

Chatbots – A Case Study

Rutika Pawar, Abhishek Joshi

Abstract—Chatbots are tiny programs that help simulate interactions with users automatically based on a set of pre-defined conditions, triggers, and/or events/ The employment of chatbots is unmeasurable, From being first-tier support for customers to help visitors navigate your website to diagnosing a disease chatbots are now been expanded to all sectors. They are widely used in various fields viz, Entertainment, Sales & Marketing, Education, Corporate, Medical and so on. Chatbot can be designed and built using various technologies such as Artificial Intelligence, Natural Language Toolkit (NLTK), Keras, long-short term memory (LSTM) which is an artificial recurrent neural network and many other technologies. This paper describes functioning and designing of some famous industry driven voice enabled chatbots such as Alexa, Siri, Google Home, Cortana along with a detailed comparative study. Through this study one can get to know the features, it's construction, pros and cons of each of the mentioned chatbots.

Index Terms— Chatbot, Artificial intelligence, Machine learning, Alexa, Cortana, Google home, Siri, Apple, nltk, Itsm..

1 INTRODUCTION

As said, chatbots now, no more are the state-of-art technology but the necessity in daily routine.

The chatbot market size is estimated to grow from \$2.6 billion in 2019 to \$9.4 billion by 2024 at compound annual growth rate of 29.7%. Peeking into figures, it very well answers the question " Why chatbots are the future of market research"?

The driving force for such a huge leap counts various factors such as advancement in technology, rising market-customer demands at low operational costs.

Various companies such as Amazon, Google, Microsoft have come up with their own versions of AI(Artificial Intelligence) enabled chatbots such as Alexa, Siri, Cortana, Tay, and to name a few or a lot.

Chatbots now no more offer a bucket of services but are now been enlarged to handle diversified tasks such as booking an appointment or recommending a hotel for vacation. It follows your command and sings a playlist of your choice and enabling various electronic devices such as tube light, Air Conditioner (AC), Television(TV), etc.

It has even segmented to banking, financial Services , Retail & E-commerce, Tourism, Healthcare ,and many other. Also, there are various organizations and medical firms which are now using chatbot as a secondary advisor who when in need examines your physical and mental conditions by asking a set of questions, and according to responses, it suggest user some advices and medicines

We often may have visited websites which have chatbot inbuilt that helps to navigate through and look around when error comes across.

Although the frontend looks just more appealing but there are some key factors which needs to considered which choosing one for usage.

One should look for a chatbot which changes and updates information and provides highly - personalized experience and automated enough to provide inputs when and where needed.

[1]According to recent "Forrester Report"; even after implementing chatbot, various firms are lagging behind to achieve the level pf personalized jobs generating for them. Out

of the total, only 21% are capable of providing highly-customized automated services with a seamless experience. in turn affecting the market rate. There are some reason due to which chatbots fail to give consistent events.

Many users tend to avoid the installation of such machine sin workspace or in neighbourhood as it enables various security threats. Data-alterations phishing , data leak to sites, etc.

If these are excluded , then chatbots are definitely worth buying to automate any of province.

2 DESIGNING OF CHATBOT

Chatbot are designed for understanding the intent of user and delivering service based on the need. But there comes work for making them intelligent so as to enable it for performing all humanly tasks and maybe beyond that.

Some of Technologies that drives these functionalities are Automatic Speech Recognition (ASR), Natural Language Understanding (NLU) and Entity resolution(ER).[2]

Machines built using these opearate on various factors and models. Accuracy of such models is increased by manually annotating data over the system. But pitfall definitely arises as sometimes the dAata fedded int system doesn't respond with correct labels manulated.

Along with it, there are various frameworks and dvelopers tools which are required to build a chatbot

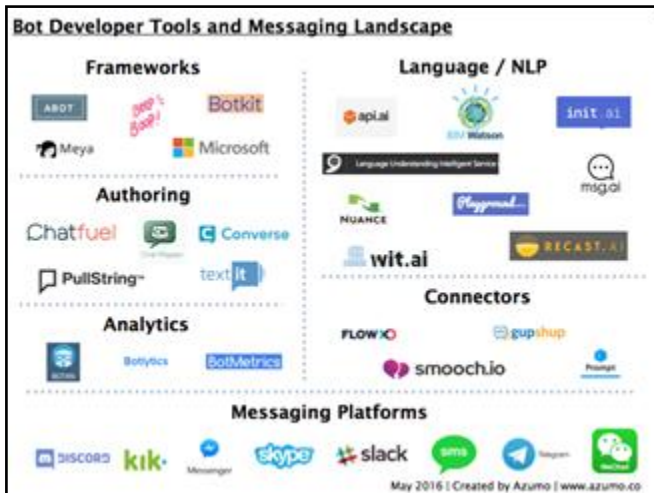


Fig. 2.1 Developer Tools for Chatbot

Developer tools can be classified as

1. Frameworks
2. Languages used
3. Authoring
4. Connectors
5. analytics
6. messaging/deployment platform

Usually while developing such machines, a strong research holds a strong foundation. The following questions are been asked & taken into consideration [3]

1. What types of commands do people make to their smart speakers (e.g. weather, music playback, smart home control, etc.) and in what percentages do they use these different features?
2. How are these devices used at different times of day or days of the week? Are there differences in the categories of commands?
3. How do these types of commands change over time as users become more familiar with these devices?
4. Do the topics change? Does the length of commands change per topic over time?
5. Are there any differences in use of these devices in different age groups or household sizes?

3 FEATURES OF CHATBOT

Taking into consideration the design and working of a chatbot; there are various features one has to deal with.

All the chatbots which are crated till now, are not equal. Some have 'X' unique features whereas the other one has 'Y' unique features.

Each system has the ability to detect intent and give output based on knowledge base as well as the algorithm's pace and adaptability. But there are some characteristics which need to be considered for comparing one chatbot with another.

Mentioned below are some characteristics which will help us to design a great bot.

- Conversational Maturity
 - Autonomous reasoning
 - Integrated with CRM
- Here CRM means Customer relationship management
- Emotionally Intelligent
 - Free to explore and adapt new changes
 - Pre-trained to the data well.

4 UNDERSTANDING VARIOUS DEVICES

4.1 Alexa

Amazon's greatest invention in field of chatbot; popularly known as Alexa has gotten in wide reach since its launch. Alexa comes under HDVA i.e. Home digital voice assistants. HDVA's perform variety of tasks such as they include playing music, ordering pizzas, shopping online, scheduling an appointment, checking weather, making a payment, controlling smart devices (e.g. garage doors, plug, thermostats).[4]

Alexa works and detects all sounds whose Sound pressure level (SPL) is more than 60 decibels.

According to a survey, after launch phase, Alexa now has been sold for over 5 million [5]. Adding to it, Alexa now offers services to user consisting of more than 10,000 skills which counts towards a huge number compared to other existing systems [6]

There are 3 kinds of Alexa devices: amazon echo, echo dot and amazon tap. Amazon echo is the 1st gen of Alexa devices. This device gets awakened upon calling her out by name "Alexa", continuing, it always stays in listening mode. Every time someone calls her up; she returns a voice command and then going back to listening mode.

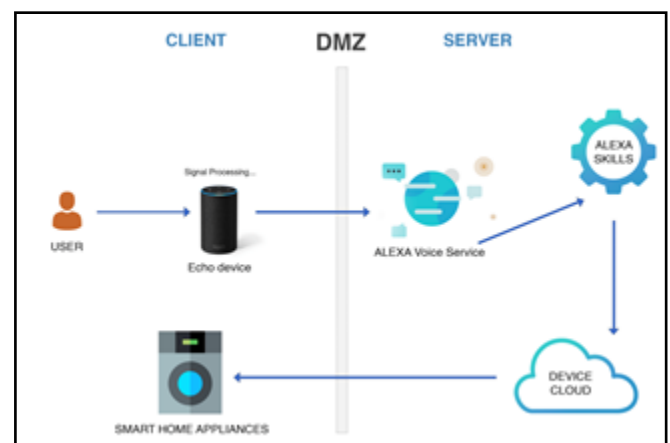


Fig. 4.1 Working of Alexa Echo Device

Amazon tap is similar but compact version and the last echo dot is the latest generation.

All the devices are connected to cloud-based voice system which after getting awakened, respond in voice message. Due to Alexa's voice service, it has an upper edge advantage of rising in market but it also introduces insecure access control. Two Vulnerabilities are explored in Alexa's access control & 1

vulnerability in smart enabled device.

The attacks may not be invasive attack where invader needs to be in victim's room or crowd attacks (many invaders reaching out victim). Often noticed, victims to digital breach are the users. As said earlier, the adversary need not be in victim's vicinity but can have control over other devices such as Google Chromecast [6], answer machines or Bluetooth speakers [7][8]

The major reasons for such threats can be:

- i) Weak single factor authorization
- ii) No physical presence-based access control
- iii) Insecure control to Alexa-enabled device cloud

To conclude HDVA's enable user to get automated in terms of living but somewhere, the security threat gets compromised. Emerging companies such as Amazon are continuously developing and upgrading, updating their system to increase security so as to prevent third party access and minimize such threats.

4.2 Google Home

According to a research conducted by Forrester, more than half of American homes will have a smart speaker device by 2022 [9]

Moreover, Google claims to be selling a Google Home device every second which accounts to a pace of 31.5 M/year.

To understand the existence and dependency of any device, we first need to check the daily usage.

On an average data of 110 days was collected per household. Analytics showed that median household issued 4.1 commands to Google home devices per day.

Concluding this usage was much higher than expected. Looking with broader perspective, 25th percentile household issued 2.5 commands per day and 75th percentile user issued 17.7 commands per day.

This showed that Google home was extensively used by people from various countries.

Looking into the technological aspect.

Command	Usage (in %)
Stop	7.3
what time is it	2.7
Pause	1.1
How much time is left	1.1
Pause tv	0.8
Skip this song	0.8
Play	0.8
Volume up	0.8
Tell me a joke	0.8
What's the temperature	0.8
Resume TV	0.7
Volume down	0.7
Next song	0.6
Next	0.5
Turn on kitchen	0.5
What's the temperature outside	0.5
Stop Tv	0.5

Turn on table	0.5
Turn off kitchen	0.5
Turn of living room	0.4
And many more.....

Table 4.1

Table 4.1 gives a brief description pf various commands used and its percentage of use.

In google home many microphones are been used to enhance the speech signal and reduce the effect of noise and reverberation. [10]

Taking into account the tradeoffs between computational complexity and performance, they proposed to use the factored Complex Linear Projection (fCLP) model [11] in the current work, which has much smaller computational complexity and similar performance to models trained in the time domain.

The fCLP model takes in complex and minimizes by applying sun-and-filter operations in first layer itself of network. the fCLP takes CFFT- complex fast Fourier transform. This CFFT is then passed to adaptive WPE frontend. the WPE then process data fed to fCLP which performs multichannel processing. Then by using Grid LSTM and passing through standard LDNN acoustic model, output target are achieved and sent back to the user [12]

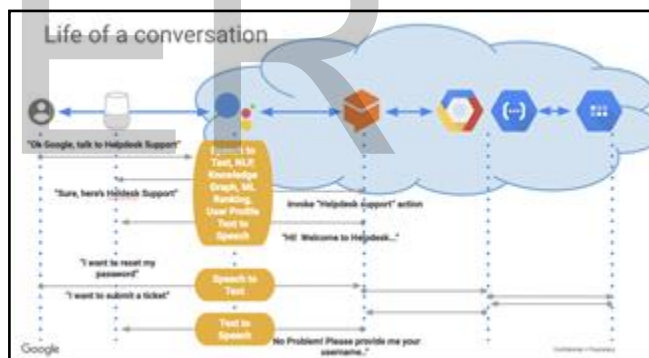


Fig. 4.2 Life of Google Home

Concluding the model, goal of getting perplexity lowered should be the aim of researches. This change can be brought by some improvements in model, algorithms, architectures also the data processed and finally computing the changes. Though Google Home is the best-selling and most accurate model yet there are some technicalities and loopholes which can be taken care of along with new adaptions

4.3 Apple's Siri

Siri developed by Apple is a personal assistant which communicates using speech. Unlike other voice/chat-based assistants Siri works on Apple based products.

Siri is been built using Speech Synthesis Work [13] The personality of Siri is to find a professional voice talent whose voice is both intelligent and pleasant. Siri works on principle of TTS text-to-speech system. TTS finds a sequence of units which matches with the input text and predicts target without

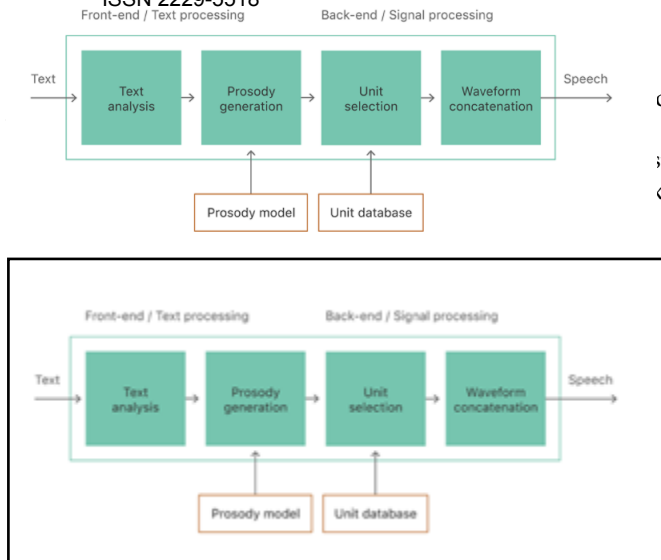


Fig. 4.3 Siri's Model

Unlike other assistants, Siri gets awakened by calling upon "Hey Siri". As soon as the device detects this and voice is triggered. Once audio is been recorded it then send to server and dataspace which can send new configurations to the model. It has a voice trigger framework which feeds audio into model, controls the threshold encountered while recording and sends some events related to wake-up.[16]

To conclude, Siri is definitely one of the strongest competitors, and day-by-day it is getting better by adopting few qualities from other models too This assistant learns on the history of the percept and makes the job done for the user, tending to minimize human efforts.[17]

4.4 Cortana by Microsoft

Cortana is voice based virtual assistant, another invention by Microsoft.

"Cortana is always looking out for you, remembering things so you don't have to, proactively suggesting things you might need to do and helping you do more with less effort," said Kiesha Clayton, senior communications manager at Microsoft.[18]

Cortana was first released in 2014 as part of Windows phone 8.1, and then it announced conversational speech recognition system in late 2017 which had 5.11% error rate, lowest so far [19] Microsoft handles a virtual notebook for Cortana which stores user's personal information and anything that's approved for Cortana to see and use. It's not a privacy control panel, but it definitely gives you a little more control over what Cortana does and doesn't have access to.

The use case of Cortana remains pretty simple. Cortana has access to device as well as cloud, hence whenever we seek to help it goes backtracking to search history and connecting to your cloud data base, interpreting the intent and then responding and showing results for the same. Cortana is ill-natured enabled device which suggests reminders connected to office 365, reads high priority mails and even helps user to find best meeting timings using voice commands

Core functions can be figured as pulling some LinkedIn profiles to creating personalizes planners and keeping a track

ch is of reservations and recommending activities. and Similar to other compatible devices, Microsoft (owner of ;ead Cortana) has another huge advantage when it comes to market ture reach as Cortana is available on all Windows computer and mobiles running on Windows 10.

5 COMPARISON OF VARIOUS MODELS

According to a study conducted by Loup ventures [20], they asked speakers, same set of 800 questions and were graded by some metrics. The metrics covered answering of 2 following questions

1. Did it understand what was said?
2. Did it deliver a correct response?

The question set comprising of 800 questions were all domain inclusive, starting from

- **Local** - Where is the nearest coffee shop?
- **Commerce** - Can you order me more paper towels?
- **Navigation** - How do I get to uptown on the bus?
- **Information** - Who do the Twins play tonight?
- **Command** - Remind me to call Steve at 2 pm today.

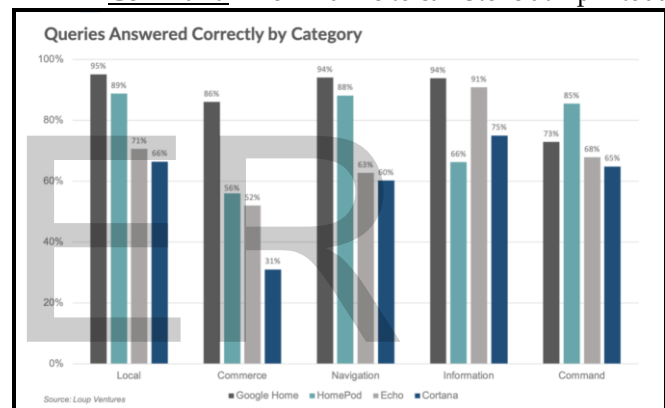


Fig. 5.1 Comparative Study of Question Types

After analyzing the trend and the responses the devices gave [Table 5.1], Google home outperformed as it answered 88% of questions correctly and understood all of them. Whereas Alexa's rate of understanding and answering correctly was 99% & 72.5% respectively. On the other hand, Cortana by Microsoft answered 63.4% of questions correctly and understood 99.4% of questions. Standing second in terms of understanding question, Apple's Siri graded 99.6% and answered 74.6% of questions correctly.

	Answered Correctly	Understood Query
Google Assistant	87.9 %	100 %
Siri	74.6 %	99.6 %
Alexa	72.5 %	99.0 %
Cortana	63.4 %	99.4 %

Table 5.1 Mapping of Results

After analyzing these figures, it was very well understood that though Alexa and Siri are been developing its own

independent device though Google's assistant got everything correct and answered after having a thorough research of the topic asked. Google home stood in the Local domain and navigation sector because of its integrity with google maps and other services. Whilst in other sectors/domains Siri, Alexa and Cortana took off.

6 CONCLUSION

6.1 Figures and Tables

With such high pace growing conversational chatbots, it becomes difficult for developers to cope up with the growing expectations of users. This paper was just an overview of designing chatbots and some minute details about specific chatbots launched by some prominent firms. Also, a comparison between all these gave a brief knowledge of their use cases along with the flaws. On a concluding note, immense research and development is been continuously occurring in this field, in turn contributing towards the advancement of technologies. Taking into consideration all the aspects and constantly evolving technology, the days are not much far when these fully automated machines will overtake all the manual work and uplift the essence of living and working in the community.

7 REFERENCES

- [1] <https://chatbotsmagazine.com/5-core-considerations-for-choosing-your-chatbot-4f4bb0856ead>
- [2] paper 1 downloads
- [3] Understanding the Long-Term Use of Smart Speaker Assistants, FRANK BENTLEY, Oath, USA, CHRIS LUVOGT, Oath, USA, MAX SILVERMAN, Oath, USA, RUSHANI WIRASINGHE, Oath, USA, BROOKE WHITE, Oath, USA DANIELLE LOTTRIDGE, Oath, USA
- [4] amazon-alexa-hits-10000-skills-plenty-room-grow/, 2017 .
- [5] "Amazon alexa hits 10,000 skills," <https://www.wired.com/2017/02/>
- [6] The Insecurity of Home Digital Voice Assistants– Amazon Alexa as a Case Study Xinyu Lei*, Guan-Hua Tu*, Alex X. Liu*, Kamran Ali*, Chi-Yu Li†, Tian Xie*
- [7] D. Petro, "Rickrolling your neighbors with google chromecast," in Black Hat, 2014.
- [8] "Hacking answering machines," <https://www.youtube.com/watch?v=bq6aV0Cxlh0>, 2012
- [9] Apple Computer. 1987. Knowledge Navigator. Available at: <https://www.youtube.com/watch?v=-jiBLQyUi38> Accessed 9/5/2017.
- [10] M. Brandstein and D. Ward, Microphone Arrays: Signal
- [11] Processing Techniques and Applications. Springer, 2001
- [12] T. N. Sainath, A. Narayanan, R. J. Weiss, K. W. Wilson,
- [13] T. N. Sainath, O. Vinyals, A. Senior, and H. Sak, "Convolutional, Long Short-Term Memory, Fully Connected Deep Neural Networks," in Proc. ICASSP, 201
- [14] M. Bacchiani, and I. Shafran, "Improvements to Factorized Neural Network Multichannel Models," in Proc. Interspeech, 2016
- [15] <https://www.indiatoday.in/technology/news/story/meet-meena-google-s-new-chatbot-that-interacts-like-humans-1641982-2020-01-31>
- [16] <https://machinelearning.apple.com/research/siri-voices>
- [17] Bishop. Mixture density networks, Tech. Rep. NCRG/94/004, Neural Computing Research Group. Aston University, 1994
- [18] [.http://ijarcet.org/wp-content/uploads/IJARCET-VOL-5-ISSUE-6-2021-2024.pdf](http://ijarcet.org/wp-content/uploads/IJARCET-VOL-5-ISSUE-6-2021-2024.pdf)
- [19] H. Zen, A. Senior. Deep mixture density networks for acoustic modeling in statistical parametric speech synthesis, ICASSP, 2014.
- [20] <https://machinelearning.apple.com/research/hey-siri>
- [21] <https://loupventures.com/annual-smart-speaker-iq-test/>

IJSER