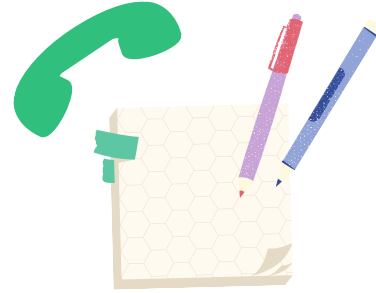


CALL SUMMARIZATION

Company: *Invoca*

Mentors: David Ron
Lauren Hedberg
Jasen Hall

Rachel Pham
Yinon Rousso



Team Name: #Koki's Kookies

Team Members: Koki Narimoto (Lead) Bryan Xu (Scribe)
Adarsh Garg Shaurye Mahajan
Sydney Lim Zion Wang

Introduction

What is the problem?

As a salesperson making hundreds of phone calls on a weekly basis, critical information gets lost.

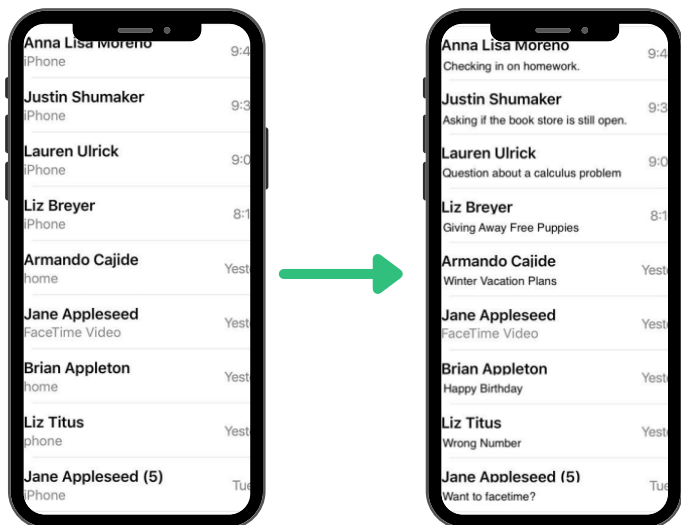
Whether this is from lack of notes or incorrect translation between conversations to text, there exists a critical need for call summarization within the field of sales folks.



What is our objective?

Remembering information from a sales call is extremely important. Regardless of how experienced a salesperson is, there will always be some information that is lost during the phone call. Thus, we need a way to refer back to our calls and recover this lost information. By using state-of-the-art speech-to-text and summarization technology, we want to give YOU the upper hand in securing sales!

Call Summarization



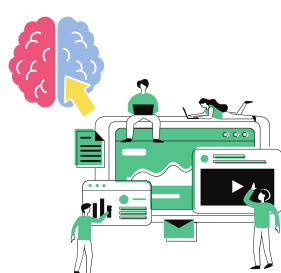
#KOKI'S KOOKIES TEAM

UCSB: Narimoto, Xu, Garg, Mahajan, Lim, Wang
Invoca: Ron, Hedberg, Hall, Phan, Rousso



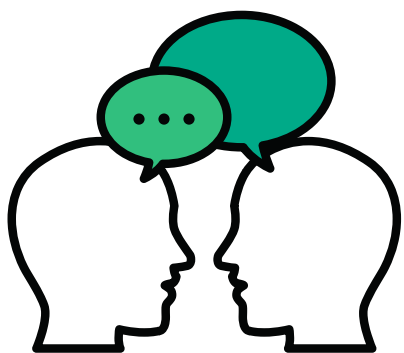
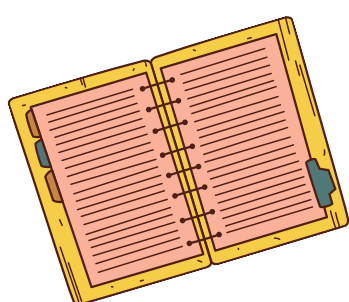
How will we implement our solution?

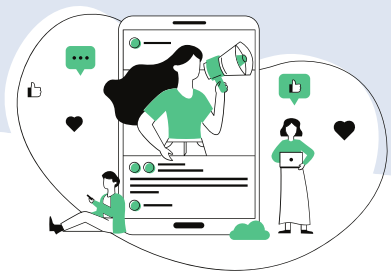
- Incorporate IBM Watson's cloud-based speech-to-text APIs to transcribe a conversation to text
- Incorporate natural language processing and machine learning techniques (i.e. sentiment analysis) to create an accurate summarization of the transcription
- Research/utilize API's to make summarizations easily readable and create visualizations to give users more options



How is the problem addressed today?

To this day, there are very few solutions to this issue despite there being an apparent demand for such a product: especially within the sales force. The main reasoning for there being a lack of existing solutions is that spoken language adds several degrees of complexity to the problem. Some reasonings for the complexity include the addition of filler words, accents, style, and background noise in conversations. Thus, the following presents some related solutions; however, none are identical to this product.



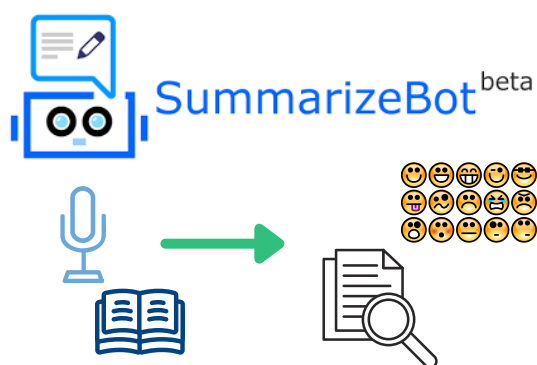


• Summarizebot

◦ SummarizeBot allows for real-time summarization and sentiment or intent analysis of audio and text data

◦ Pros:

- Real-time audio processing and summarization
- Keyword extraction
- Variable summary size
- Supports multiple languages including
 - English, Chinese, Russian, Japanese



◦ Costs:

- | | |
|---|--|
| <ul style="list-style-type: none"> ▪ Free tier (14 days only): <ul style="list-style-type: none"> • 5000 included requests • Up to 5 API calls per minute • 3 MB max file size | <ul style="list-style-type: none"> ▪ Standard tier (\$179/month) <ul style="list-style-type: none"> • 120,000 included requests • Up to 20 API calls per minute • 10 MB max file size |
|---|--|
- Custom Tier
- Unlimited requests
 - Custom API Calls limit

• Cognistx

◦ Cognistx is for companies to allow them to look for specific things that the consumer is saying. It can also format that speech into text form to allow them to have easy access to the information. Can also do call summarization and understand the overall call sentiment.



COGNISTX™

◦ Pros:

- Able to pick out specific answers and convert them to text
- Can fully summarize and find basic emotions of the call
- Intended for customer service



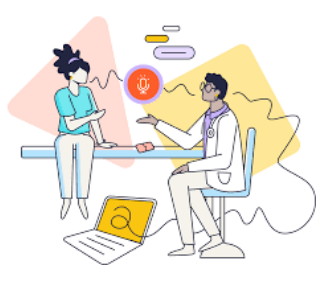
◦ Cons:

- Only a prototype



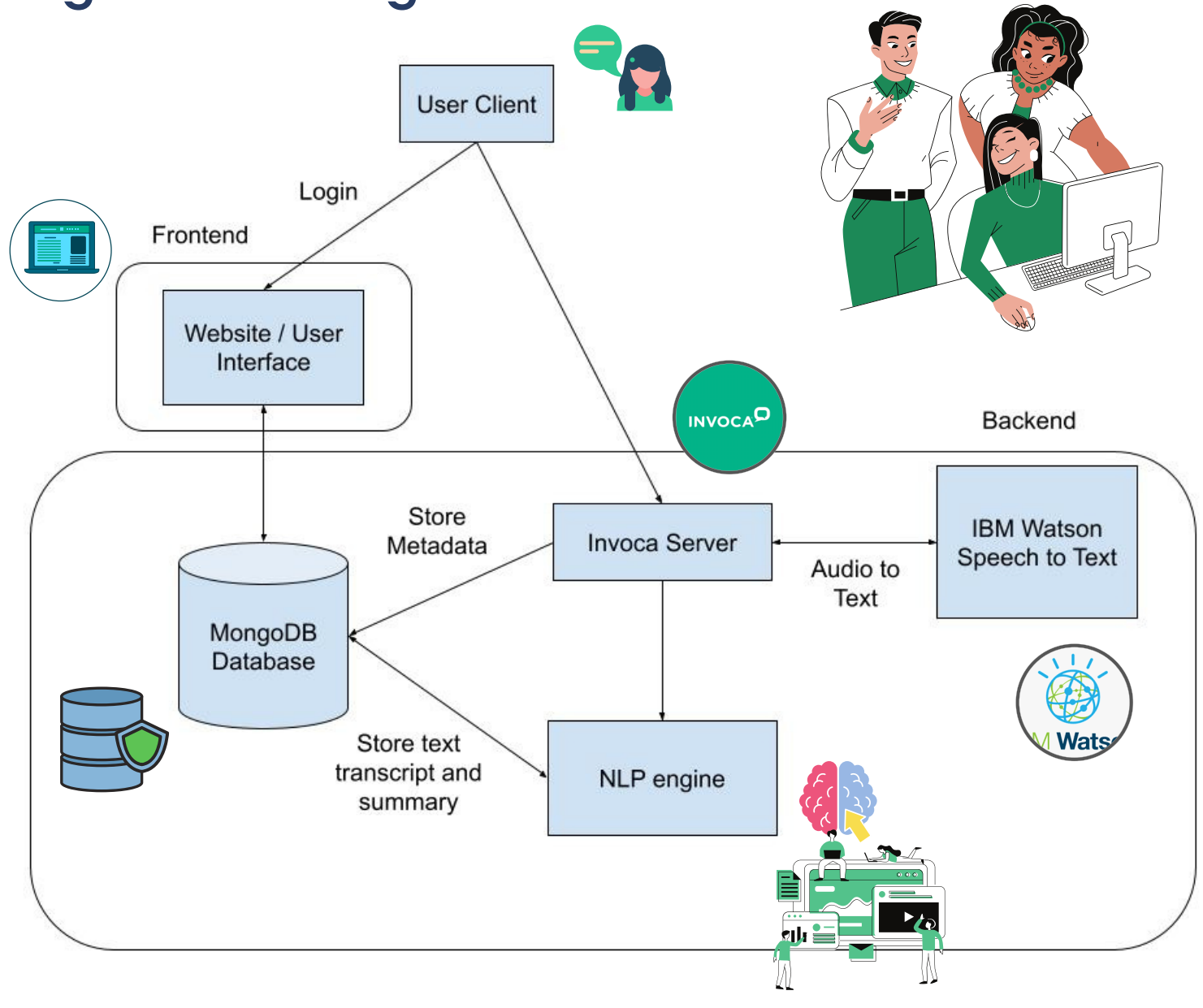
• Abridge

- Abridge is a service that transcribes medical appointments. It easily gives summarizations of doctor's visits and returns a format that makes it easy to share with patients
- Pros:
 - Free app for patients
 - Fast transcriptions and summarization
 - Highlights and emboldens keywords to make summarizations easier to read
- Cons:
 - Only limited to the medical field



System Architecture Overview:

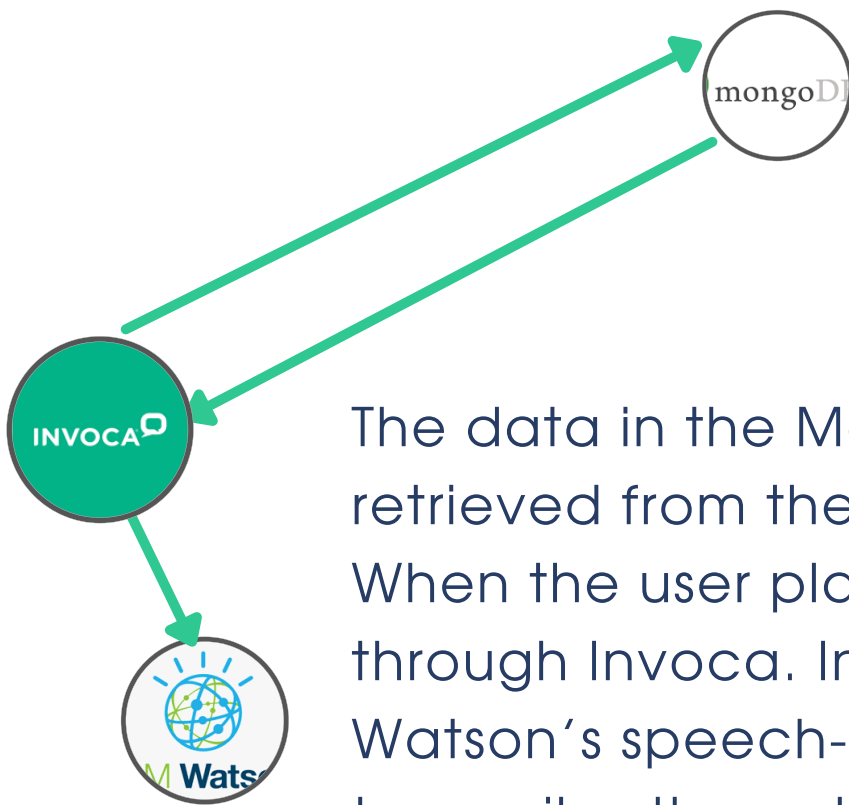
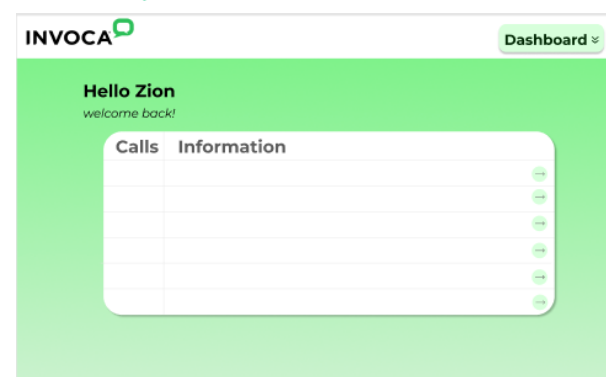
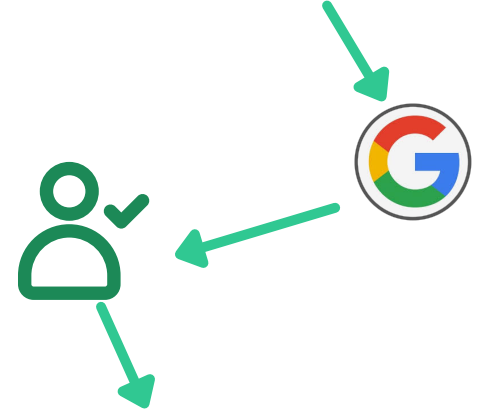
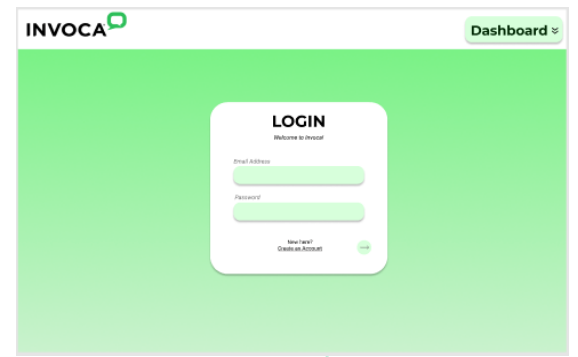
High Level Diagram





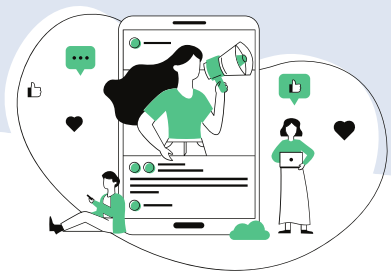
User Interaction & Design

The user accesses the front end through the website and user interface. They login with their credentials using their Google accounts which will be authenticated with OAuth. The user arrives at their dashboard, which contains a table of their call logs and summaries. This call data is retrieved from the MongoDB database

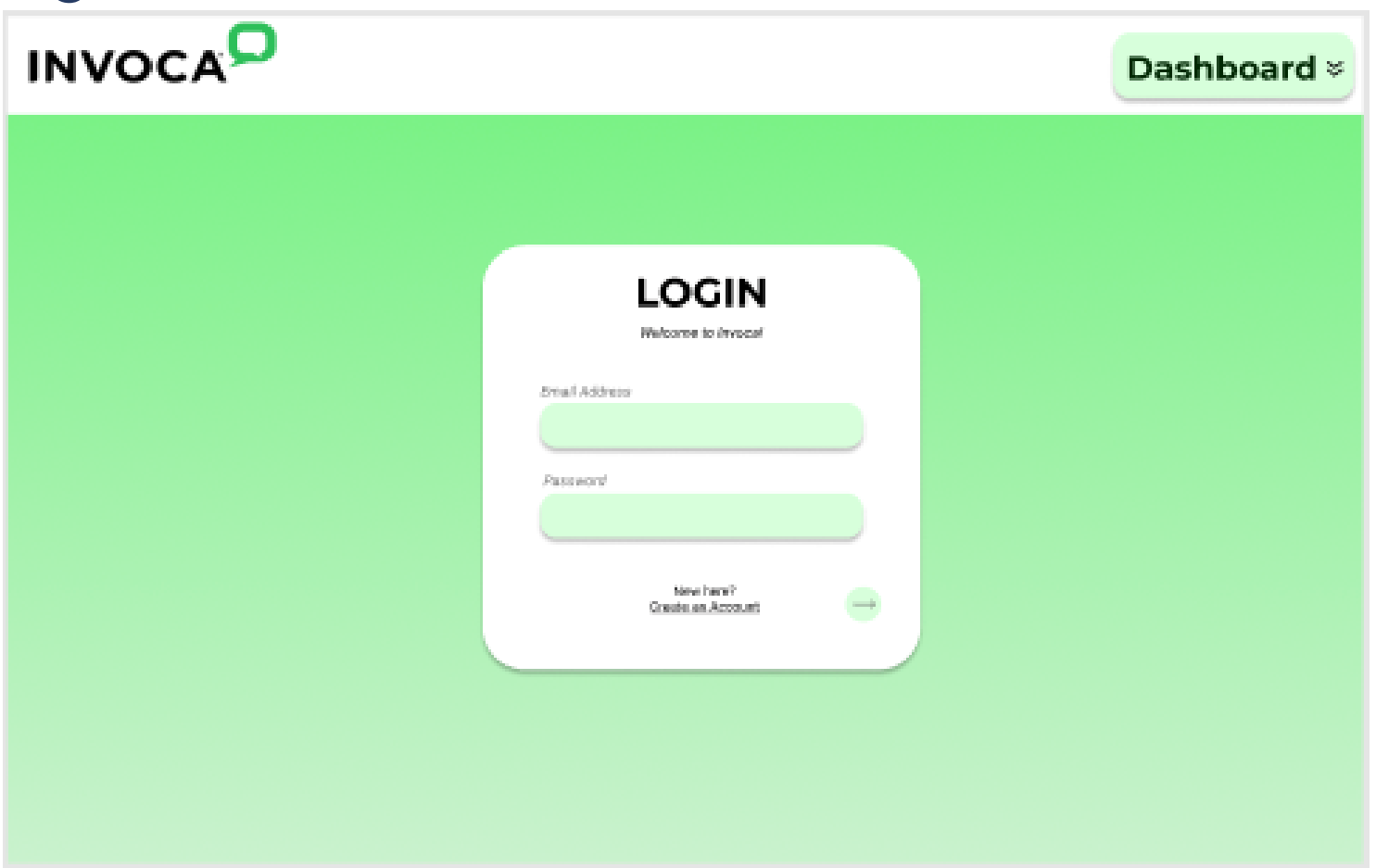


The data in the MongoDB database is retrieved from the Invoca servers. When the user places a call, it is routed through Invoca. Invoca uses IBM Watson's speech-to-text API to transcribe the call. We receive the transcribed call from Invoca's servers. First, we store the metadata directly in our database. This includes information such as the call time, call length, and customer's phone number. The full transcription of the call is sent to our NLP engine, where we will create a concise summary of the call as well as determine the customer's sentiment. Both the full text of the call and the summarization are stored back in our database.

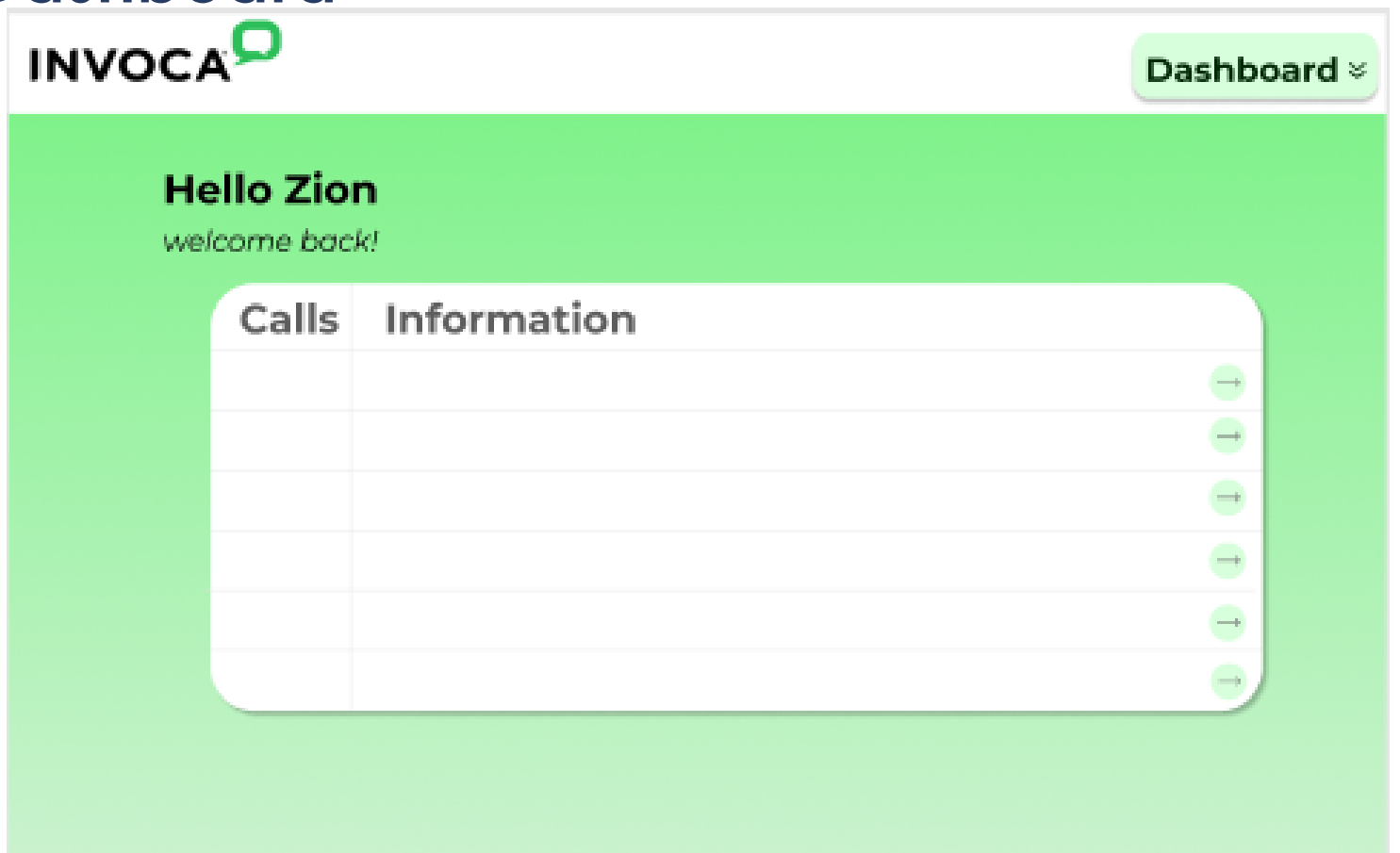




Log In



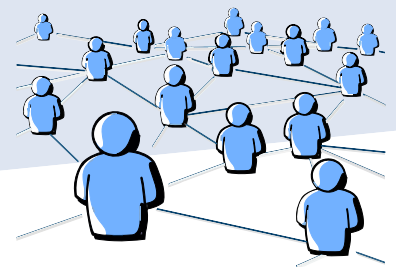
Dashboard



<https://www.figma.com/file/ARbWTHFeA41dbLWZoqnhIX/invoca-project?node-id=0%3A1>

Our UI is meant to emulate the UI of previous Invoca projects so that long-time customers can feel a sense of familiarity with this new project. We are planning to add more pages as we continue the project but for now, we want a very simple UI to display the more complex backend we are working on.

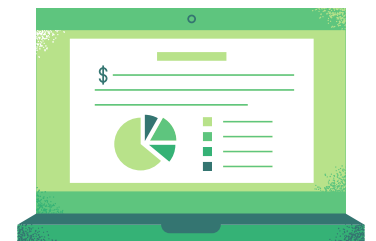
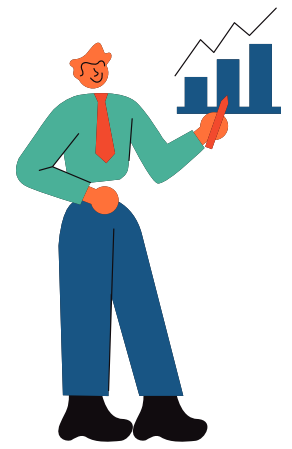




Requirements:

10 user stories

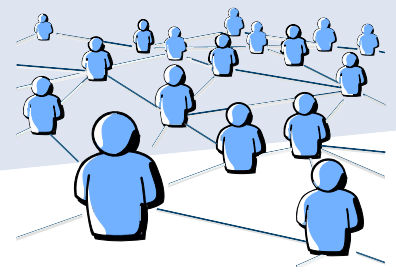
- As a user, when I open/close the app, call summarizations are left intact.
- As a user, when I look at the call summarizations, call summarizations should be neat and not obscure phone numbers/contact names
- As a user, I can click on the call summary to see the entire transcription of the call.
- As a user, I can see a suggested name if the phone number is not in my contacts.
- As a user, I can delete conversations that I no longer deem necessary.



- As a user, I can upload a CSV file of my call logs so that it is easier to view call summaries
- As a user, I can login/register utilizing google OAuth to ensure the user is permitted to access the database/resources
- As a user, I can update the call log if the produced call log was missing some information from the notes the user took during the call
- As a user, I can see the average sentiment of my calls
- As a user, I am assigned a phone number at registration that I can use to receive call summarization

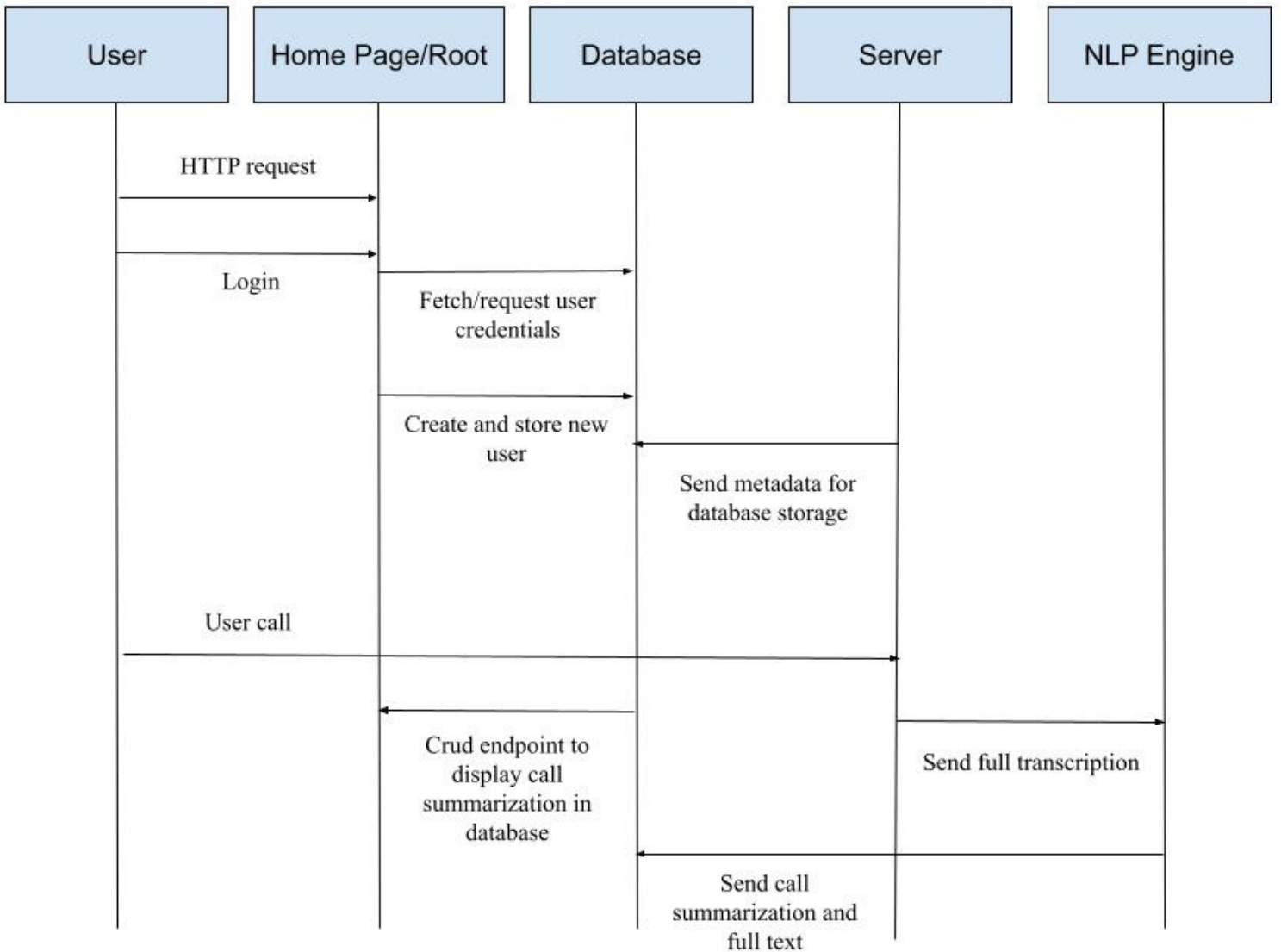


Proto-typing code, tests, metrics (5+ user stories): Github commits/issues



System Models

Contexts, sequences, behavioral/UML, state

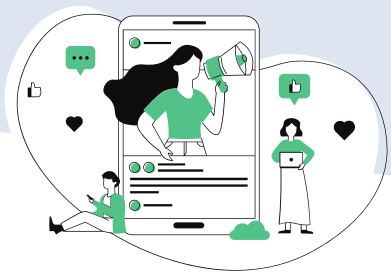


The user sends an HTTP request to access the website and user interface. They login with their credentials using their Google accounts which will be authenticated with OAuth. The credentials are sent to our database to see if they are an existing user or not. If not, they are prompted to create an account. Once logged in, the user arrives at their dashboard, which contains a table of their call logs and summaries.



The data in the MongoDB database is retrieved from the Invoca servers. First, we store the metadata of the calls directly in our database from the servers. This includes information such as the call time, call length, and customer's phone number.





The full transcription of the call is sent to our NLP engine, where we will create a concise summary of the call as well as determine the customer's sentiment. Both the full text of the call and the summarization are stored back in our database. All of the call data, including the summarization, the full transcript, and the metadata, is then retrieved from the database and displayed onto the user's dashboard.



Appendices

- Frontend:
 - React JS
- Backend:
 - Node JS
 - Express JS
- API
 - IBM Watson's cloud-based speech-to-text APIs
 - Invoca's native API for call routing
- Testing / Deployment
 - Heroku
 - Github
- Database:
 - MongoDB
 - Mongoose

