

City University of New York (CUNY)

## CUNY Academic Works

---

Publications and Research

New York City College of Technology

---

2019

### On Using AI Bots for Voice Controlled Augmented Reality Applications

Juan Estrella

*CUNY New York City College of Technology*

[How does access to this work benefit you? Let us know!](#)

More information about this work at: [https://academicworks.cuny.edu/ny\\_pubs/970](https://academicworks.cuny.edu/ny_pubs/970)

Discover additional works at: <https://academicworks.cuny.edu>

---

This work is made publicly available by the City University of New York (CUNY).

Contact: [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu)



# On Using AI Bots for Voice Controlled Augmented Reality Applications

Student: Juan Estrella | Advisor: Benito Mendoza

Computer Engineering Technology Department

## Abstract

Our research project focuses on exploring the integration of Artificial Intelligence (AI) in Augmented Reality (AR) applications. Specifically on using Speech Recognition or Natural Language Processing for controlling virtual AR objects and enhancing the human-computer interaction. We present an empirical study that compares currently available alternatives for creating an AI Bot to implement voice controlled systems. We selected the alternative that meets the criteria of openness, usability, easy integration, and cost.



## Introduction

- AI has the potential to benefit society in the realms of manufacturing, medicine, security, entertainment, marketing, and many others.
- A subfield of AI is Natural Language Processing and Speech Recognition; making computers understand what humans say and mean.
- AR refers to the technologies that superimpose digital content, generated by computers, over the user's view of the real world.
- AR applications for industrial use such as manufacturing, and equipment maintenance, inspection, and repair are in experimental phase.
- There has not been a breakthrough in industrial applications of AR. On one hand, the technology is new, and on the other hand, the software developing tools are limited in scope.
- An AR application could be significantly served by incorporating AI into it.
- However, for developers, it is difficult to find an entryway for incorporating AI into AR apps.

## Discussion

- AI and AR seem ideally suited to one another. In fact, AR relies on AI to be effective (computer vision is a subfield of AI and AR relies on it).
- AI has a vital role to play in the construction of intelligent adaptive interfaces. Object recognition and tracking, and gestural input.
- Eye tracking and voice commands as a means for manipulating the virtual environment, are the close following steps.
- Next step, speech recognition, including classification and language translation.
- The trend of harnessing AR and big data to breed new interesting applications is starting to have a tangible presence. For example, some apps are incorporating real-world object tagging and advanced data visualization.
- We believe AI will enable AR interfaces to become truly multidimensional; this will generate a whole new layer of perception.

## Our Work

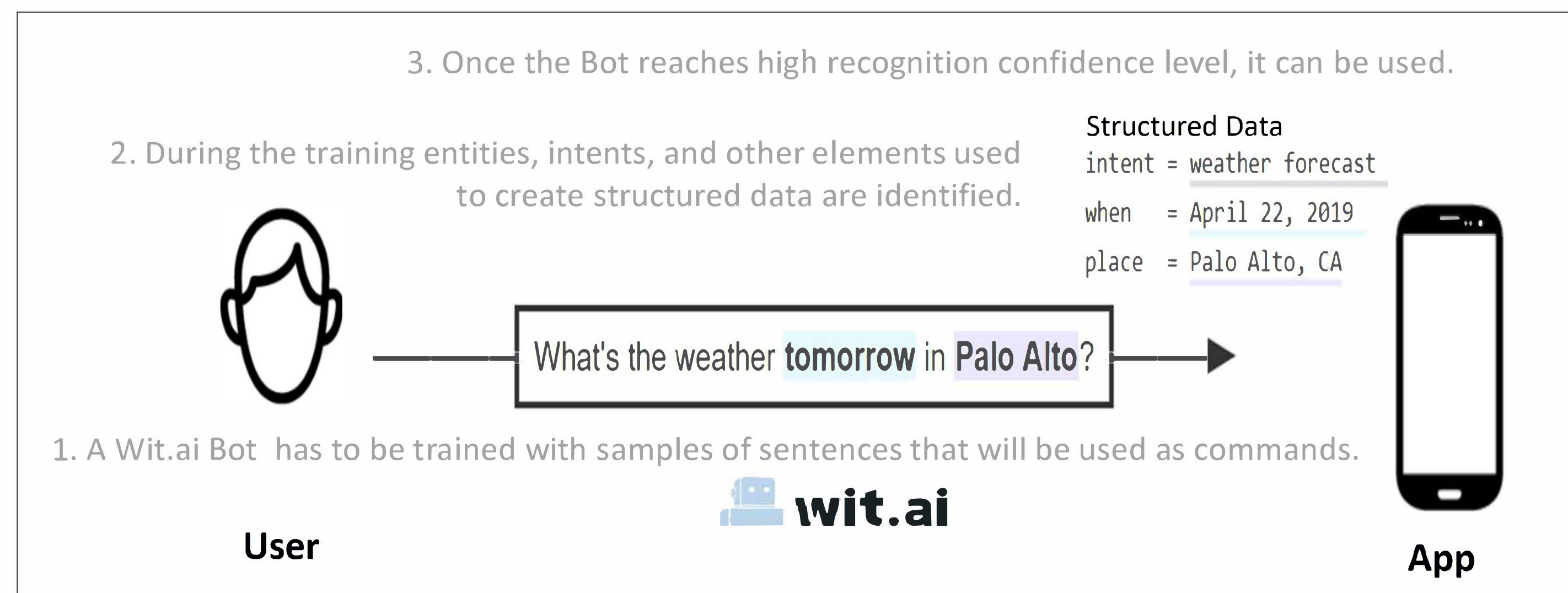
- We present an AR application in which an online AI Bot process user's commands to control the behavior of cars in a virtual showroom.
- The technology in this app can be used on a wide range of Internet Of Things apps such as monitoring systems, medical, marketing and advertisement, entertainment, etc.
- In manufacturing settings, this technology has the potential to increase work performance; workers could command machines to do several tasks while they work on something else.
- We explored different alternatives for implementing the AI Bot: Google's Cloud Speech to Text, Microsoft Azure, Clarifai, and IBM's Watson API. However, they all come at a high cost to use for the general public. We selected Wit.ai, a free and straightforward cloud service owned by Facebook.

## Materials and Methods

To build this app we integrated the following technologies:

- Unity: A Video Game Engine to create the virtual elements.
- Vuforia: An AR software development kit (SDK) for mobile devices.
- Wit.ai: A natural language API and cloud service capable of turning spoken sentences into structured data.

With the combination of these three applications, we can deploy this voice controlled app on Apple iOS, Android, and Windows AR Devices such as the HoloLens



## Results and Conclusions

- Our App responds to commands such as "Open the Drivers Door," "Open the Hood," "Change Color to Red," "Start Engine," or "Start Video,"
- The integration of Wit.ai with Unity consists of a script that triggers specific animations based on the structured data returned by the AI Bot, as a result of converting speech sentences to structured text.
- Wit.ai, is easy to integrate with Unity and Vuforia. The combination is robust.
- The disadvantage is that Wit.ai requires an Internet connection.



## References

- Diego Herrera. "Discover Voice Controlled AR Apps|Unity & A Cloud Based AI." Udemy, 2017.
- Nahal Norouzi, Gerd Bruder, Brandon Belna, Stefanie Mutter, Damla Turgut, and Greg Welch. "A Systematic Review of the Convergence of Augmented Reality, Intelligent Virtual Agents, and the Internet of Things." Book chapter in Artificial Intelligence in IoT, Fadi Al-Turjman Editor. Springer International Publishing, 2019

## Acknowledgements

- Undergraduate Research Program
- CUNY Research Scholars Program
- GRTI 20 Grant, Project 4 "The AREngEd Project: Augmented Reality for Engineering Education"