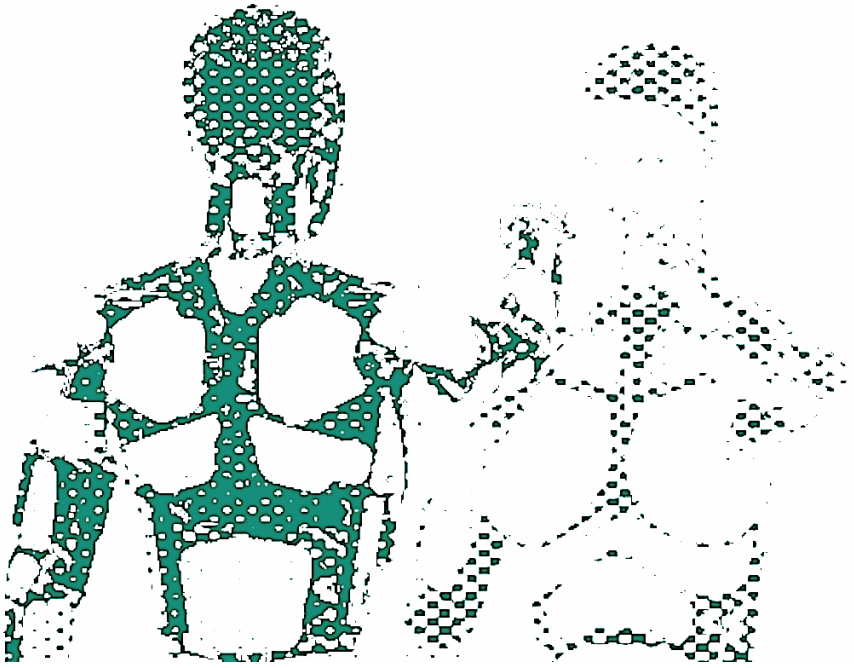
Intelligent
Scalable
System
Modelling



XIVA Assistant

Augmenting Your Combined Intelligence

White paper

Abstract

XIVA lays the foundation for your organisation to eventually automate all of its communication by learning in real time from your human agents and replicating their responses. We like to think of XIVA not as chatbot, but simply a bot. Which can scrape through your databases to fetch files, crawl through websites to generate answers it currently doesn't hold in its dialogue flow, it can do transactions or authentication of customer all by itself. XIVA's sole purpose is to perform every function that a human agent does and improve it over time.

Our chatbot is built via combining powerful Language models such as BERT, GPT-2, RASA and others, we use these Pre-trained models to build custom language models of each organisation that are optimised for their language needs. These models along with our own Machine learning models enable us to offer multiple language such as English, Urdu or Roman Urdu. XIVA is also able to digress between different languages and topics and has so far been able to achieve 0% bot detection rate.

Unlike other chatbot platforms, we offer a Bot-Builder Dashboard where our clients can build their own chatbot and dialogue flows according to themselves which the system later re-learns and improves on its own.

This paper will first outline the summary of the proposal and possible features and channels that can be included within the scope of this project. The paper will later elaborate on the details of each feature and the basic working of the XIVA Assitant to create a better understanding of the technology that runs it, and identifying further unique cases and problems that it can solve.

The Challenge: Growing Customer Expectations

As customers change, so must industry. Delivering efficient yet cost-effective customer service has always been challenging, but customers are now more demanding than they ever have been.

Customers today want information fast. They have little patience for a sub-par experience and they expect answers on their terms and their schedule. This is true across all demographics, but particularly acute amongst younger consumers. 33% of Millennials are only willing to wait 1 to 3 minutes to get a response to a customer inquiry before they drop.1 56% of Millennials have switched from one company to another because of underwhelming customer service.2 Fundamentally, if consumers are not given immediate, accurate answers they are more likely to leave for a competitor.

Organizations today are looking to solve these problems by digitalizing the customer's experience through self-help customer service solutions. Enabling customers to get the support they need themselves is highly desirable to most firms. Done well, self-service is not only a lower cost channel, it also has the potential to be the fastest way of resolving an issue. Early approaches to self- service technology have not always been successful. IVR (Interactive Voice Response) systems in particular have been contentious as, while cost-effective, they often have a negative effect on customer satisfaction.

Operational Support Costs

- 33% of Millennials are only willing to wait 1 to 3 minutes to get a response
- – 52% have hung up on a customer service call
- – 56% of Millennials have
- switched from one company to another because of underwhelming customer service
- – 270 billion service calls are made annually

The cost of providing round the clock support is traditionally high. In the current model, where the call center plays such a dominant support role, customers frequently experience problems resolving their issues, often being passed from agent to agent. Yet it is three times more expensive to handle and escalate a call than to resolve the query on the first point of contact. Put in perspective, there are 270 billion service calls annually, and roughly 50% of first calls go unresolved. Not only are customers unsatisfied, but companies also lose money trying to figure out how to help. Organizations also struggle to reduce the average handling time of each call or chat. Many firms find it challenging to get the right information to the customer as quickly as possible, greatly increasing the cost per call.

Furthermore, when customers cannot quickly get what they need, they are likely to defect to a rival. According to RCR Wireless News, for example, a study released by the WDS found that "Only 13% of customers show the level of loyalty that would prevent them from leaving for other competitive offers or because of service disruptions."

Even if an end product works well, customers may leave solely based on customer care. RCR Wireless News also reports that a customer is "2.5 times more likely to be a switch-risk if they don't feel valued or trust their operator" and "1.6 times more likely to become a switch-risk if they contact customer support."5 Retaining customers becomes nearly impossible when users are not willing to wait for their answers. Customers want to feel valued and that Assistant

Elements of XIVA

XIVA works on three main elements:

1.Natural Language Understanding(NLU)
2.Natural Language Generation (NLG)
3.Bot Builder

NLU
NLU converts text to vectors to identify the intention of the sentence and converts the incoming text to tokenizers with the extraction of entities.

How? A part-of-speech tagger or POS-tagger is used to tag each work with a part of speech like noun, verb, etc. then the Chunker chunks into groups the nouns with words related to them.

For example- you ask XIVA "what will be the weather like tomorrow?". The first process vectorises it and finds out intent which is "request weather". The next step uses tokeniser, POS tagger, and Chunker to finally extract the entity "tomorrow"

So we get — " request weather" and "tomorrow" which is pretty much what the question's motive was.

94% of marketers believe that "personalization of the digital experience is critical to current and future success."

Natural language generation(NLG)
NLG is a subset of AI and is powerful owing to its feature to take input of non-linguistic format and convert it into a human-understandable form.

Bot-Builder:
As the name suggests it is used to manage so that when the extracted data from intent and entity is received, it doesn't mess up.

RASA is a flexible framework that allows the developer to improve the accuracy of the models using various transform

Comparative Analysis

IBM Watson

82%

Precision by IBM Transformer Model

-

No Computer Vision

1 language

IBM watson comes only in English

XIVA Assistant

89%

ML Integration via BERT, GPT-2, RASA and Custom ML models of every organisation

4 use cases

of using Computer Vision for better customer expereince

More than 3 Language

Our Chatbot available in English, Roman Urdu and Urdu. Because of BERT could be expanded into other languages.

Some of the Use Cases

Lead Generation

Lead Generation has never been easier, now with the use of chatbots on existing chat platforms, to better reach your audiences. You can collect, sort, categorise and target users to generate leads for your various products. Our Lead Generation Chatbot comes in both English Urdu and Roman Urdu Languages.

FAQ's

Frequently asked questions do not require human intervention, but can significantly improve the users experiences over time. Whenever you acquire a new user it is common for them to require help familiarising themselves with your platform. You can now do that easily with XiVa.

Share Documents

Now find easy ways to disperse information with your users in a safe and secure manner.
With our state of the art Conversational Ai assistants that can share pertinent documents in the form of PDF's with your audiences in a safe and secure way, while also collecting data of the users in the process to evaluate your audiences.

Human Resourse Bot

Now find easy ways to disperse information with your employees in a safe and secure manner.
With our state of the art Conversational Ai assistants that can share pertinent documents in the form of PDF's with your audiences in a safe and secure way, while also collecting data of the users in the process to evaluate your audiences.

Booking Management

XiVa brings you the opportunity to be where your audience is, and easily allow them to make reservations in advance, collect the necessary details and provide a seamless solution to all their booking and reservation needs.

Interactive Website Redirection

XIVA bot can scrape through your data or website to generate responses for customers if they're not available within its knowledgeable

Online Transaction

CXIVA bot can can be connected via na API to conduct online transactions through a two factor authentication or any other verification means,

Block ID or Account

Banking Customers can Block their account or transactions via the bot

Account Opening

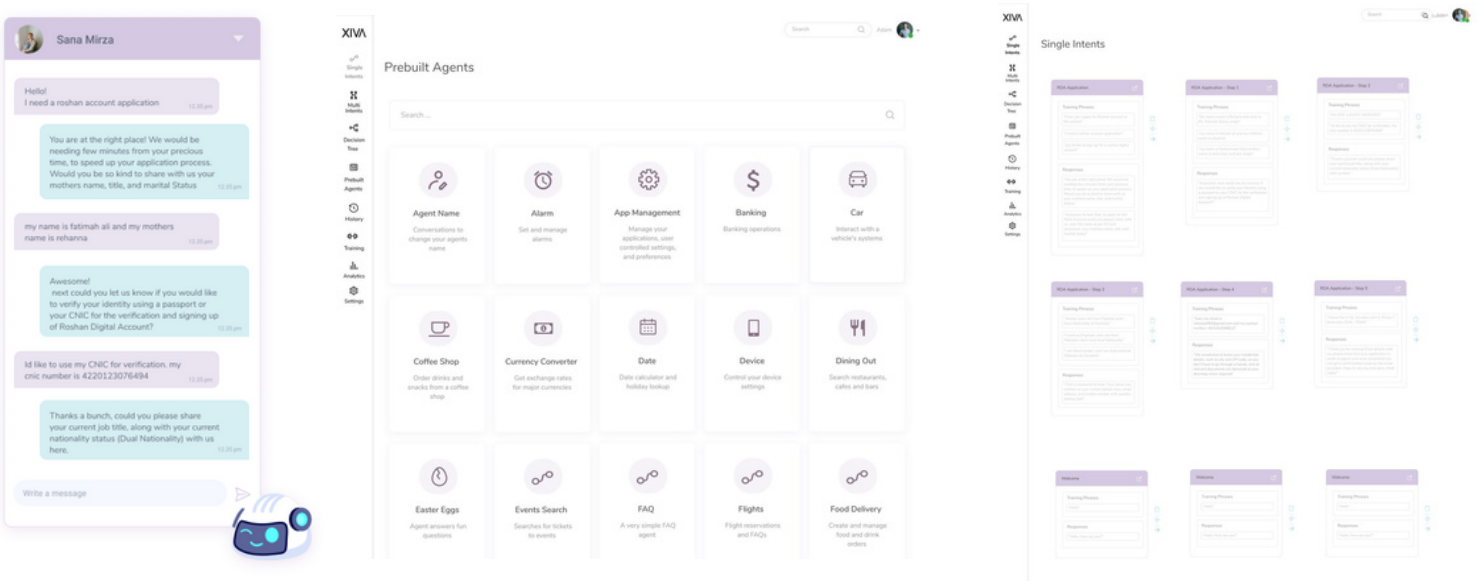
Account opening can be done via the bot, in some cases, customer can take pictures of handwritten forms which are read through computer vision and using Intelligent Character Recognition, information is passed onto Clients DB.

Xiva ChatBot Technical Architecture :

Our chatbot works using the Rasa Framework. It is further developed using BERT(Bidirectional Encoder Representations from Transformers) to improve accuracy. Our chatbot has 96% accuracy. Our chatbot is connected to the middleware. We are using NLP SpaCy for the named entity recognition(NER). To tokenize the data we are using a white space tokenizer. Apart from this, a state-of-the-art duckling HTTP entity extractor is being used to extract the numeric entities by the bot.

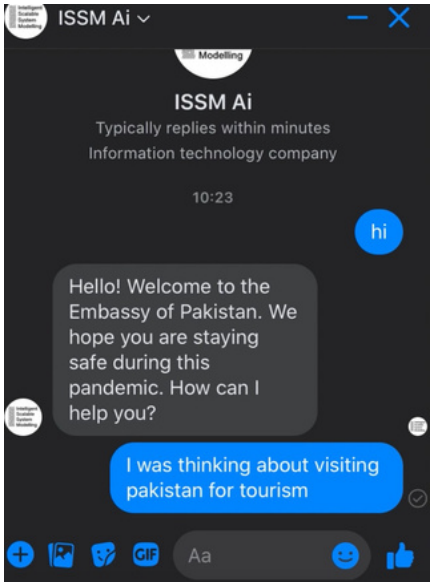
Stage-1
Bot-Builder:

The Bot-builder enables our clients to build their own dialogue flows, label intents and create responses.These include normal FAQ's and also action based responses where the bot fetches a response from an API or Database. BotBuilder has proven to be effective as it enables customers to create variations of the same queries for the bot to better understdand and respond. This also allows our clients to self serve rather than rely on their vendors to update dialogue flows.



Stage-2
Channel Integration

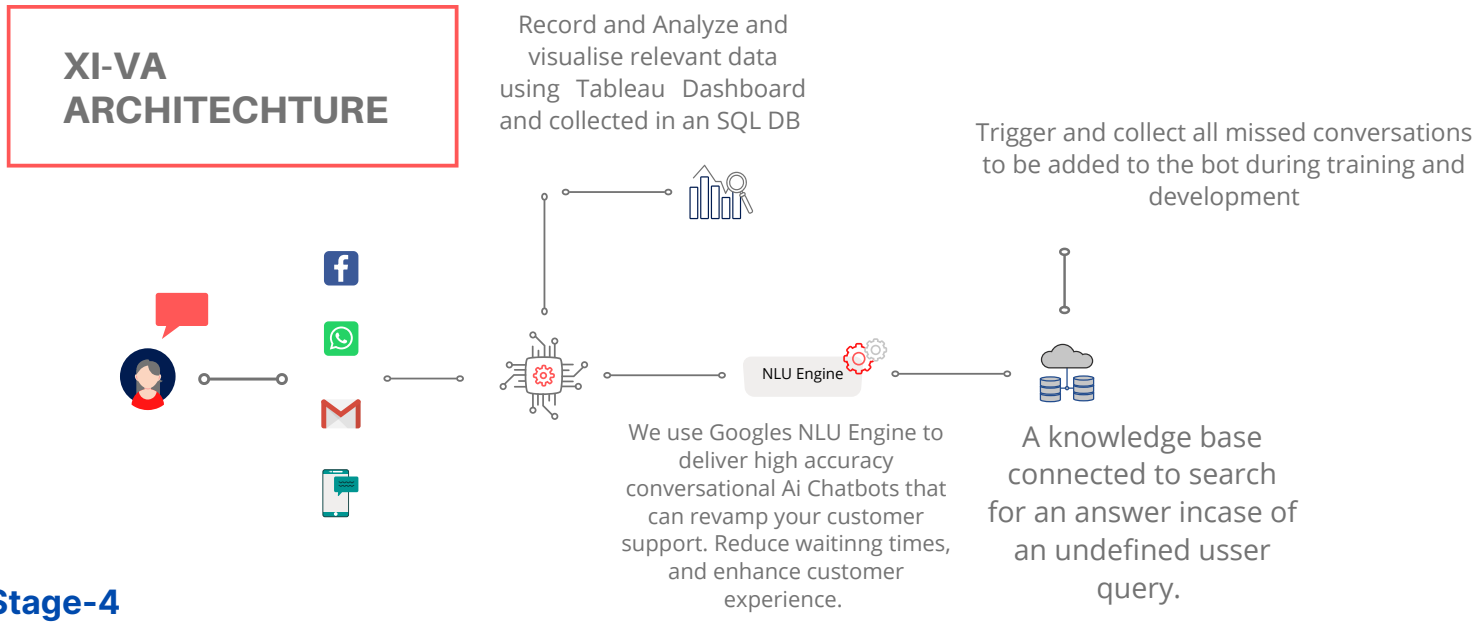
Once the Chatbot is ready, we integrate your chatbot with your social media channels such as WhatsApp, Facebook, Gmail etc, the bot can also be integrated with customer's website and App'sor other tools such as CRM or ERP depending on available API's and requirements.



Stage-3

Middleware :

We integrate API with the chatbot using the middleware. The middleware consists of REST APIs that enable us to send the data from the bot to specific platforms and also allow us to perform various actions such as dumping the data into databases.

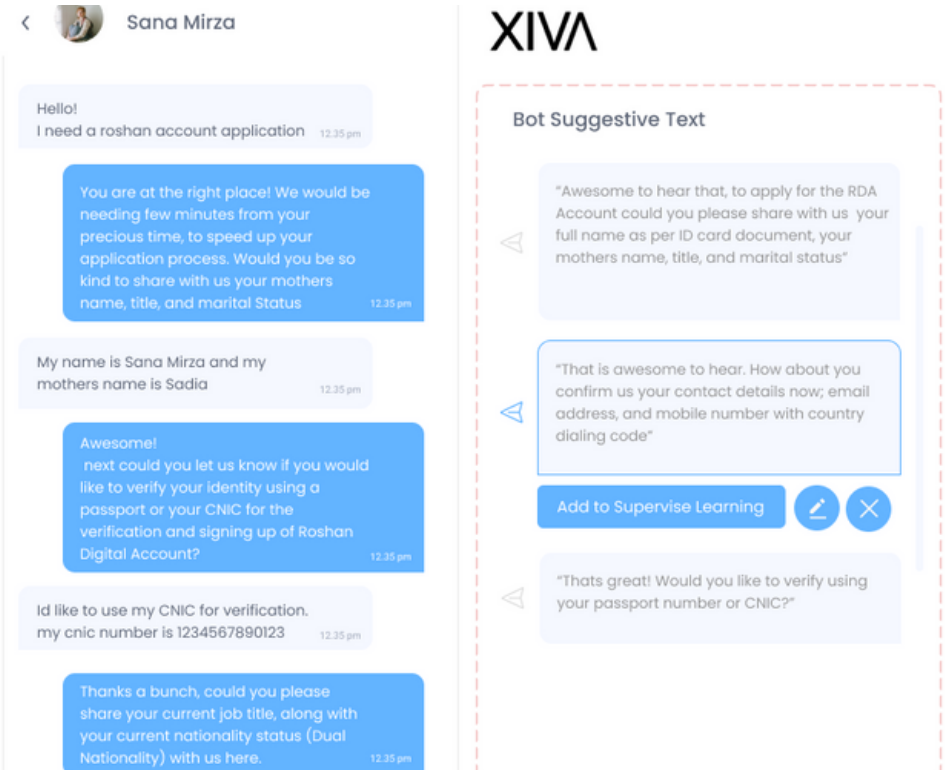


Stage-4

Agent Handover :

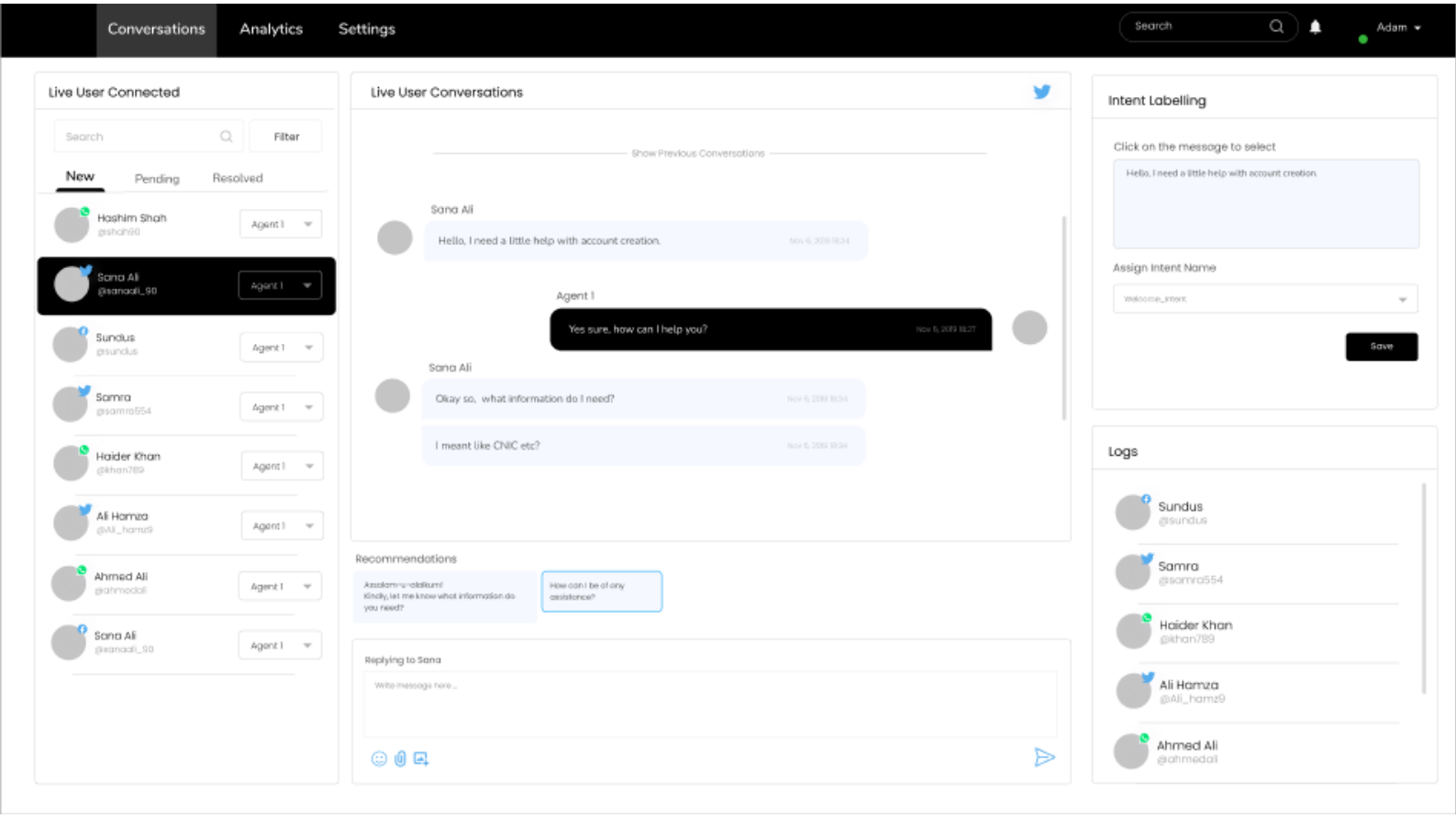
Agent Handover is also handled in the middleware. When the bot does not understand a specific question that is asked by the user and is unable to answer or if the confidence is less than 50%(User Defined) the bot hands over the conversation to the live agent. Session is closed once the conversation is transferred to the agent and if the user has to talk to the bot again, a new session will be made.

The conversation is transferred to the agent on different platform whether that is Genesys, E-ocean, or Infobip. These platform-owners expose us to their API or SDK that is triggered when a conversation is to be routed from the middlewareto the agent and vice versa. Once the platform exposes API and SDK, we define rules in terms of conversation re-rotation based on different scenarios. These scenarios may include, (1) Bot2Agent, or (2) Agent2Bot, and so on until the conversation flow is completed and the session is ended by the human agent.



Stage-5
Ticket Management System (TMS) :

The purpose of having a TMS is so you can manage all the live agents and the live customer that have been redirected by the bot due to any specific reason. TMS contains different features which are very useful. One of these feature is that whenever a customer is redirected to the live agent from the bot, that specific intent is stored and whatever the live agent responds the bot checks the response of that specific agent and cross checks against all the different responses from different agents and takes mode of it and trains the bot. Now the bot will not redirect the customer for that intent again and respond on its on based on the training from live agents response. TMS also contains star-rating system and a sentiment analysis. The star-rating system is so that the agent can rate the customer that if he was being rude or was in fact well-behaved. The basic purpose of sentiment analysis is for the machine to classify that what is the current mode of the agent as well as the customer. It is divided into 5 categories highly negative, negative, neutral, positive, highly positive. TMS shows you the sentiment of each individual statement as well as the combined sentiment for the whole chat.
TMS contains an analytics dashboard which shows you the different analytics extracted from the bot it self like which intent was recently called, which was the most used sentiment, how many users have interacted with the bot etc.



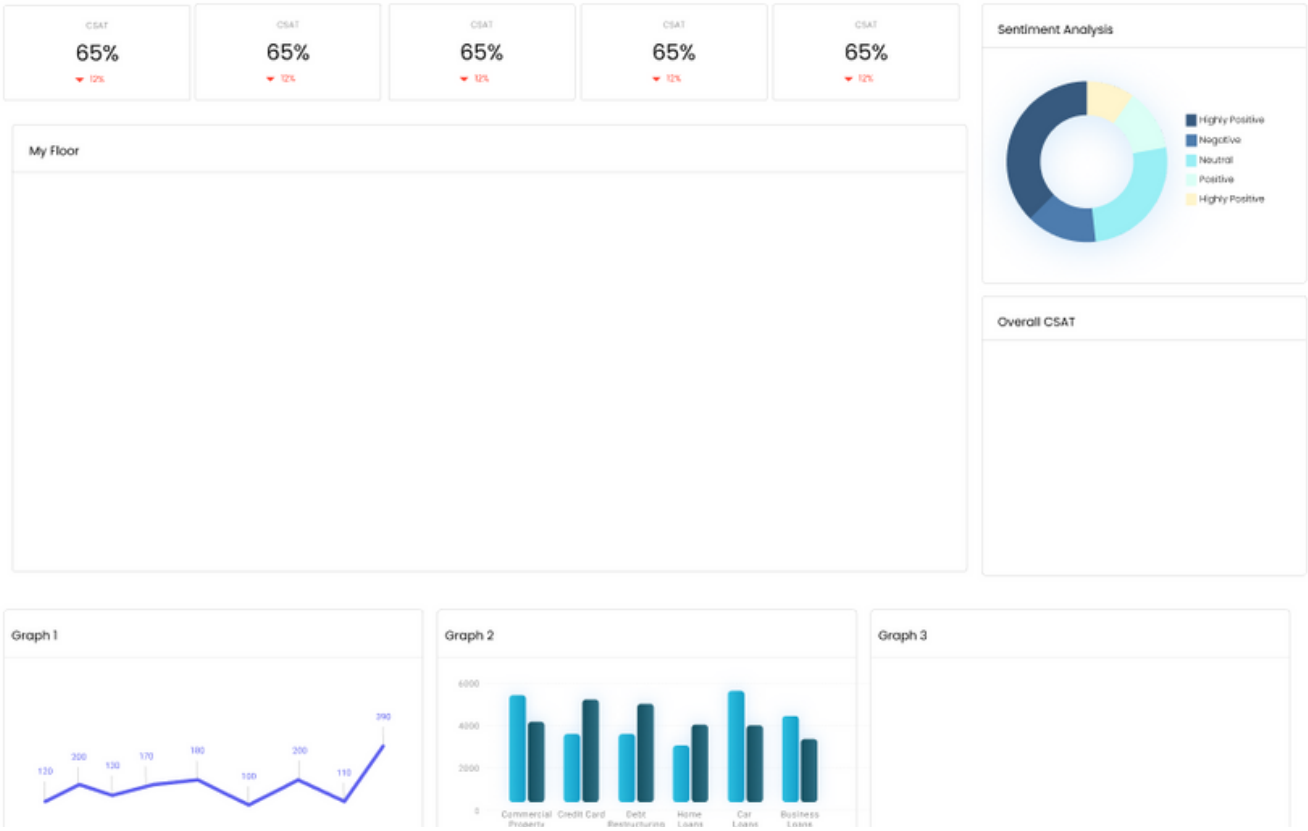
TMS also provides the following features that help the live agent in their confabulation with the customer:

- Profanity Filter
  - Profanity filter avoids agents as well as customers to use foul language/ Curse Words.
  - Profanity filter works not only in English but Roman Urdu as well.
- Machine Learning
  - TMS uses Supervised learning as well as Unsupervised learning to provide the user with a seamless experience throughout its conversation.
    - Supervised Learning
      - TMS allows users to make the bot learn new phrases using supervised learning, the users will be able to label the intent.
    - Unsupervised Learning
      - TMS learns on its own using unsupervised learning. The bot will identify a question that is being frequently asked and where the bot is unable to answer, the bot will learn to answer this question on its own using an unsupervised approach.

Stage-6
Analytics Dashboard:

Analytics Dashboard will analyze and visualize statistics based on the data that is being generated by the chatbot.
Following are the kind of analytics that will be shown:

- Agent Analysis
- Bot Analysis
- Costumer Satisfaction Percentage
- Consumer Analysis
- Sales Analysis
- Competitor Analysis



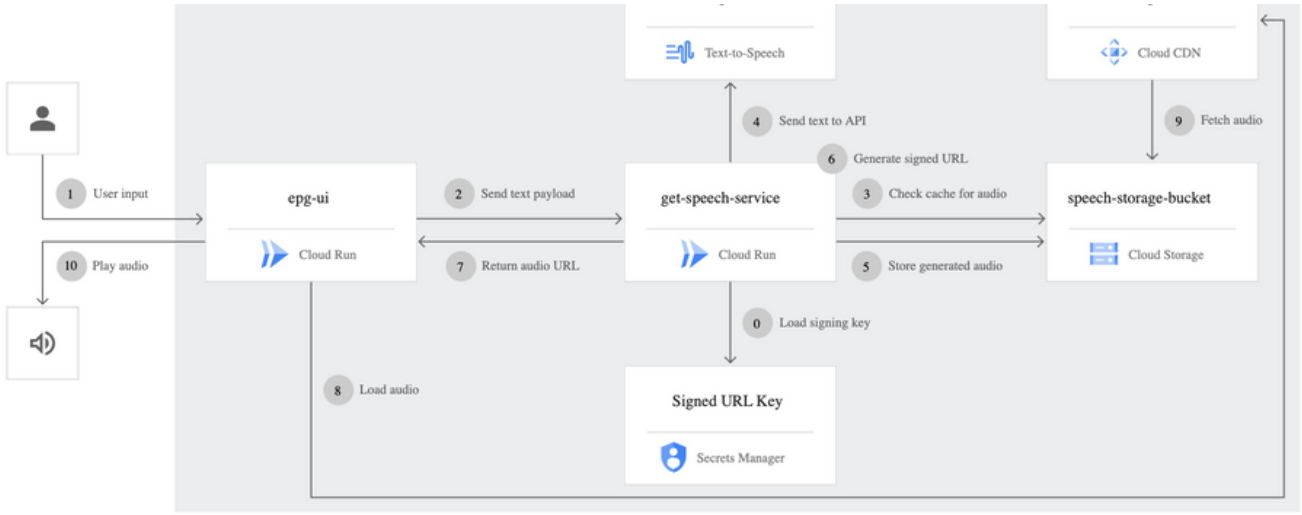
XIVA Speech Bot

The Solution- XIVA Speechbot

Our Speechbot is built on the same architecture as our chatbot, the only differentiating element is speech to text layer. Once the speech has been transcribed to text, our transformer models like BERT and NLP/NLU framewrok RASA contextualise those words and reply with existing audio messages. Just like our chatbot, these models along with our own Machine learning models enable us to offer multiple language such as English, Urdu or Roman Urdu. XIVA speechbot just like our chatbot is also able to digress between different languages and topics during a conversation.

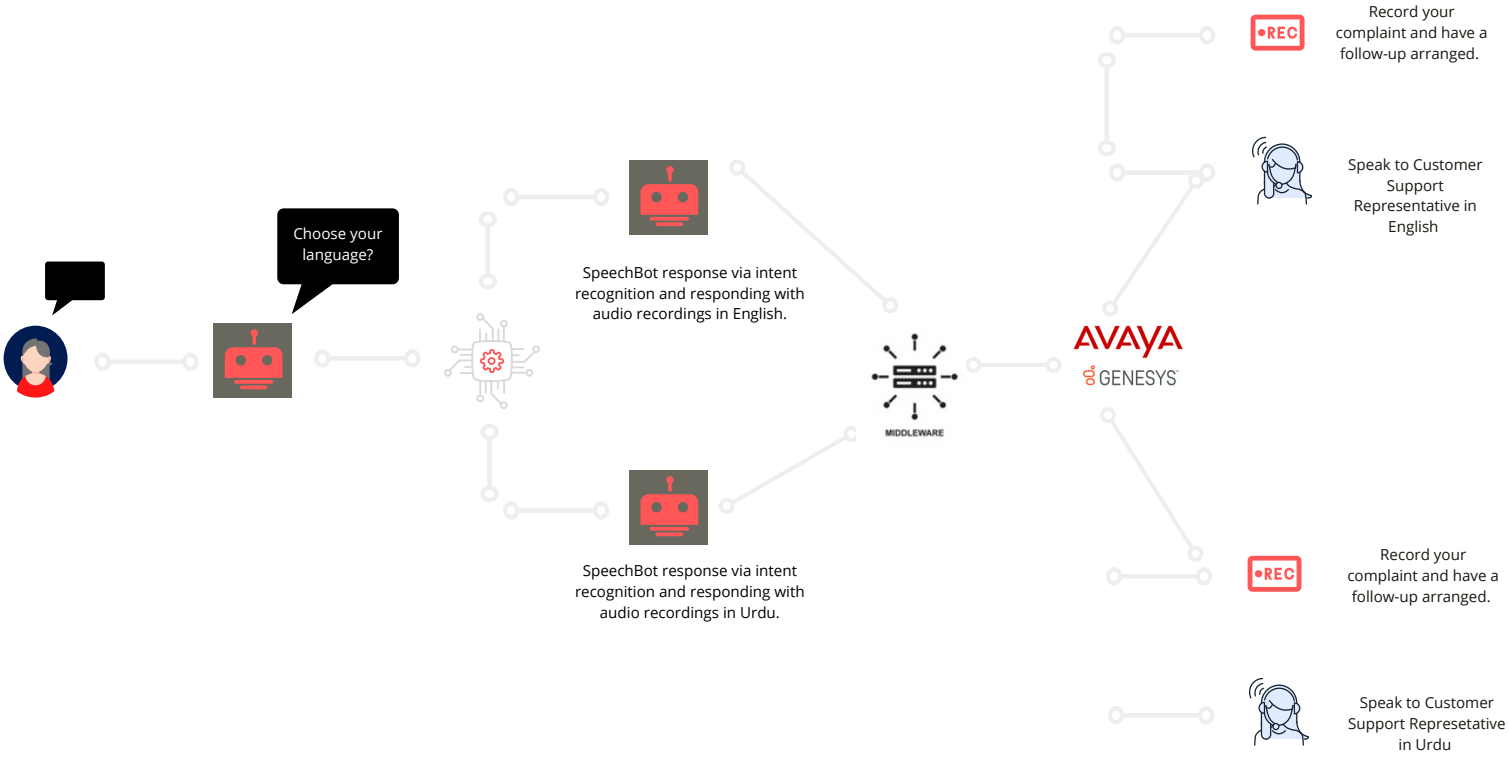
Our Speechbot is also integrated within our **Bot-Builder Dashboard** where our clients can create their own audio responses which the system later re-learns and improves on its own.

Note: Deployment and bot creation of speech bot is same as that of Chatbot.



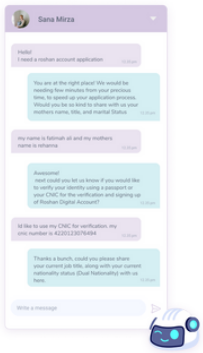
XIVA Speechbot Architecture

Speech Bot - User Journey



Phase - 1
Fully Conversational AI Chatbot
Deployment and Mechanism

1

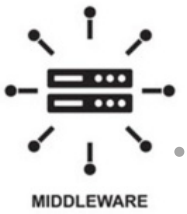


Dialogue Flow and intent labelling done Via the Bot Builder to form a chatbot

2

Integration of chatbot with different social media channels and Apps

3



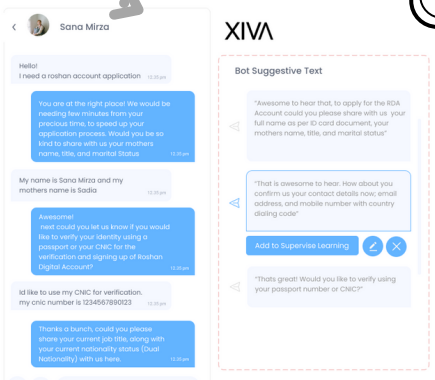
GENESYS

AVAYA

salesforce

Middleware re-routes the conversation from XIVA bot to Agent Platform.

4



Agent Can view previous conversations and see's possible responses from XIVA for its active learning.

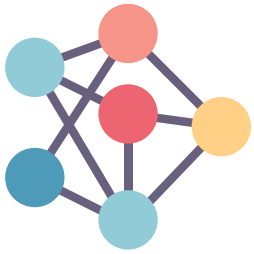
5

Relearning is done from the agent's responses

Phase -1 for Speechbot

Fully Conversational AI Speechbot

1



Speechbot ML model deployed in the first phase

2

ML Mode is re-trained on conversations of an Agent and the voice bot starts emulating those responses within its knowledge base.

3

During Agent Handover, the conversation of the call centre agent is transcribed and the responses of the agents are logged in the knowledge base. Just like the chatbot, the ML model is re-trained via unsupervised learning to generate these audio responses as its own.

Phase -2 Text + Speech

Augmenting and Emulating Human Responses

Deployment and Mechanism

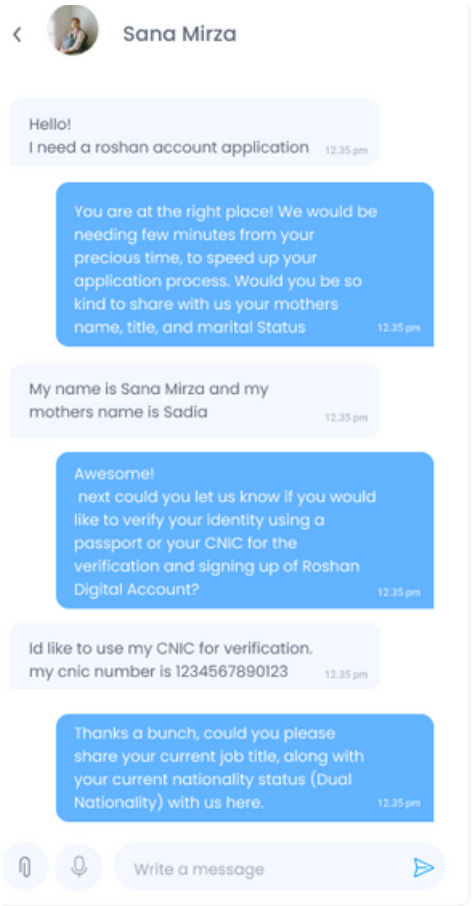
1

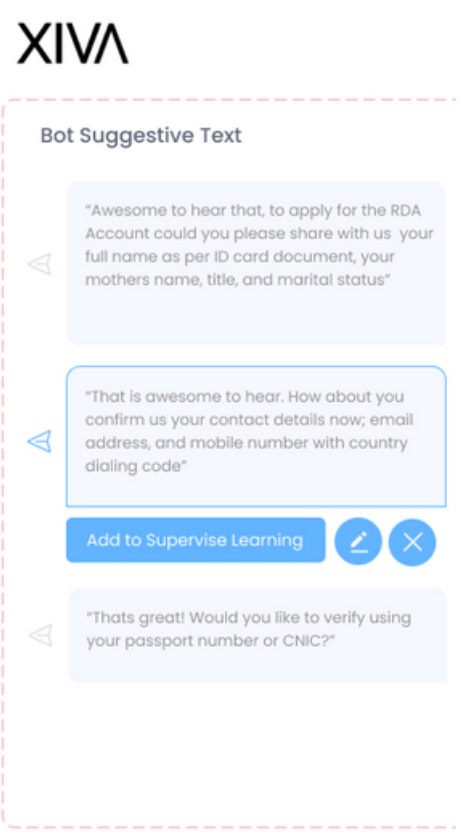
ML model deployed in the first phase

2

3

ML Model re-trained on historical logs of customer's data to further expand response volume by the Bot in order to expand to even those cases that are re-routed to agent.





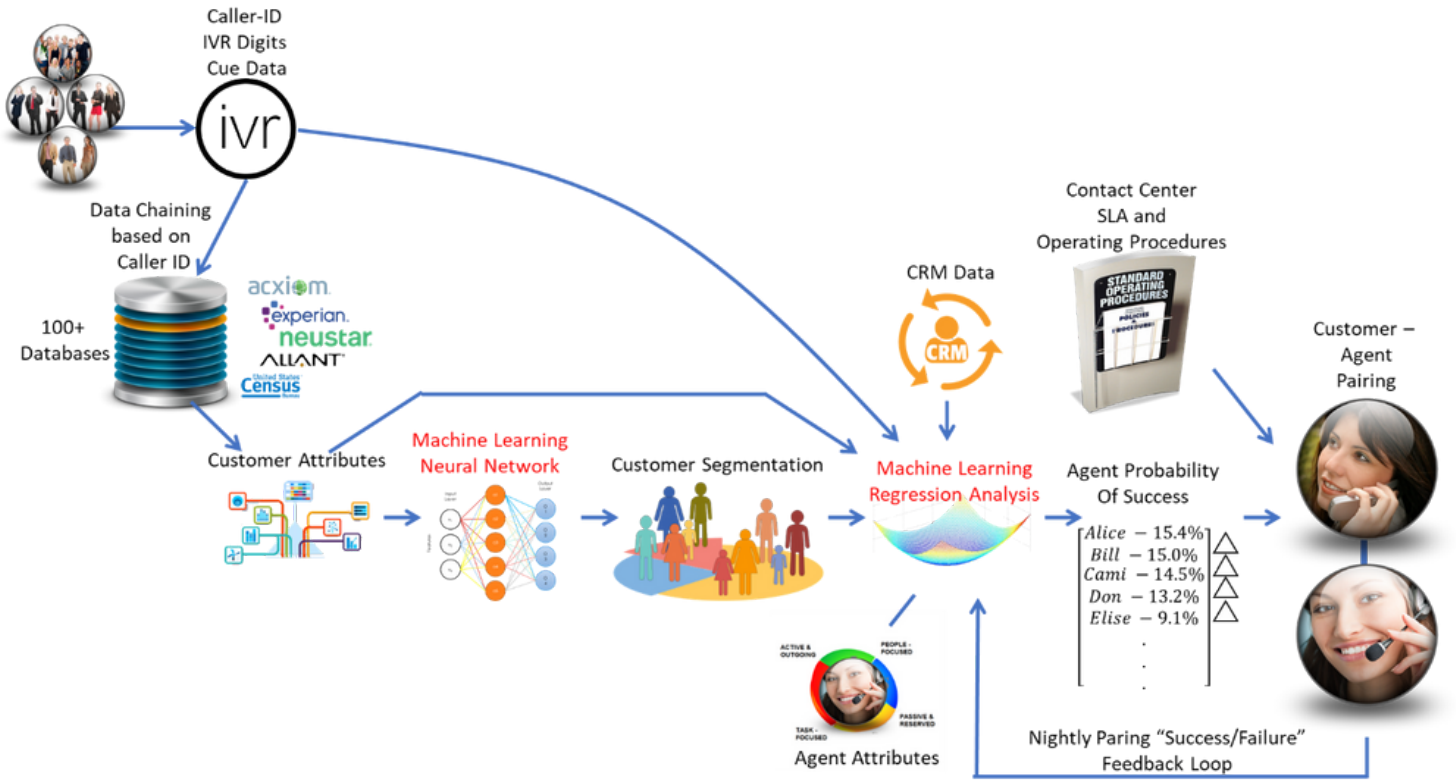
During Agent Handover, the conversation by the Human Agent are shown possible responses by the XIVA Assistant, if these responses are elected, the model re-learns them as its possible responses.

Over time as XIVA suggestions are selected by the agent, the agent handover decreases over time as the bot learns more and more from these responses, therefore fertiling the ground for complete automation of Customer Service.

Phase –3 Text + Speech

AI First Customer Service Features

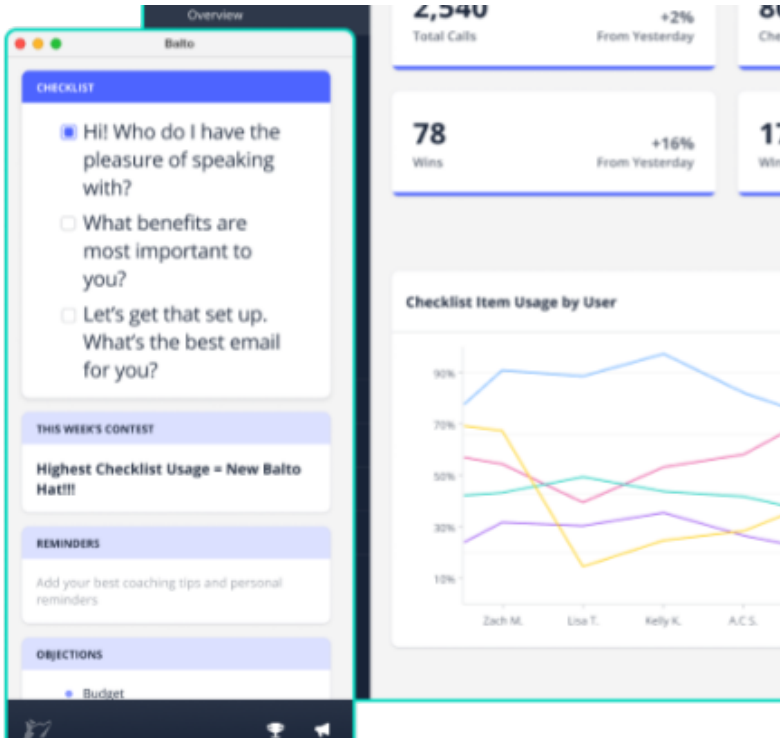
Behavioural Pairing of Agents and Customers



Based on the historical conversation of each customer and via data cataloguing of customer data, XIVA will start forming a data lineage of customer profile based on individual data points and co-relate them on type of query, department etc This customer segmentation and re-routing can also be used by XIVA bot to generate tailor made responses for each customer to optimse experience .

Recommendation Prompts for Agents

By further cataloging Human Agent's data and performances, XIVA can start adopting and learning best practices and responses and recommend them to the agent order to improve customer experience and in some cases increase sales.



ISSM Labelling Solutions
Csquare Pvt Ltd
rafay@issm.ai
Afnan@issm.ai

Document made by :
Abu Abdullah Ashar - Chief Product Officer
Rafay Asim - Team Lead XIVA
Muhammad Afnaan - Lead NLP and NLU Engineer

The End

XIVA