



# AISSMS

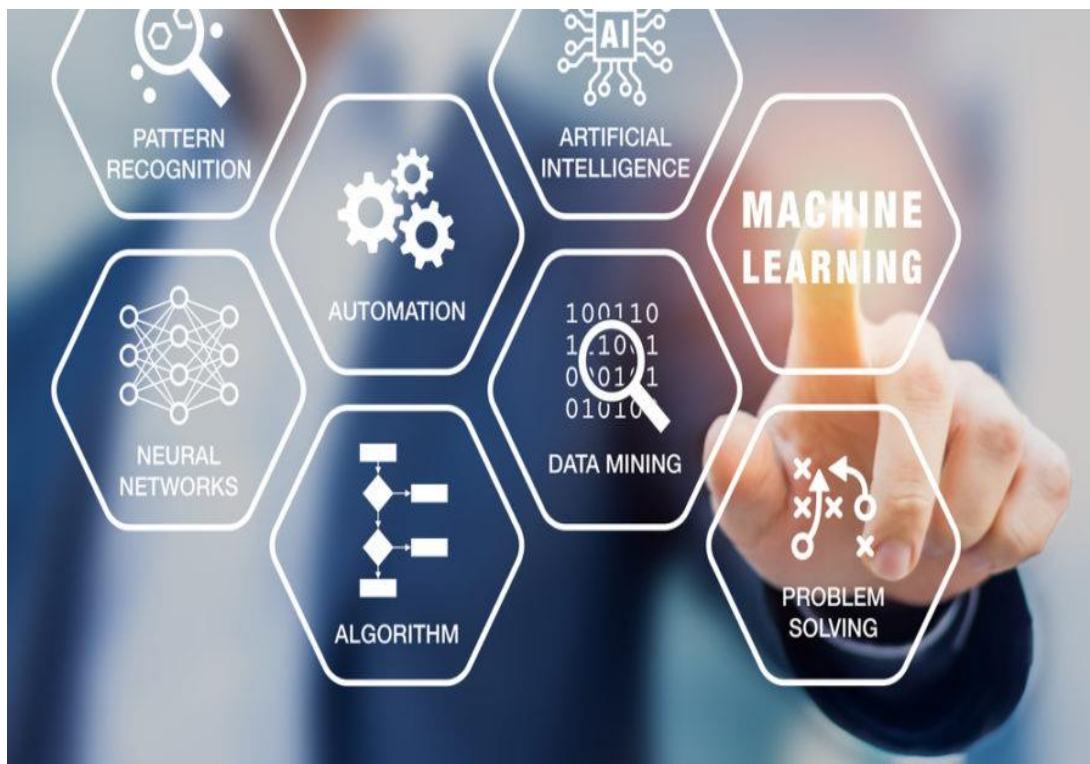
INSTITUTE OF INFORMATION TECHNOLOGY

ADDING VALUE TO ENGINEERING



Department of Computer Engineering

## Technical Magazine 2018-19



## **Objective behind Technical Magazine**

Department of Computer Engineering is very happy and proud to publish technical magazine of year 2018-19. We have gathered technical articles from our students worked as intern in Tech Mahindra IT industry. These articles gives guidelines to students regarding what is expected in IT industry and how various technologies are applied for the projects in IT industry.

Department has set objective to bring technical competency among the students. Department is taking efforts for the same since second year of these students. Department arranges various expert lectures,workshops,industrial visits,learning contents beyond syllabus for the students. All these activities are planned to make students aware of current need of IT industry. Outcome of these efforts is reflected through their final year projects ,placement and admission to higher studies.

We had collected project details from our student who worked in Tech Mahindra as intern . We are sharing experience technical work of these students with our students through this magazine. Our objective behind sharing this information is to motivate students and to create awareness among them about current need in IT industry.

Coordinator

S.P.Pimpalkar

HOD

S.N.Zaware

## Contents

1. About Tech Mahindra Limited
2. About Maker's Lab
3. Student's work experience
4. Dynamic HTML code generation  
Pratiksha Jatti  
Arbaaz Shaikh  
Sujoy Patil Isha Doshi
5. Translation of English words from a German-English bi-lingual text for Natural Language Processing  
  
Gautami Mudaliar  
Tejasvi Gadakh
6. Japanese NLP MeCab Tool  
  
Jincy Biju  
Mrinal Bhangale
7. Counter Raffle Unity3D Software  
Jaymala Pawar  
Smita Muke  
Parv Javheri
8. Augmented Football Stadium Application  
  
Prajwal Chinchmalatpure  
Shashank Hunashimarad
9. Watson Chatbot  
NikulGovardhan Patel  
Shashank V Hunashimarad  
Prajwal Chinchmalatpure  
PratikshaJatti  
MrinalBhangale

## 1. About Tech Mahindra Limited

**Tech Mahindra Limited** is an Indian multinational provider of information technology (IT), networking technology solutions and Business Process Outsourcing (BPO) to various industry verticals and horizontals. Tech Mahindra represents connected world, offering innovative and customer-centric information technology experiences, enabling Enterprises, Associates and the society to Rise™. It is a 4.35 billion company with 112,000 + professionals across 90 countries, helping over 903 global customers including fortune 500 companies. Tech Mahindra's convergent, digital, design experiences, innovation platforms and reusable assets connect across a number of technologies to deliver tangible business value and experiences to our stakeholders. Tech Mahindra is amongst the Fab50 companies in Asia (Forbes 2016 list).

Tech Mahindra is a part of USD 19 billion Mahindra Group that employs more than 2 lakh people in over 100 countries. The Group operates in the key industries that drive economic growth, enjoying a leadership position in tractors, utility vehicles, after market, information technology and vacation ownership.

### *Vision*

Company's vision is to be among the top three leaders in its chosen market segments while fostering innovation and inclusion. It will consistently achieve top quartile growth by contributing to customer's success, by enabling employees to realize their potential and by creating value for all stake holders.

## 2. About Maker's Lab

Tech  
Mahindra



Makers Lab is the Research and Development wing of Tech Mahindra in which Innovators are working on cutting edge technologies in sync and for our customers. The lab was set up in 2014, and has since come up with a Patented digital solutions that are being used either internally at Tech Mahindra or by clients. Our Innovators are continuously working on solution to deliver great customer experiences.

Makers Lab has created solutions like Entellio, X-Retail, Komify, and HIRO – Home Intelligence ROBOT which has won critical

awards and recognition. We develop the entire Enterprise / Business Solution in-house. Makers Lab is currently active in four locations across India viz. Pune, Chennai, Hyderabad and Bengaluru, with Pune housing the largest facility with 40+ Innovators. Our international presence is in Adastral Park (UK), and Dallas (USA). A new Innovation Lab will start in Munich (Germany) in 2018.

We co-innovate with customers to set up Innovation Cells in their premises in a BOT mode.



### VISION

To support digital disruption by empowering students from select forums to participate in Tech Mahindra Makers Lab initiatives.

### MISSION

Helping students build Smart Machines of the future.  
Charting a path of commercially viable products developed using AI, Mobility, IoT, AR/ VR, BI, Robotics, etc and future technologies.

### OBJECTIVE

To closely collaborate with Academia for research and development projects in Digital technologies.

### **3. Student's work experience**

Maker's lab is the research and development area of tech Mahindra, It is under the Leadership of my guide, Mr. Nikhil Malhotra who is the global head. Maker's lab has Branches in Chennai, Bangalore, Hyderabad, Pune and also Dallas in USA and Ipswich in England and with more to come. They specialize in the modern trends in information Technology such as Machine Learning, Artificial Intelligence, Internet of things, Augmented and Virtual Reality.

Although I am under the guidance of Mr. Nikhil Malhotra, I am working under Mr. Saurabh Shaligram, who have helped me in the most Important project.

My mentor also likes my work and is a helpful figure, giving thoughtful advice on Various fields and encouraged me to also take a look at machine learning as it is helpful.

So, in general, Maker's lab is an amazing place to be there to work and it is also a great Place for industrial visits due to the various products they have to showcase.

## 4. Dynamic HTML code generation

Pratiksha Jatti

Arbaaz Shaikh

Sujay Patil

Isha Doshi

The project is based on the concepts of Convolutional Neural Network(Deep Learning).It is a class of deep neural network most commonly applied for analysing visual imagery. A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, RELU layer i.e. activation

### **Task Definition**

The project is aimed at creating a web page in Hyper Text Mark-up Language using Machine Learning.

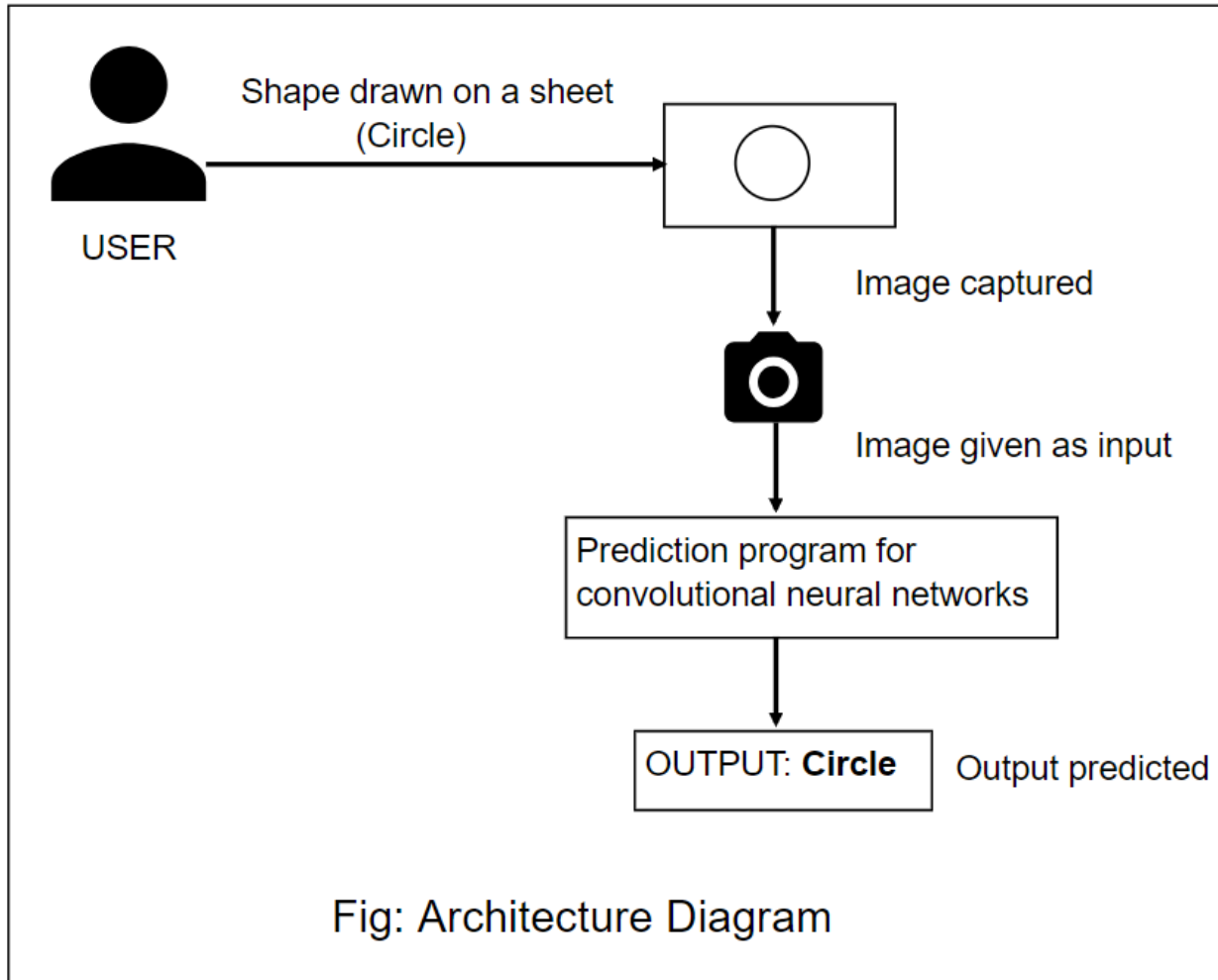
The basic functionality of this project is- The user will draw a basic wire frame structure of a website on a screen/board/white sheet, which would be captured by camera and the HTML code will be generated on the machine.

This project starts with website study which includes studying various websites (social media, ecommerce, educational etc) and understanding the structures and creating 50 wireframes using moqups webapp. Further the project proceeds with dataset Creation which includes deriving basic shapes that are mostly used in website designing. Further the shapes were classified in different categories –basic and advanced. Then dataset of basic shapes(8000 images) were created.(Square, Rectangle, Circle, Radio, Images, Search-box, Checkbox).

### **Approach**

The project could be done with different approaches :

- **Opencv**: This approach was inappropriate because irregular shapes or some distortion in the shapes are not recognized by Opencv.
- **CNN** : This was appropriate approach as it could recognize distorted and irregular shapes with better accuracy than opencv.





## **5. Translation of English words from a German-English bi-lingual text for Natural Language Processing**

Gautami Mudaliar

Tejasvi Gadakh

### **Introduction**

As we know, in today's day almost every application used on the Internet has a chat bot along with it. A chat bot is an artificial intelligence (AI) software that can simulate a conversation (or a chat) with a user in natural language through messaging applications, websites, mobile apps or through the telephone. Chat bots are available in a variety of languages to increase the scope of its usage.

But, as most of the users of such applications are millennials. They tend to use certain English words in the sentences that do not belong to that particular language, which results in the inaccuracy of the results provided by the bot.

This paper is presented for English and German (en-de).

In this particular paper, methods are specified that would carefully filter out the English words. Translation of those words would be done and then would be inserted back into the text sequence in a proper grammatical manner.

This paper presents effective approaches for the translation of English words present in a German sentence. Bi-lingual sentences consisting words of German and English becomes intractable for the machine to understand.

Our embedding models are trained to produce similar representations exclusively for GermanEnglish bilingual sentence pairs.

Hence, with the effective use of the approaches, it would be possible to reconstruct the sentence completely into German. The quality of the resulting embedding is evaluated by comparing the accuracy of the translated text with that of the given input text.

## Problem Statement

To resolve the English interventions in a German-English bi-lingual text and providing a grammatically correct equivalent German sequence for effective Natural Language Processing in chat bots or other messaging applications.

## Background

As the influence of English language is increasing in various languages, it has also found its home in the German language. There are many similarities present in both the languages. For example :The internet revolution started in America and many of the terms have inevitably found their way into German, e.g. *eMail*, *Browser*, *onLineBanking*, *surfen*etc. However, some English computer terms are used at the expense of perfectly good German alternatives. *Chatten* instead of *plaudern*, *downloaden* instead of *herunterladen*, even the word *Computer* itself instead of *Rechner*. Even many advertising fields in Germany use English words or sentences.

Predominantly, machines are trained only for one language exclusively. The problem arises when this blend of languages is fed as an input to the machine. Ambiguity within the machine arises, which causes incompetent results.

To cater to this particular issue, approaches have been specified in this paper that would help resolve the gap.

There are various Natural Language Processing Toolkits available for the aid of programmers to provide assistance with NLP. The prominently used toolkits in the described methods are NLTK and Textblob.

NLTK: Natural Language Toolkit is a platform used to build programs that work with human language data for achieving NLP. It provides easy-to-use interfaces to over 50 corpora and lexical resources such as WordNet, along with a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum. NLTK is a very useful tool to learn the basics of NLP.

TextBlob: TextBlob is a Python(2 and 3) library which is used for simple processing of textual data. It provides simple API for diving into common natural language processing tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation and more. Textblob, being available for a variety of languages; has been principally used for English and German in this paper. The 'textblob-de' is German extension of TextBlob. All directly accessible textblob de classes are initialized with default models for German. The various functions provided by Textblob for NLP are: Tokenization, Noun phrase extraction, POS-TaggingWords inflection and lemmatization, N-grams and Sentiment Analysis.

## **Solution**

This section introduces all the three approaches researched about, along with the outcomes of their analysis. The approaches are listed below:

### **Translation using the translate() function of Textblob**

Given an input bi-lingual text having a combination of German and English words, this method attempts to use the inbuilt translate function of the Textblob toolkit to retrieve a piece of text into a complete, grammatically acceptable German text.

Translating the German-English text directly into German, leads to certain inaccuracies. The function returns the language which occupies majority of the text. Since major part of the extract is identified as German by the machine, the complete text is assumed to be into German. Thus returning the input string as it is.

This can be fixed by using the translate function twice. The primary translate(to='en') function, renders the text into English. This formulation causes the text to be interpreted completely into English, devoid of any German word. Once an equivalent text in English is retrieved, this text is further processed by the translate(to='de') function. Ultimately, this converts the text into the required entirely German sentence, that the machine will be able to interpret.

This particular procedure is the most efficient approach from the ones specified in the paper.

## **Translation using POS Tagging of words**

Loaded with a bi-lingual text of German and English words, this technique uses the NLP features of Textblob. POS-Tagging is the prominent one used in this technique. POS-Tagging is a process of marking up a word in a text (corpus) as corresponding to a particular part of speech, based on both its definition and its context.

Through various observations, it has been analyzed that the English words present in a German sentence, when passed through the `postags()` function; are either tagged as a Foreign Word ('FW') or as a Form of Noun ('NN'). Filtration of these words, will yield a list of all the English words present.

These filtered words are furthermore translated and replaced into the input string. After which, `languagecheck()` package is used that determines and corrects the German sentence grammatically.

The end result would be a grammatically right German sentence. It is an effective process but its time complexity is comparatively much more than the other given approaches.

## **Translation using an English and German Dictionary Corpus**

Another approach for this translation is by using an English Dictionary and a German Dictionary. The dictionary corpora are incorporated into the code. They act as an indicator for checking if a particular word in the sequence is an English word or German word.

If the extract when passed through the check condition, outcomes as an English word, translation would be performed on the word. The equivalent translation of the specific word would be replaced in the string sequence.

The procedure is a functional one with productive result, but is of a bit higher time complexity.

## **6. Japanese NLP :MeCab Tool**

Jincy Biju

Mrinal Bhangale

This report is about NLP (Natural language processing) for Japanese Language and the various tools available for Japanese NLP.

### **Scope:**

Understanding the various NLP tools for Japanese and finding their efficiencies in analyzing the Japanese sentences and also the N-best solution of each.

### **1 Natural Language Processing:**

Natural language processing (NLP) is an area of computer science and artificial intelligence concerned with the interactions between computers and human languages, in particular how to program computers to process and analyze large amounts of natural language data.

### **2 Japanese NLP Tools:**

There are different types of Japanese NLP tools like:

Dictionary Based Tokenizer

It Contains:

MeCab

Chasen

JUMAN

KuroMoji

ML Based Tokenizer

It Contains:

TinySegmenter

Micter

KyTea

### **3 About: Dictionary Based Tokenizer:**

#### **1.MeCab:**

MeCab is an open source morphological analysis engine. It is based on a generalpurpose design that does not depend on language, dictionary, or corpus. Conditional Random Fields (CRF ) is used for parameter estimation, and the performance is improved compared to the Hidden Markov Model adopted by ChaSen . Also, it runs faster than ChaSen and JUMAN on average

Features:

- Dictionary, corpus independent generic design.
- High analysis accuracy based on conditional random field (CRF ).
- Faster than ChaSen .
- Various script language bindings (Perl / ruby / python / java / C #).

Dictionary for MeCab

- IPA dictionary, IPA corpus based on parameter estimation by CRF dictionary.
- JUMAN dictionary, dictionary parameterized by CRF based on Kyoto corpus.
- Unidic dictionary, based on BCCWJ corpus dictionary estimated by parameter on CRF .

#### **2.Chasen:**

ChaSen is a morphological analyzer tool used for analyzing morphemes, developed at Technology. ChaSen was developed from an earlier version of JUMAN. It uses Hidden Markov Models (HMM) to learn character and partsof- speech transitions based on words and parts-of-speech labels in a training corpus. It can learn models for various dictionaries including Juman's dictionary and IPAL. ChaSen was further modified as MeCab with an improved learning algorithm and faster implementation.

#### **3.JUMAN:**

JUMAN is the most widely used morphological analyzer. It was developed at the Kyoto University. It uses a large hand-built dictionary with detailed parts-of-speech.

#### **4.Kuromoji:**

Kuromoji supports standard morphological analysis features such as :

1. Word segmentation - segmenting text into words (morphemes)
2. Part-of-speech tagging - assign word-categories (nouns, verbs, particles, adjectives, etc.)
3. Lemmatization - get dictionary forms for inflected verbs and adjectives

4. Readings - extract readings for kanji

## **ML Based Tokenizer:**

### **1.TinySegmenter:**

TinySegmenter is a compact Japanese tokenizer originally written in JavaScript. It works with Python 2.6 and above.

### **2.KyTea:**

Kyoto Text Analysis Toolkit, is a general toolkit developed for analyzing text, with the main focus on Japanese, Chinese and other languages. The two main features of KyTea are Word Segmentation and Tagging. Word segmentation is separating unsegmented words into smaller units or morphemes. Tagging is parts-of-speech tagging of words.

### **3.Micter:**

Micter is a machine learning based word segmenter. Word splitting is a bit difficult task in agglutinative language like Japanese. Micter could be used for such languages (if you can arrange learning data.) Micter's learning algorithm is Support Vector Machine.

MeCab The reason for choosing MeCab tool is that it is an open source library. MeCab uses discriminative CRF model instead of the generative HMM model which is used by ChaSen. A CRF model can be considered as a generalization of HMM. CRFs outperform HMMs on a number of realworld sequence labeling tasks. So, MeCab is more efficient than ChaSen. We use MeCab to perform a parts-of-speech tagging of the input sentences and then eliminate the non-significant morphemes, such as auxiliary verbs, particles, conjunctions, etc. It is also faster than ChaSen and JUMAN.

## **7.Counter Raffle**

### **Unity3D Software**

**Jaymala Pawar**

**Smita Muke**

**Parv Javheri**

This report is about a game “ Counter Raffle” that has been developed to randomly select a winner from the participants. Secondly, implementation of a 3D game named “Block Runner” using Unity software has been done.

#### **Scope:**

Block Runner is a basic game build under the unity training that is deployable as an android application. Counter Raffle is an application that selects a winner. The winner is been selected randomly via the IMEI code when the counter is set to zero .

#### **TECHNOLOGY :**

##### **1 Augmented Reality**

**AR** is an interactive experience of a real-world environment where the objects that reside in the real-world are "augmented" by computer-generated perceptual information . Augmented reality apps are written in special 3D programs that allow the developer to tie animation or contextual digital information in the computer program to an augmented reality "marker" in the real world. When a computing device's AR app or browser plug-in receives digital information from a known marker, it begins to execute the marker's code and layer the correct image or images.

##### **2 Virtual reality**

**VR** is an interactive computer-generated experience taking place within a simulated environment. It incorporates mainly auditory and visual feedback, but may also allow other types of sensory feedback like haptic. This immersive environment can be similar to the real world or it can be fantastical. Augmented reality systems may also be considered a form of VR that layers virtual information over a live camera feed into a headset or through a smartphone or tablet device giving the user the ability to view three-dimensional images.Current VR technology most commonly uses virtual reality headsets or multiprojected environments, sometimes in combination with physical environments or props, to generate realistic images, sounds and other sensations that simulate a



user's physical presence in a virtual or imaginary environment. A person using virtual reality equipment is able to "look around" the artificial world, move around in it, and interact with virtual features or items. The effect is commonly created by VR headsets consisting of a head-mounted display with a small screen in front of the eyes, but can also be created through specially designed rooms with multiple large screens.

### **3 SOFTWARE USED:**

#### **1 UNITY 3D:**

Unity is a cross-platform game engine developed by Unity Technologies , first announced and released in June 2005 at Apple Inc. Worldwide Developers Conference as an OS X exclusive game engine. As of 2018, the engine has been extended to support 27 platforms. The engine can be used to create both three-dimensional and two-dimensional games as well as simulations for its many platforms. Several major versions of Unity have been released since its launch, with the latest stable version being Unity 2018.3.6. Unity gives users the ability to create games in both 2D and 3D, and the engine offers a primary scripting API in C#, for both the Unity editor in the form of plugins, and games themselves, as well as drag and drop functionality. Prior to C# being the primary programming language used for the engine, it previously supported Boo, which was removed in the Unity release, and a version of JavaScript called UnityScript, which was deprecated in August 2017 after the release of Unity 2017.1 in favor of C#. The engine has support for the following graphics

APIs: Direct3D on Windows and Xbox One; OpenGL on Linux, macOS, and Windows; OpenGL ES on Android and iOS; WebGL on the web; and proprietary APIs on the video game consoles. Additionally, Unity supports the low-level APIs Metal on iOS and macOS and Vulkan on Android, Linux, and Windows, as well as Direct3D 12 on Windows and Xbox One.

## **8. Augmented Football Stadium Application**

**Prajwal Chinchmalatpure**

**Shashank Hunashimarad**

### **Introduction**

Augmented reality has been a hot topic in software development circles for a number of years, but it's getting renewed focus and attention with the release of products like Google Glass. Augmented reality is a technology that works on computer vision-based recognition algorithms to augment sound, video, graphics and other sensor-based inputs on real world objects using the camera of your device. It is a good way to render real world information and present it in an interactive way so that virtual elements become part of the real world.

Augmented reality displays superimpose information in your field of view and can take you into a new world where the real and virtual worlds are tightly coupled. It is not just limited to desktop or mobile devices. As mentioned, Google Glass, a wearable computer with optical head-mounted display, is a perfect example.

A simple augmented reality use case is: a user captures the image of a real-world object, and the underlying platform detects a marker, which triggers it to add a virtual object on top of the real-world image and displays on your camera screen. Augmented reality is a technology that works on computer vision-based recognition algorithms to augment sound, video, graphics and other sensor based inputs on real world objects using the camera of your device. It is a good way to render real world information and present it in an interactive way so that virtual elements become part of the real world.

Augmented reality displays superimpose information in your field of view and

can take you into a new world where the real and virtual worlds are tightly coupled. It is not just limited to desktop or mobile devices. As mentioned, Google Glass, a wearable computer with optical head-mounted display, is a perfect example.

A simple augmented reality use case is: a user captures the image of a real-world object, and the underlying platform detects a marker, which triggers it to add a virtual object on top of the real-world image and displays on your camera screen.

---

### Using augmented reality in football stadiums:

The success of Pokemon Go in the summer of 2016 marked the first real breakthrough for AR into the mainstream: it was a game that changed the face of the technology in the way that Google Glass was supposed to do.

Since then, the technology hasn't been about something that overlays information about the world onto your visual field as much as it's been about using your phone and its camera as a portal door from the real world into the digital one.

When it comes to sport, though, it can provide a new lease of life: freshening the experience of watching most sports, which date from a time before digital technology and haven't been updated very much. Going into a stadium can be one of the few places you can find yourself disconnected from smart technology; in some ways, watching at home can bring you closer to the action than watching in-person. But AR could change the stadium experience completely, making it a hugely different proposition from TV broadcasts whilst still maintaining all the modern conveniences. Today's world is all about 'experiences' we're being told. That has become something of a cliched phrase, but at its root is a pushback against the fact that most people have lived so much of their lives behind the veil of the smartphone; more and more we are tied to it by the need to stay connected. AR has the power to create a stadium which presents a totally different proposition to fans than watching at home, and makes your team's home into what is almost a different world.

So, our primary aim is to basically turn the stadium into a whole new augmented environment, using an app as the controller with which every fan in the stadium can transform their experience.

### **Use cases of the application**

Two primary use cases for the application are as follows:

1. A counter raffle game for all fans present in the stadium.
2. An AR powered food ordering system for convenience during gameplay.

### **Flow of Module 1 (Raffle Game):**

1. All participants are required to scan a poster / banner of match sponsor to participate in the game.
2. As soon as the participant scans the poster, a counter will start running in his android device.
3. A virtual football image will be circulated through all participant devices. (Only a single device will have the image at a given point of time)
4. Participant having the circulated football image at end of counter will be declared as the winner.
5. The device will capture first reaction of the game winner using device's front camera.
6. Gifts and prizes will be awarded to winner of competition by sponsors of the match.

## 9.Watson Chatbot

**NikulGovardhan Patel**

**Shashank V Hunashimarad**

**Prajwal Chinchmalatpure**

**PratikshaJatti**

**MrinalBhangale**

### PROPOSED SYSTEM

In proposed system we can add extra features in existing systems to improve serviceability of bots and quality of bots. We can add features like Text-to-Speech by which chatbot will be able to read the output produced by itself. We can add Speech-to-Text feature in existing chatbot so that user can easily interact with bot. These features will help humans to experience human like response from bot. This features will improve users interest in engaging with bot. We can increase datasets so that bot will be able to give response for any type of query asked by user. We can use various ML algorithms such as NLP for effective retrieval of response from datasets. The bots would be able to navigate to different sites according to user requests.

### INCLUDED COMPONENTS

- **IBM Watson Assistant:** Build, test and deploy a bot or virtual agent across mobile devices, messaging platforms, or even on a physical robot.
- **IBM Watson Discovery:** A cognitive search and content analytics engine for applications to identify patterns, trends, and actionable insights.
- **IBM Watson Natural Language Understanding:** Analyze text to extract meta-data from content such as concepts, entities, keywords, categories, sentiment, emotion, relations, semantic roles, using natural language understanding.
- **IBM Watson Tone Analyzer:** Uses linguistic analysis to detect communication tones in written text.

### MODULES

- Jarvis-bot
- Codebots-plugins
- Obie
- Sourcing bot
- Statsbot

## **APPLICATION AREAS**

As we know Chatbot is a new way of handling user or customers request or feedback, Chatbot can be used in social sites like Government sites, Messaging sites etc. Chatbots can be used as an agent to handle the query of customers in business site of companies. Chatbot can be used as assistant to manage day-to-day activities, reminders, meetings and it should be able to perform some extra tasks ex: when user gives input “call xyz”, so it should perform this task.

