

# **DEVELOPMENT OF ANDROID APPLICATION FOR TOURISM INDUSTRY**

**(Third Eye)**

*Submitted in partial fulfillment of the requirement for the award of the  
degree*

*Of*

**Bachelor of Technology  
In Computer Science Engineering and Information  
Technology**

*By*

**Adriealle Dsouza(15csu013)  
Anchit Bhalla (15csu024)**

**Under the supervision of  
Mrs. Neha Sahu**



**Department of CSE & IT  
The NorthCap University, Gurgaon  
March 2018**

## INDEX

CHAPTER -1	Abstract	2
CHAPTER -1	Existing Problem	2
CHAPTER -2	Introduction	3
	2.1(i) Image Recognition	3
	2.1(ii) Text Recognition	3
	2.1(iii) Virtual Artificial Assistant	3
CHAPTER -3	Working	4
CHAPTER -3	Modules	5
CHAPTER -4	Functional Requirements	7
CHAPTER -5	Progress	11
CHAPTER -6	Gantt Chart	12
CHAPTER -7	Conclusion	12

## **1.Abstract**

Today almost all mobile devices have more processing power than was used to put the first man on the moon. There is a constant the need for technology and minimising the language barrier wherever we go. Built-in network connectivity make it increasingly appealing for mobile phone users to snap pictures of objects and then obtain relevant information about the captured objects.

For example, an Australian tourist sees a landmark (in this case, the Taj Mahal) that he would like to know more about. He uses his mobile phone to capture an image of the landmark and sends the photo to the server. The server performs recognition on the image, then sends the user information about it—for example, the landmark's name, history, and activities. The user can then conduct further information retrieval about the landmark according to his needs.

In today's market applications like Google Goggles,CamFind,etc which use technologies like OCR, Image Recognition and Cloud Computing somewhat help us to provide solutions to own daily requests. But, still there are physical and technological constraints when it comes to using them efficiently.

### **1.1 Existing Problem**

When we go for a vacation to a new place, we don't know the language or the area. For example, an Australian tourist sees a landmark (in this case, the Taj Mahal) that he would like to know more about or visits a nearby restaurant but is unable to order because the menu is in the local language. He is only able to solve this problem by using his smartphones' camera along with a image recognition application. We're trying to automate this.

## **2 Introduction**

### **2.1 Purpose**

The purpose of this document is to capture, in natural language and at a functional level, the description and requirements of Third Eye which is an image recognition application. This is a functional description of those features required to address current tourism industry requirements. A short discussion accompanies each requirement, to add the background and framework necessary to explain the functionality. It also describes nonfunctional requirements and other factors necessary to provide a complete and comprehensive description of the requirements for the software.

### **2.2 Scope**

This software system will be an android application for any tourist wishing to travel well informed and discover new places . More specifically to design and develop a simple and intuitive system which shall cater the tourist needs of any of the selected countries. The application shall provide features to the user with translation, monument queries , navigation and image recognition of foreign objects.

### **2.3 Definitions, Acronyms and Abbreviations**

#### **2.3.1 Image Recognition**

Image recognition, in the context of machine vision, is the ability of software to identify objects, places, people, writing and actions in images.

#### **2.3.2 Text Recognition**

Optical character recognition (also optical character reader, OCR) is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image.

#### **2.3.3 Virtual assistant (artificial intelligence)**

A virtual assistant, also called AI assistant or digital assistant, is an application program that understands natural language voice commands and completes tasks historically performed by a personal assistant or secretary, include taking dictation, reading text or email messages aloud, etc for the user.

Popular virtual assistants currently include Amazon Alexa, Apple's Siri, Google Now and Microsoft's Cortana.

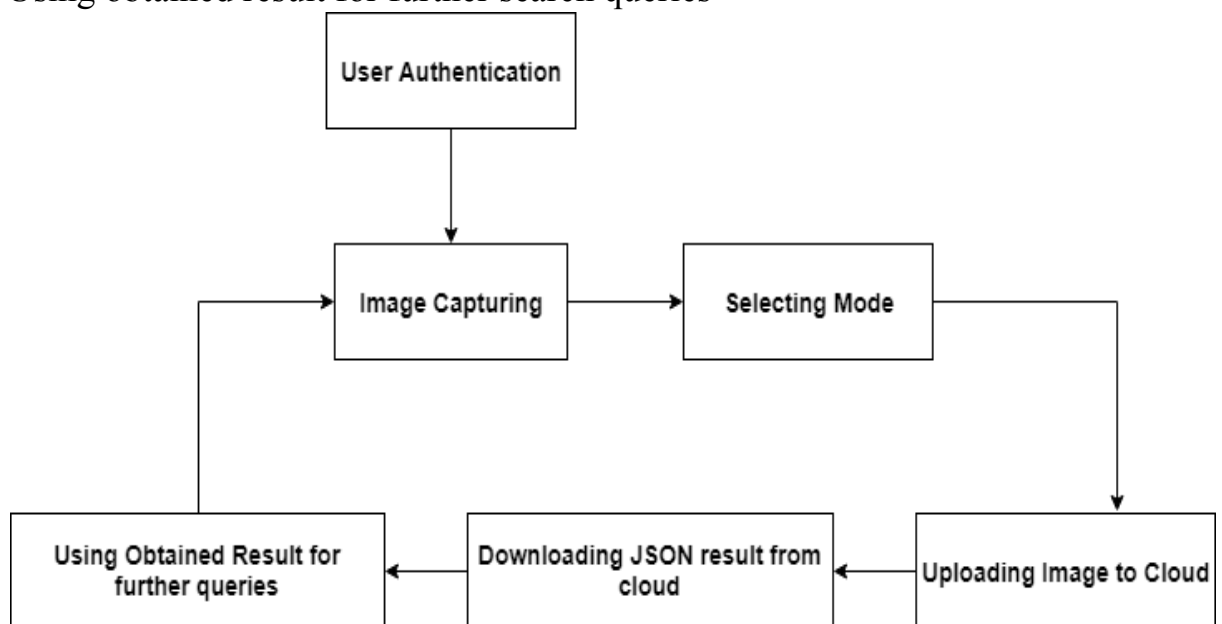
## 2.4 Overview

This is a working document and, as such, is subject to change. In its initial form, it is incomplete by definition, and will require continuing refinement. Requirements may be modified and additional requirements may be added as development progresses and the system description becomes more refined. This information will serve as a framework for the current definition and future evolution of the Open Eye.

## 3. Overall Description

### 3.1 Working

- First user authentication
- Image capturing
- Selecting mode
- Uploading image to cloud
- Downloading JSON result from cloud
- Using obtained result for further search queries



**Fig 1. Working Block Diagram**

## API Usage: Detect Objects in an Image



**Fig 2. Cloud request diagram**

### 3.2 Working Modules

#### 3.2.1. Basic Android App Development:

In this module we are using Android Studio Software to make a basic design of our App. We also implement an authentication mechanism so that data of each user can be saved for further use in research.

#### 3.2.2. Image Capture and Recognition:

First, we are using the camera equipped with the smartphone to capture an image. Then we integrate Google Cloud Vision API in the App.

Google Cloud Vision API enables us to understand the content of an image by encapsulating powerful machine learning models in an easy to use REST API. It quickly classifies images into thousands of categories, detects individual objects and faces within images, and finds and reads printed words contained within images.

#### 3.2.3. Detailed Search Module:

We are using Google Custom Search API to further search about the results from the image. This API lets us retrieve data about any keyword or image.

#### 3.2.4. Text to Speech Module:

For interaction with the App we are using Google's Text to Speech API to convert the text we find in the image to speech.

#### 3.2.5 Artificial Intelligence Assistant:

For better and easy interaction between the user and the App we are going to build a AI Assistant with the help of API.AI and make it understand and learn some simple commands using Neural Networks API. This assistant will recognize some simple commands given by the user and will perform tasks accordingly. For

example, if the user says “search more” the App will search more details about the App.

### **3.3 Product Perspective**

Open Eye is meant to serve as a common platform where tourist tasks and activities can be carried out conveniently. Our goal is to develop a replacement to the existing dependency model with a self learning and application model making it more users friendly and to promote the tourism industry.

#### **3.3.1 System Interface**

Android Application platform will be used where user will provide inputs in the form of Json images or queries. The actual program that will perform the operations is written in JAVA.

#### **3.3.2 User interface**

The new system shall provide a very intuitive and simple interface to the user so that the user can easily navigate through pages and other tourist related activities.

#### **3.3.3 Software Interface**

##### **a)Android Studio**

The official IDE for Android. Android Studio provides the fastest tools for building apps on every type of Android device. World-class code editing, debugging, performance tooling, a flexible build system, and an instant build/deploy system all allow you to focus on building unique and high quality apps.

##### **b)Google Cloud Vision API**

Google Cloud Vision API enables developers to understand the content of an image by encapsulating powerful machine learning models in an easy to use REST API. It quickly classifies images into thousands of categories, detects individual objects and faces within images, and finds and reads printed words contained within images. You can build metadata on your image catalog, moderate offensive content, or enable new marketing scenarios through image sentiment analysis. Analyze images uploaded in the request or integrate with your image storage on Google Cloud Storage.

##### **c)DialogFlow**

Dialogflow is a Google-owned developer of human–computer interaction technologies based on natural language conversations.

The process a Dialogflow agent follows from invocation to fulfillment is similar to someone answering a question, with some liberties taken of course. In the example scenario below, the same question is being asked, but we compare the "human to human" interaction with a conversation with an Dialogflow agent.

### **3.3.4 Authentication Interface**

For Authentication users will be allowed to create custom accounts with email or log in with an existing google account.

### **3.3.5 Database Interface**

For database we'll be using Firebase Realtime Database to provide view history option to users.

### **3.3.6 Memory Constraints**

Memory constraints will come into play when the size of database grows to a considerable size.

### **3.3.7 Operations**

The product shall have operations to protect the database from being corrupted or accidentally altered during a system failure.

## **4. Functional requirements**

This section includes the requirements that specify all the fundamental actions of the software system.

### **4.1 The User**

#### **4.1.1 Functional requirement 1.1**

ID: FR1

TITLE: Download mobile application

DESC: A user should be able to download the mobile application through either an application store or similar service on the mobile phone. The application should be free to download.

RAT: In order for a user to download the mobile application.

DEP: None

#### **4.1.2 Functional requirement 1.2**

ID: FR2

TITLE: Download and notify users of new releases

DESC: When a new/updated version or release of the software is released, the user should check for these manually. The download of the new release should be done through the mobile phone in the same way as downloading the mobile application.

RAT: In order for a user to download a new/updated release.

DEP: FR1



### **4.1.3 Functional requirement 1.3**

ID: FR3

TITLE: User registration-Mobile application

DESC: Given that a user has downloaded the mobile application, then the user should be able to register through the mobile application. The user must provide user-name, password and e-mail address. The user can choose to provide a regularly used phone number.

RAT: In order for a user to register on the mobile application.

### **4.1.4 Functional requirement 1.4**

DEP: FR1

ID: FR4

TITLE: User log-in-Mobile application

DESC: Given that a user has registered, then the user should be able to log in to the mobile application. The log-in information will be stored on the phone and in the future the user should be logged in automatically.

RAT: In order for a user to register on the mobile application.

DEP: FR1, FR3

### **4.1.5 Functional requirement 1.5**

ID: FR5

TITLE: Retrieve password

DESC: Given that a user has registered, then the user should be able to retrieve his/her password by e-mail.

RAT: In order for a user to retrieve his/her password.

DEP: FR1

### **4.1.6 Functional Requirement 1.6**

ID: FR6

TITLE: Select Image

DESC: The user selects the image he/she wants to find out about.

RAT: In order to get an image from the user

DEP: FR1

### **4.1.7 Functional Requirement 1.7**

ID: FR7

TITLE: Search Image

DESC: The image selected will be uploaded on the cloud and searching is done by the system on finding its content.

RAT: In order to search the contents of the image.

DEP: FR1

#### 4.1.8 Functional Requirements 1.8

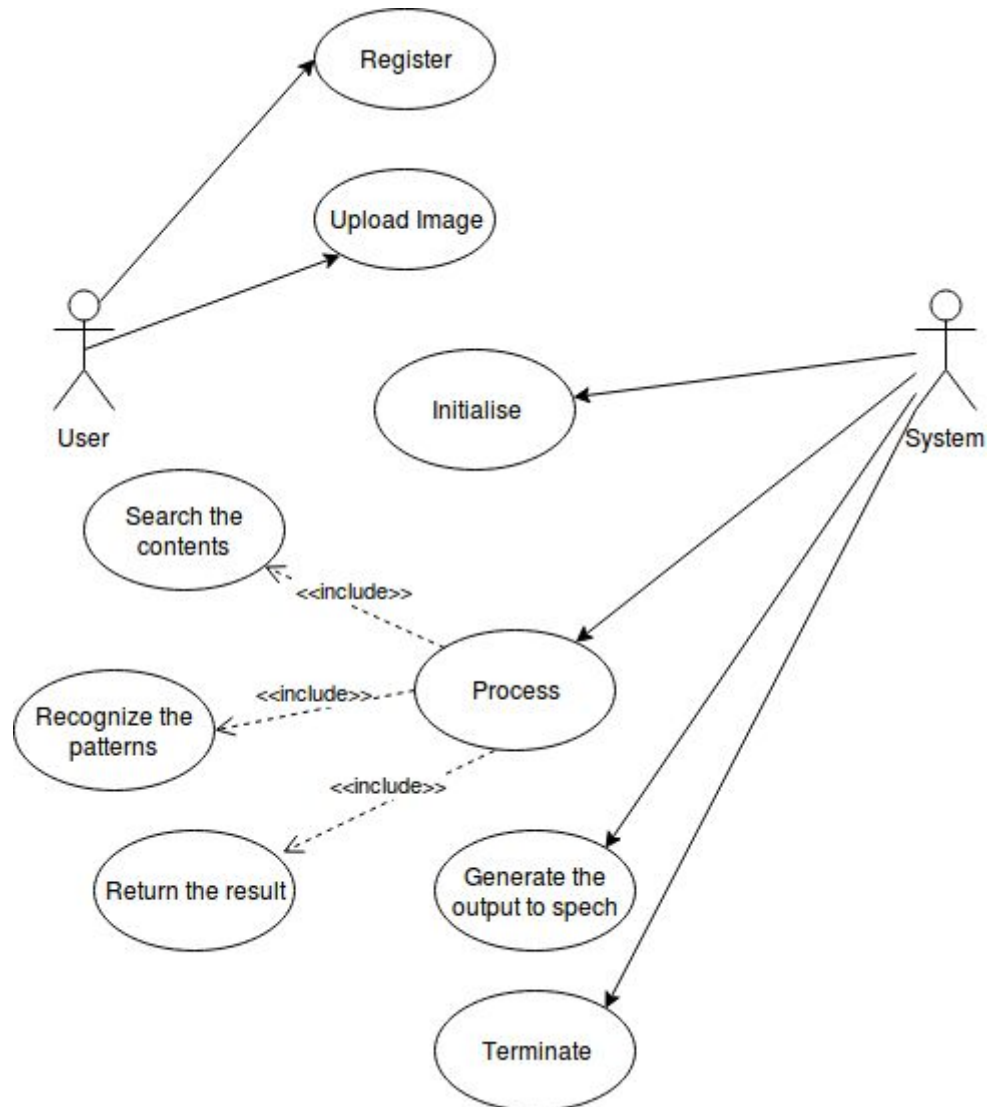
ID: FR8

TITLE: Result to Speech

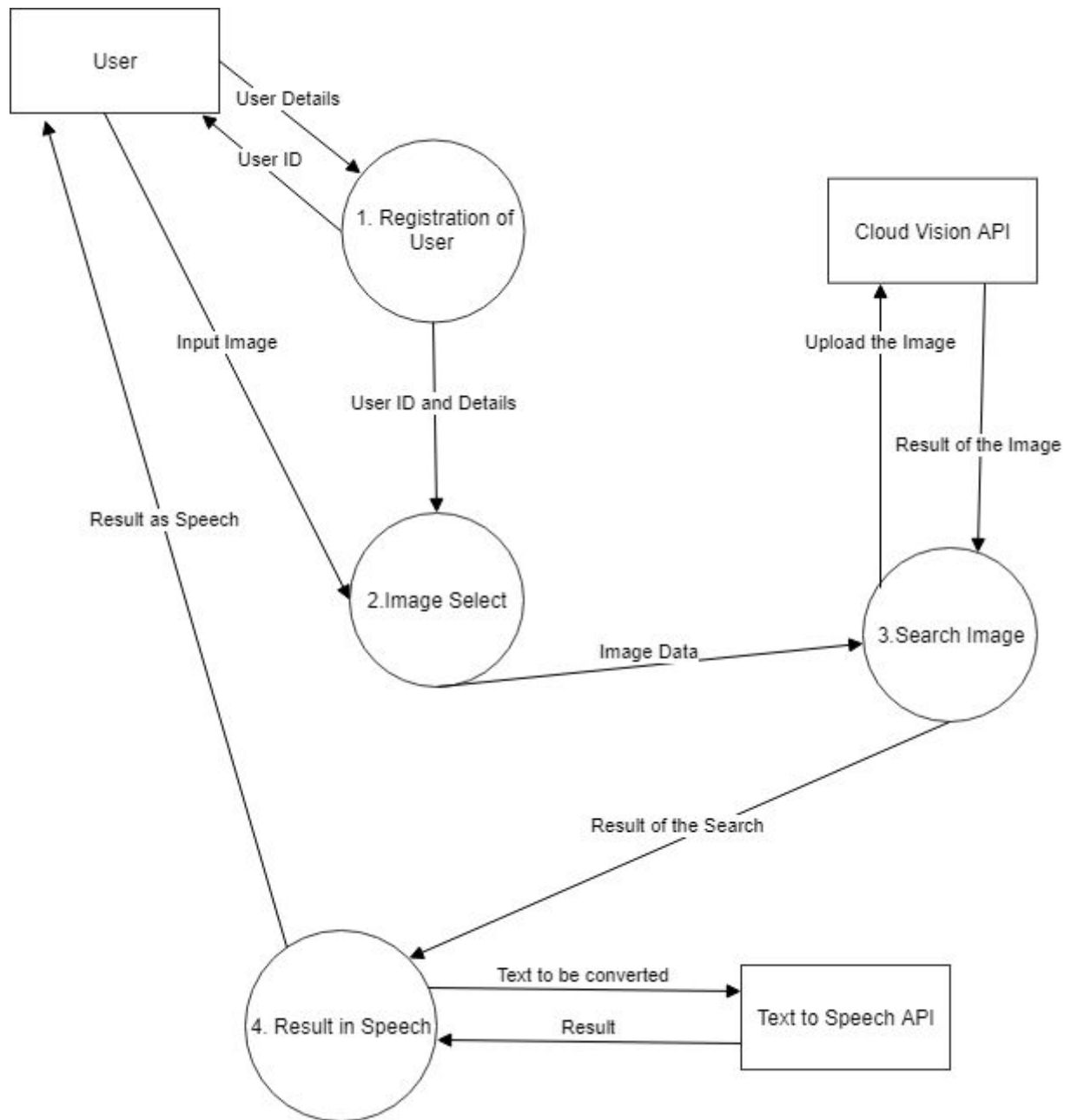
DESC: The result that we get from the image uploaded is now spoken using text to speech conversion.

RAT: In order to tell the result

DEP: FR1 ,FR7

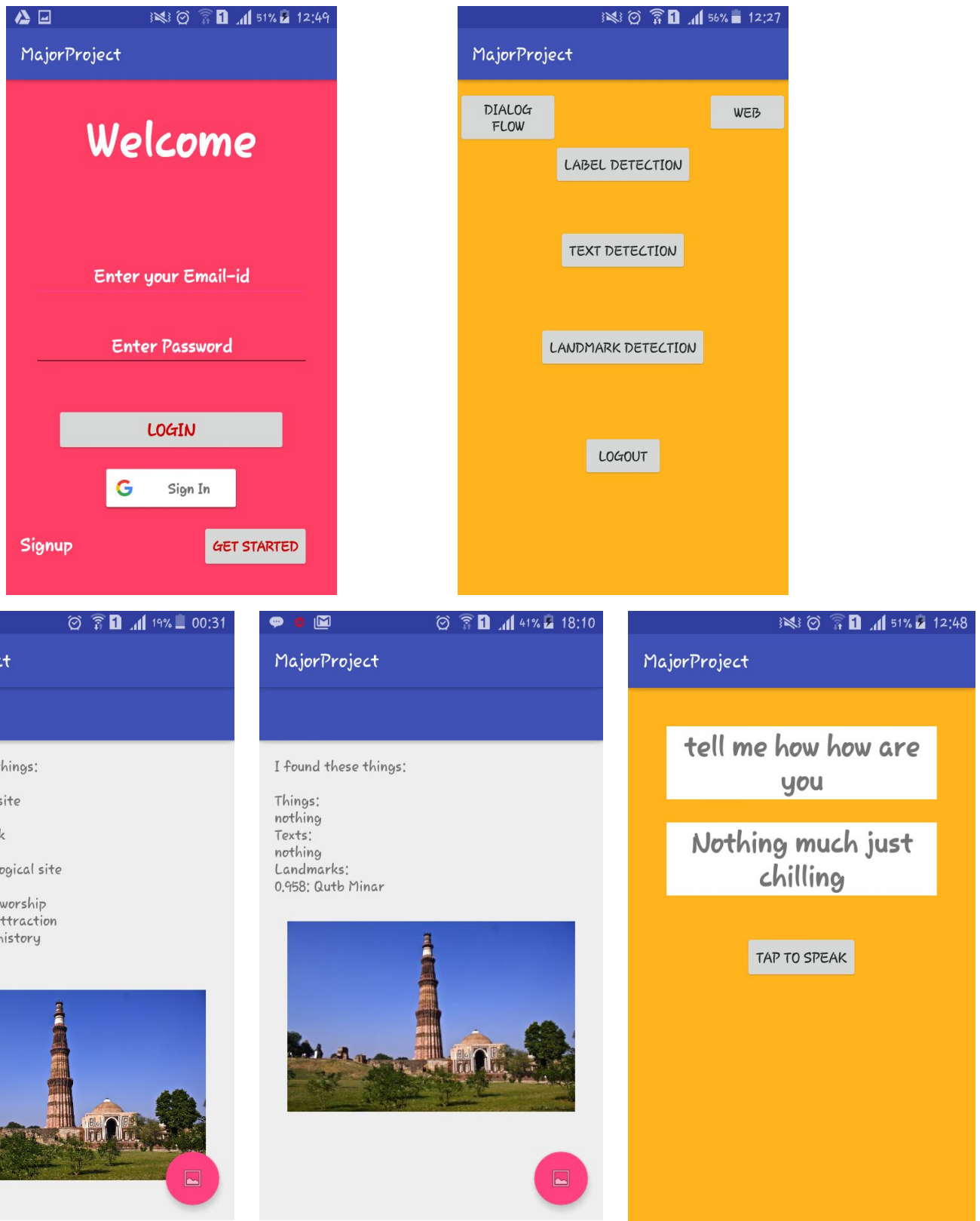


**Fig3. Use Case Diagram**



**Fig 4. DFD Diagram**

## 5. Progress ( Modules 1, 2, 5 : Prototypes)



**Figures 5 (Clockwise): Login page, Stub, Label detection, Landmark detection, DialogFlow**

## 6. Gantt Chart

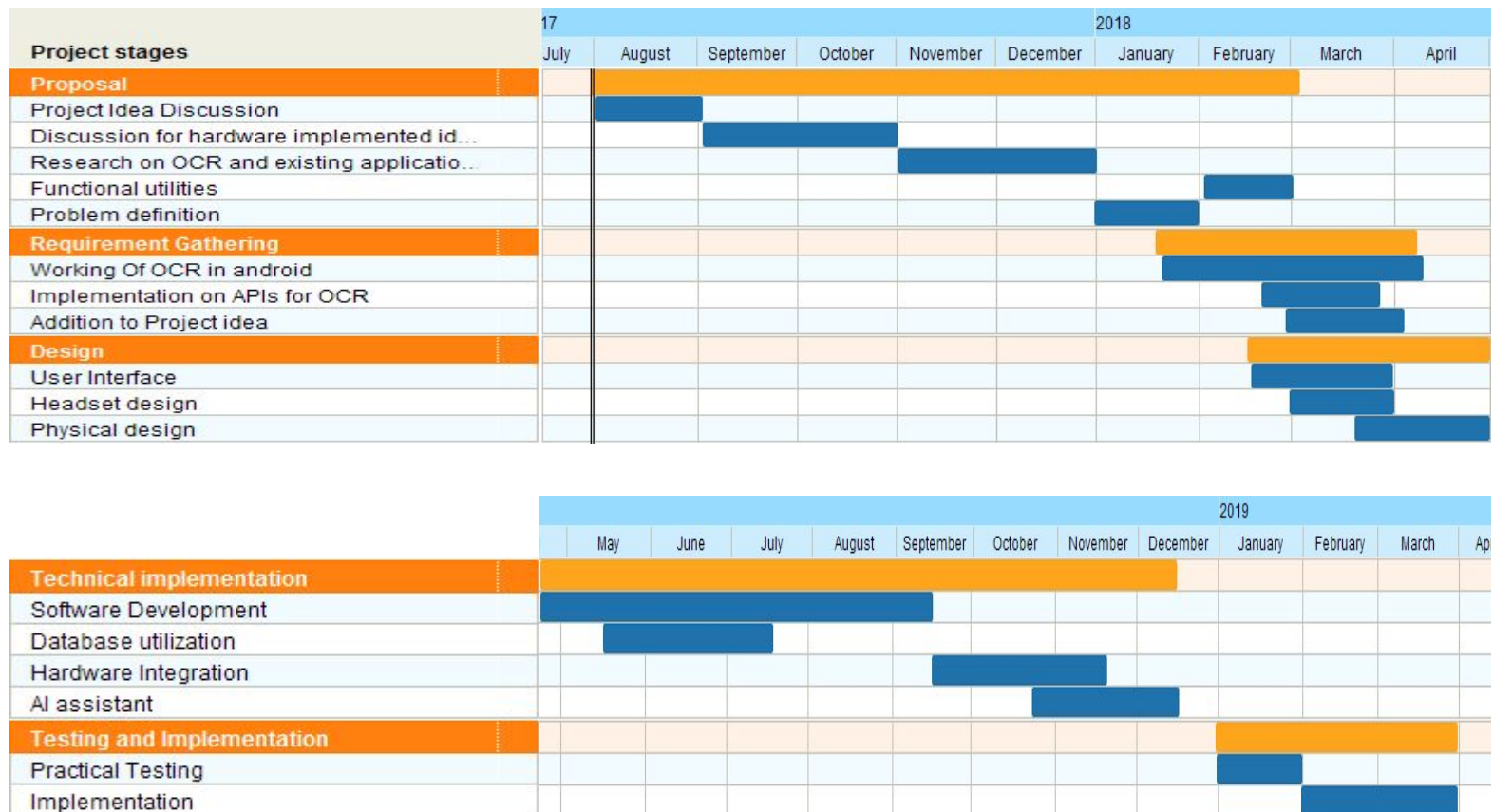


Fig 6. Gantt Chart

## 7. Conclusion

We are trying to solve the problem mentioned before in our App. The App will capture an image and using image recognition software we will be able to scan the image and get data from it. The data then will be searched from the web and the details will be told via speech to the user.

Additional features like External Camera and Assistant are integrated in the App for better interaction and easy to use for the user.