Durga Prasad Rangavajjala – 300123236
Yamuna Satheesh Kumar – 300115516
Vidhi Mistry – 300101658

Ethical Analysis

Design Problem: Amazon developers listening to audio recordings of users

Design Description

The system used here is the Amazon Echo which is a smart speaker developed by Amazon. The Echo device connects to the voice-controlled intelligent personal assistant service Alexa, which will respond when you say "Alexa". When the wake-word is said the device starts to record the conversation of the user and stores it which can be used by the developers to improve or update the system. These conversations can get recorded without the user's knowledge which can impact the privacy of the users. Alexa is incorporated in many devices; however, Echo Smart speakers have Alexa technology at core, and hence the words Echo smart speaker and Alexa are used interchangeably in this report.

Key Value Tension: User's Privacy vs Amazon's Accessibility to Data

A report in TechCrunch says nearly 70% of US smart speaker owners use Amazon Echo devices. Moreover, growth of 13.7% in smart speakers is predicted, reaching a total of around 83.1 million Amazon Echo users in US [1]. With these numbers just limited to US, number of worldwide Amazon Echo users could be anticipated to be even more. It's worthwhile to study the value – 'privacy' that is important to everyone in this horde, and the practices adopted by the Amazon with respect to Alexa voice services.

Description of Key Value Tension

With every utterance of wake word of 'Alexa', or any conversation with Alexa using Echo speaker, the user generates a voice recording. These voice recordings are of value to both the user as well as the Amazon. They are of value to the user because it is personal data. On the other end, Amazon needs this data to improve Amazon services including the speech recognition of Alexa. Complications arise as the user might not want to share the voice recordings or he might not be knowing about Amazon accessing the data for Amazon's own gain. This brews the tension between the User's privacy and the Amazon's data accessibility.

With people setting up Amazon Echo speakers in many corners of their home including private spaces like bedrooms and common spaces like kitchen, guest rooms, to name a few, it is worthwhile to investigate how Echo carries out the conversation within those different spaces. With echo speakers being set up in bedrooms, offices and other private spaces, it spurs questions of privacy, which is discussed in the next section.

Ethical Issue: Privacy

Often overlooked is the aspect of human involvement in the machine learning especially behind technologies like Alexa. Not many know that every conversation with smart speakers is recorded, stored, and certainly not aware of the fact that Amazon's employees can listen to fragments of those recordings. In an investigation report by Bloomberg on April 10,2019, Amazon was accused of employing people around the world to listen to voice recordings from Echo users' homes [3]. The report made clear that some of the things user's say aren't just processed by computers but by real people. The report mentioned Amazon spokesman confirming hiring of contractors around the globe to transcribe and annotate the voice clips in order to improve the speech recognition systems. Charlie Osborne, a cybersecurity journalist shows her concerns in her article stating that idea of an unknown human listening in is enough to make her feel uneasy even though she has mundane interaction recordings with Alexa. [2]

Even though Osborne states her most of the conversations are mundane with the Alexa, her uneasiness might stem from the uncertainty that some private information might have got recorded along with the conversations. As hundreds of conversations occur in a household, and sometimes simultaneously, some background conversation might get recorded along with the conversation made with Alexa which user might not be aware of. Directly spoken to Alexa, or recorded indirectly in the background, private information knowingly or unknowingly may get recorded, for example information like: address, phone numbers, social security number, confidential medical or financial information, or any sort of information that user might not want to share with an unknown human. Even though with the Amazon's claim of using only a small percentage of the voice recording for parsing by humans and claiming the recordings being anonymized, they often contain enough information to identify or embarrass the user. Bloomberg reported the employees sharing amusing recording finds in their internal workplace chat rooms. Sharing of personal recordings even in closed - employees only group, unknown to the data owner is simply privacy breach. Moreover, in a controversy, though not related to Alexa and Echo speakers but similar smart speaker technology by its competitor Google, a news service - VRT News identified the users by listening to 1,000 leaked recordings of user's Google Assistant queries [4]. The news firm said that they could easily hear sensitive information in those recordings which made the job of identifying the users easier. If users could be traced back from their recordings, it becomes easy for malevolent actors to pull off frauds, harassment threats or even identity thefts. Such privacy concerns are even higher when Alexa Echo speaker is placed in a private space like bedrooms or offices, as often these are the places where people speak their mind out, engage in private conversations and acts that are not supposed to be recorded and listened without their knowing. Most people would refrain from using such a technology that invades their privacy.

With all the benefits and improvements, a collection of recordings could reap to the Amazon's speech recognition services, there lies potential risks to the privacy of the users on the other side. As the issue of privacy pertains to all the users, it's a critical ethical issue; and it should be addressed in order to maintain smart speaker technology market.

Design Decisions

This section lists some of the feasible design decisions for Amazon Echo speaker that would ease the key value tension and address the ethical issue arising from setting up Amazon Echo in variety of spaces including private spaces. Each design decision's value outcomes are also discussed.

1. *Amazon Alexa responding to two wake words*: As of now Alexa is responding to a single wake word and whenever the wake word followed by instruction is given, it is being processed and stored in cloud. This procedure is not prioritizing the difference between normal query and personal conversation with Alexa. So, our design decision is to have two wake words, where one wake word can be used for general purposes and data after processing will be stored in cloud and this data can be freely used by developers. Other wake word is for having personal conversation with Alexa, where the conversation is processed and stored locally in Alexa and deletes all the recordings once the conversation is completed. This wake word works like Incognito in browsers ensuring the user's privacy, data storing in cloud. However, user has to make a critical decision about which wake word to use prior to initiating a conversation with Alexa. Since, the personal conversation is not stored anywhere, it may affect the Alexa's accuracy and performance.

So, users need to be made aware of the consequences of using both the wake words for effective utilization of wake words.

Value Outcome: User's privacy increases whereas Amazon's data accessibility decreases because Amazon now only has access to the conversations made using the one of the two wake words.

2. *Privacy levels based on room*: Rather than giving same data sharing and storing options in Amazon Alexa, we can also have a design decision to restrict some features based on the room where Echo smart speaker is set up. When it is installed for the first time in the house, user's is asked to configure the number of rooms, type of each room with some defaults whether to share the data with developers and for testing purposes. When the Amazon Alexa is set up in a room that has data sharing option set to private then conversations with Alexa in that room will never be shared with anyone. According to Andrew Burt, managing user consent is very difficult and they can change their preferences anytime [8]. So, design decision gives the editable preference setting option for the user to change the privacy option for room at any point of time ensuring user's privacy over the amazon accessibility to data. Thus, the consent management should be dynamic enough to accommodate these changes.

   Value Outcome: User's privacy increases whereas Amazon's data accessibility decreases because Amazon now only has access to the conversations made from room or spaces with sharing option set.

3. *Data sharing based on user's permission*: The Alexa starts recording each conversation of the user. These recordings get stored on the cloud without the user's knowledge. We can have an alternate design decision to restrict the use of these recording by the developers. When the conversations are recorded and stored in the device a notification should be sent to the user regarding the storage of conversations. When the developers want to use these recordings to improvise the device, a message or notification should be sent to the user seeking the permission of the user. Then it is up to the user to give permission for the developers to use the recording or deny the request. Before sharing the recording to the developers, user permission is needed. So, the design decision gives preference to user's permission on sharing those recorded conversations before it is used by the developers. However, this approach can significantly delay the developing and testing tasks as they wait for the user's permissions. Moreover, user would have to go through each recording before giving permissions and it might become tedious as number of recordings increase.

   Value Outcome: User's privacy increases whereas Amazon's data accessibility decreases because Amazon now only has access to conversations manually reviewed as shareable by the users.

References

[1] S. Perez. (2020). Nearly 70% of US smart speaker owners use Amazon Echo devices. TechCrunch.com.
[2] C. Osborne. (2019). Amazon employees listen in to your conversations with Alexa. ZDNet.com
[3] Day, et al. (2019). Amazon Workers Are Listening to What You Tell Alexa. Bloomberg.com
[4] Hee, et al. (2019). Google employees are eavesdropping, even in your living room, VRT NWS has discovered. VRT.be
[5] B. Rubin. (2019). Amazon now lets you stop human review of your Alexa recordings. Cnet.com
[6] A Ng. (2019). Amazon Alexa keeps your data with no expiration date, and shares it too. Cnet.com
[7] C. Coons. (2019). Amazon responds to Sen. Coons' concerns about consumer privacy practices for Alexa devices. Coons.senate.gov
[8] A. Burt. (2020). How will the GDPR impact machine learning? O'Reilly.com