates chatbots for various areas such as health, business, or education. Chatbots have become widely integrated with third-party applications. This integration allows these third-party platforms to collect end-users' personal information so that chatbots can provide personalized marketing and experiences to the end-users. It is unclear, however, which third-party applications are integrated with various chatbot development platforms and how explicit their involvement is to the end-users. We address this knowledge gap

*Produces the permission block, and copyright information. See the specific order to use the table cells to include the authors in the order you want yourself and your co-authors to be listed. Use footnotes sparingly, avoid using them. There is a **white text** number 1 after the ABSTRACT heading to maintain this ACM copyright block space.

CHI'19 Extended Abstracts, May 4-9, 2019, Glasgow, Scotland, UK.

ACM ISBN 978-1-4503-5971-9/19/05.
DOI: https://doi.org/10.1145/3290607.XXXXXXX

KEYWORDS

Chatbots; Privacy; Third-Party applications; User Information;

by first examining nine text-based chatbot development platforms and the information given at the set-up process regarding various parties' access to end-users' information. Our findings indicate that third-party applications gain access to end-user information through sign-in, deployment, and developers' preferences. We discuss the limitations of not making access and control of end-user information explicit by third-party applications. Future work includes analysing privacy policy and terms of service documents to gain implications about the involvement of third-party applications when conducting research with chatbots.

INTRODUCTION

Chatbots have become prevalent for service-based use integrated with multiple social media, e-commerce, health, and other business platforms. Chatbots interact with their users in a conversational manner using natural language. Chatbots are used for many purposes including information access, education, therapy, and, particularly, customer service. End-users use chatbots for their efficiency and convenience, user experience, and for a sense of novelty. Accordingly, the CHI community increasingly designs, develops, and evaluates chatbots in our research. This technology has major advantages in offering novel user interaction, but it can also carry additional privacy risks associated with users' personal information. Prior research showed consumers' privacy concerns with conversational agents in the form of chatbots. This is because users interact with the bot in a natural conversation form, and sensitive information can be unintentionally shared by the users as they converse with the bot. Folstad et al. found that users had privacy concerns when interacting with chatbots making personalized recommendations [1]. Security researchers identified security vulnerabilities that would allow attackers to exploit and breach chatbots' data. Multiple third-party applications are also involved in performing natural language understanding/processing, speech to text, and artificial intelligence to process end-users' natural conversations with the bot. Third-party applications are those applications that are owned and operated by organizations that are external to the parent organization of the chatbot platform.

Studies explored what privacy concerns end-users have about chatbots and how it affects adoption, but little has been done to examine how to inform the end-users about how their data may be used in a complex integration of multiple parties of applications in enabling a chatbot. How end-users' information and data are processed in chatbots is still a 'black box' to the user, with users having little idea as to what happens beyond the conversations they have with the chatbots. Those who develop chatbots play an important role in understanding what end-user data is collected and analyzed by third-party applications, especially if the chatbot will be used for research and involves human research participants. However, it is unclear how much sufficient information is given upfront to the researchers who develop the chatbots about how end-users data will be handled by the chatbots and third-party applications.

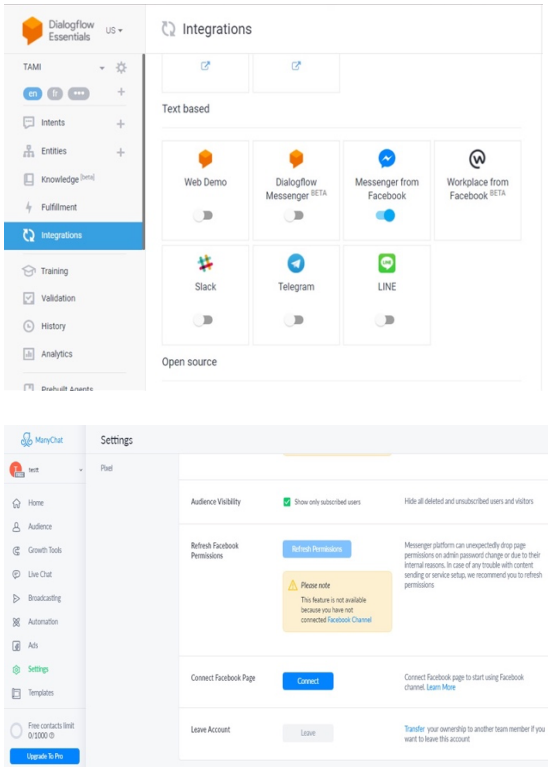

Figure 1: Screenshots of the set-up process of Google Dialogflow and Manychat chatbot development platforms analyzed for this study.

To investigate what access to end-user data third-party applications have when setting up a chatbot, we captured the process of creating chatbots from nine chatbot development platforms. We then conducted thematic analysis on any access or control over end-user information and third-party applications' involvements. We present our findings on the involvement of third-party applications and any information on the information types captured about end-users during the set-up process of chatbots.

METHOD

The research question we aim to answer is:
RQ: What control do third party applications have over end-user data as mentioned in the set-up process of a chatbot?

Data Collection

We searched online forums on chatbot developers and search engines to get a list of popular chatbot development platforms. The search resulted in thirteen platforms, out of which we excluded the chatbot development platforms that did not offer graphical user interfaces (GUI), because most platforms offered GUI and we wanted to standardize the data collection and focus on those that lay developers could easily create chatbots with. We also excluded those that were voice-based only and did not offer a text-based chatbot and those that were only available to organizations not individual developers. Table 1 shows the list of the nine chatbot development platforms that were included in the study. The researchers then navigated through the process of creating the chatbots, taking screenshots to document the process.

Table 1: Inclusion Criteria for Chatbot Development Platforms analyzed in this study

ID	Platform name	Developer GUI available	Voice vs Text-based	Available for Individual developers
P1	Google Dialogflow	Yes	Both text and voice	Yes
P2	Amazon Lex	Yes	Both text and voice	Yes
P3	ManyChat	Yes	Text	Yes
P4	Chatfuel	Yes	Text	Yes
P5	Wit.ai	Yes	Text	Yes
P6	SAP Conversational AI	Yes	Both text and voice	Yes
P7	IBM Watson	Yes	Both text and voice	Yes
P8	Microsoft Bot Framework	Yes	Text	Yes
P9	QnA Maker	Yes	Text	Yes

Data Analysis

Figure 1 shows a sample of the screenshots used for the analysis. First, the researchers inductively analyzed the screenshots in Nvivo 12 using thematic analysis [21] to generate three major themes based on the roles and interaction of third-party application with the end-users and developers. The themes resulted in third-party applications being used to sign in to chatbot, having access to end-user information depending on developer preferences, and being used for deploying chatbots. Second, from the collected themes, the researchers constructed summary tables showing the third-party application role and access to information in the different chatbot development platforms.

RESULTS

Our study identified largely three ways third-party applications are engaged: during the sign in process for the developers and end-users in setting up the chatbots, through developer decisions about what end-user information will be collected by chatbots, and during the deployment of the chatbots on third-party platforms. Table 2 summarizes the findings as a result of our analysis of the set-up process of nine chatbots.

Table 2: Inclusion Criteria for Chatbot Development Platforms analyzed in this study

ID	Platform Name	Requires end-user sign-in with 3rd Party	Developer control over access to end-user information by 3rd Party	3rd Party Integrations Via
P1	Google Dialogflow	No	Enable logging of end-user data; Bind end-user to data logging T&C	Deployment
P2	Amazon Lex	No	Enable logging of end-user data; Choose if COPPA is applicable	Deployment
P3	ManyChat	Yes (Facebook)	Agree with terms of service	Deployment; Export
P4	Chatfuel	Yes (Facebook)	Agree with terms of service	Deployment; Export
P5	Wit.ai	Yes (Facebook)	Agree with terms of service	Deployment
P6	SAP Conversational AI	No	Categorize end-users as vulnerable/non-vulnerable; Choose type of information collected from end-users; Agree with privacy policy and terms of service	Deployment
P7	IBM Watson	No	Agree with privacy policy and terms of service	Deployment
P8	Microsoft Bot Framework	No	Not mentioned	Deployment (through Microsoft Azure)
P9	QnA Maker	No	Not mentioned	Deployment (through Microsoft Azure)

Third-Party Deployment platforms for Google Dialogflow (P1): Oneclick TelephonyBETA (Avaya, SignalWire, Voximplant, AudioCodes), Telephony (Genesys Cloud, Twilio), Text based (Web demo, Messenger from Facebook, Workplace from FacebookBETA, Slack, Telegram, LINE), Open Source (Kik, Skype, Spark, Twilio IP Messaging, Twilio(Text Messaging), Twitter, Viber)

Third-Party Deployment platforms for Amazon Lex (P2): Facebook, Kik, Slack, Twilio SMS

Third-Party Deployment platforms for ManyChat (P3): Facebook Messenger, SMS, Email

Third-Party Deployment platforms for Chatfuel (P4): Facebook Messenger, Chatbot owner's Website

Third-Party Deployment platforms for Wit.ai (P5): Facebook Messenger

Third-Party Deployment platforms for SAP Conversational AI (P6): Users channels (Webchat, Amazon Alexa, Line, Messenger, (Skype, Microsoft Teams, Cortana via Microsoft Azure), Slack, SAP CoPilot, SAP Jam Collaboration, Telegram, Twilio, Twitter) Fallback channels (Intercom, SAP Contact Center)

Personal Information Collected from End-Users Due to Third-party Sign-in to Chatbots

P3, P4, and P5 required end-users to sign in through a third-party application, Facebook, to be able to interact with their chatbot. Then this third-party application will collect end-users' profile picture (P3, P4), language (P3, P4), gender (P3, P4), and name (P3, P4). While more information may be collected, that information was not made explicit during the set-up process. Although Wit.ai (P5) required end-users to sign-in through Facebook, it did not specify what information about end-users was being collected.

Developer Control over Third Party Application Access to End-User Information

The developers of chatbots are given control over third-party access to end-user information during the set-up process. Two chatbot development platforms (P6, P7) required developers to agree with their privacy policy, and five chatbot development platforms (P3, P4, P5, P6, P7) required developers to agree with their Terms of Service before moving ahead in the process of setting up their chatbot. The remaining four chatbot development platforms (P1, P2, P8, P9) did not require developers to agree with their privacy policy or terms of service to proceed further. These privacy policies and terms of conditions often include terms of third-party applications for end-user data. Only two chatbot development platforms, P1 and P2, explicitly stated that developers can enable logging of end-user data by the chatbots. P1 also stated that developers can bind end-user without consent to the data logging terms and conditions of the chatbot platform, including that of third-party applications. Other chatbot platforms provided no information about how developers can affect end-user data collection process by the chatbot platforms and third-party applications.

Deployment and Export of Chatbot on Third-Party Platforms

All nine chatbot development platforms had options for the developed chatbots to be deployed on multiple third-party platforms. This integration would increase chatbots' reach to more end-users with the help of the third-party platforms. The third-party platforms include Facebook, Slack, Twilio, Line, Amazon Alexa, Twitter, to name a few. The side bar lists all the third-party deployment platforms. Apart from deployment, chatbots and their data can also be exported to third-party applications for easy replication of chatbots. P3 had a feature of exporting its data to Google sheets, Shopify, Klaviyo, ActiveCampaign, MailChimp, HubSpot CRM, and ConvertKit. P4 could export its data to Google Sheets, Email, and Zapier. During this set up process, we observed no information on which information types will be collected from the end-users as part of chatbot deployment or export.

Third-Party Deployment platforms for IBM Watson (P7): Facebook Messenger, Slack, Intercom, Phone, Custom Application

Third-Party Deployment platforms for Microsoft Bot Framework and QnA Maker (via Microsoft Azure Cloud service) (P8 and P9): Add featured channel (Configure Direct line channel, Teams), More channels (Alexa, Direct Line Speech, Email, Facebook Messenger, GroupMe, Kik, LINE, Skype, Slack, Telegram, Twilio (SMS))

DISCUSSION

The findings shed light on multiple ways third-party applications involve in the developers' and end-users' interactions with a chatbot, and how CHI researchers should be mindful when developing and deploying chatbots and testing them with human research participants. Developers were given control of third-party applications access to end-user information in order to set up a chatbot, which in turn would approve end-user data to be stored in those applications' services. A few required end-users to also sign in with a third-party application in order to access the chatbot. All of the chatbots used third party applications to deploy themselves. Furthermore, the types of information being collected about end-user from the chatbots to these third-party applications was unclear. Any involvement of third-party applications in developing a chatbot development platform would mean some form of access and sharing of end-user data. However, without the analysis of the privacy policy and terms of use, this information was not made explicit to the developers when setting up the chatbot.

Such complicated integration of third-party applications in chatbots is often implicit for the end-users. In order to use these chatbots, end-users are required to sign in through a third-party application. For CHI researchers, when working with human research participants who will help evaluate the chatbots, the lack of control over human subject data pose potential challenges in key research ethics principles of data protection and privacy, such as informed consent, voluntariness, and coercion.

The findings also showed Facebook emerging as a common third-party application across the chatbot development platforms. But a recent incident of Facebook sharing end-users' personal data without consent to a consulting firm Cambridge Analytica have sparked concerns over end-users' data privacy [3]. This, in turn, raises concerns about how end-users' personally identifiable information is handled while chatbots that are widely being deployed without users' knowledge of its integration with third-party applications. This work in progress contributes to the developers' and end-users' thinking about the future ramifications of what they are consenting to while signing in to access these chatbots and the third-party applications associated with the chatbots. Our study is limited in its scope, since we only analyzed the set-up and registration pages of the chatbot development platforms, and not the privacy policy and terms of service documents.

REFERENCES

[1] Asbjørn Følstad, Cecilie Bertinussen Nordheim, and Cato Alexander Bjørkli. 2018. What Makes Users Trust a Chatbot for Customer Service? An Exploratory Interview Study. In Internet Science, 194–208.

[2] Kevin Granville. 2018. Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens. Retrieved January 11, 2021 from https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html?auth=login-google

[3] Anselm Strauss and Juliet Corbin. 1994. Grounded theory methodology. Handbook of qualitative research 17, 1: 273–285.