

Geoffrey Shmigelsky

---

GeorgiaTech, Educational Technology, Fall 2016

Project Proposal

Educational Cognitive Agent in Virtual Reality

## **Abstract**

The goal of this project is to create an Educational Virtual Reality Cognitive Agent (agent) for the purpose of teaching a short class. The agent will serve both as the instructor and the tutor. A Chatbot Engine (chatbot) will power the cognitive component based on Wit.AI, API.ai or IBM Watson Bluemix. Material for the class will be parsed from Wikipedia and imported into the chatbot. Wikipedia's table of contents and paragraphs will make up the course lecture, while facts provide the Question / Answer structure. The topic will be the Serengeti (content), but in theory, any topic will work. Unity3D will provide the models, sounds, lighting, materials, physics and other tools to create the Virtual Reality (VR) experience. The chatbot will render in VR either as a person or some other avatar. A student will wear an HTC Vive Virtual Reality headset, communicate to the VR tutor using voice, and experience an interactive lecture.

*Keywords:* virtual reality, cognitive tutor, Wikipedia, chatbot, Unity 3D, HTC Vive

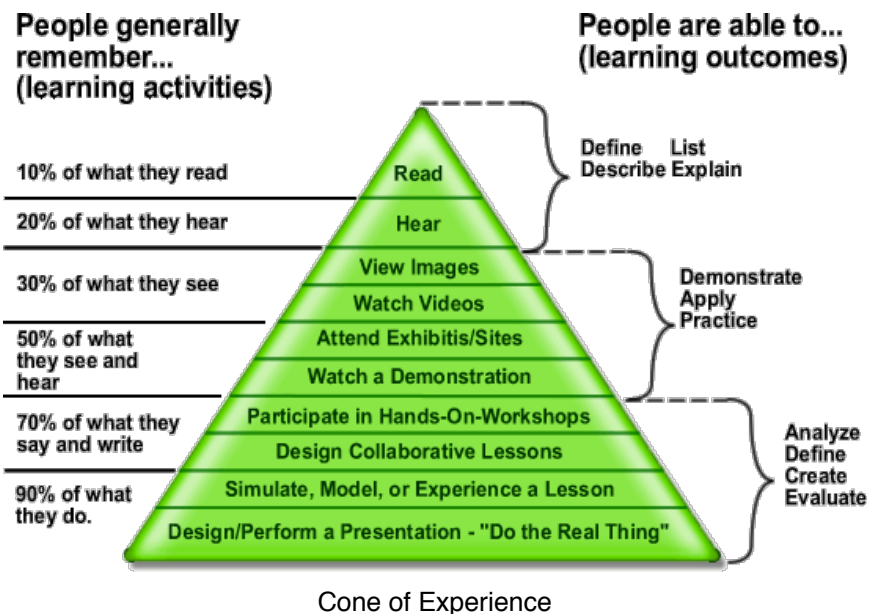
## Educational Cognitive Agent in Virtual Reality

**Problem Description**

According to Edgar Dale and his Cone of Experience, reading and viewing provide a recall rate of 10-30%. The challenge with Wikipedia's 5 million+ articles is that they have a structure of text and images. Users experience educational content by reading it. Considering the time and energy required to learn, I consider a 10-30% retention to be a sub-optimal educational experience. The question is;

*How to elevate educational content into a 90% retention rate?*

Youtube and Khan's Academy have potentially increased the learning retention rate to the 50% range with the use of video. In 2011, Clive Thompson's article [How Khan Academy is Changing the Rules of Education](#), he highlights the positive impact on-demand Internet based video education.



In Davis's and Summer's 2014 paper titled "[Applying Dale's Cone of Experience to Increase Learning and Retention](#)", the role of experiential education in learning outcomes is clear; both traditional and non-traditional students retain far more skills. More so those skills are successfully transferred to the workplace if the students experience the lesson.

## My Solution

In 2016, Bill Gates in an article in The Verge, explains how education could be customized to meet the needs of each student. Gates sees a future where each student has an AI mentor/tutor proactively helping every step of the way.

*“The idea that you could talk to a [virtual] advisor that would understand different misconceptions and arbitrary linguistics around it, that’ll certainly come in the next decade.”*

In 1984, Benjamin Bloom quantified the impact of tutoring in the paper “The Two Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One tutoring.” He identified that tutor can elevate a student 2 standard deviations. Effectively taking a student from the middle of the pack to the 98th percentile.

*I propose a learning environment where the student will experience the benefits of a one-on-one instructor/tutor and experiential education in Virtual Reality.*

Specifically, I propose to train a chatbot to act as a cognitive agent. Using Wikipedia as a knowledge source, the chatbot can be both an instructor and a tutor. It can give lectures as a professor would, gesturing and moving like a human. Likewise, it can answer questions about a topic in a conversational manner. The student will interact with the agent using both Voice to Text/Text to Speech and the HTC hand controllers.

The Wikipedia topic will be the Serengeti and related pages. Virtual Reality is to be created and managed by Unity3D. The student will put on the HTC Vive headset and appear to be standing on the Serengeti plains of Africa. The experience will be immersive, with weather effects, a rising sun, dust, and trees swaying in the wind. Animals such as Zebras or Elephants will come into and out of the scene as required by the chatbot. The student will be able to walk around the scene, peering over bushes, under rocks or in the water.

## Comparable Projects

There are two platforms to consider when discussing alternative projects; Chatbots and Virtual Reality:

### Chatbots



The Math Homework chatbot provides high school students with an interactive bot to cover algebra and pre-algebra topics. It is interactive with images and text. Access is on the Kik platform. It demonstrates today's chatbot technology in education.

### Virtual Reality

VR education is starting to take off with new projects every week. A notable one is LifeLiqe, which uses 3D models to educate in VR. The second is the impressive The Body VR, shown right. It is an immersive VR experience where you learn about the body by exploring it. It is a cutting edge example of how VR can be used to teach.



### Comparison

The challenge with a chatbot is it lacks engagement. With an interface of text and images, but cannot be any more interactive. It is limited to the functionality of a messenger interface like Facebook Messenger or Slack.

The challenge with VR is the content tends to be scripted. The content to date is sequential with a start and a finish. There is no way for a student to stop and ask questions.

To my knowledge, no one has combined these two technologies into one solution for education.

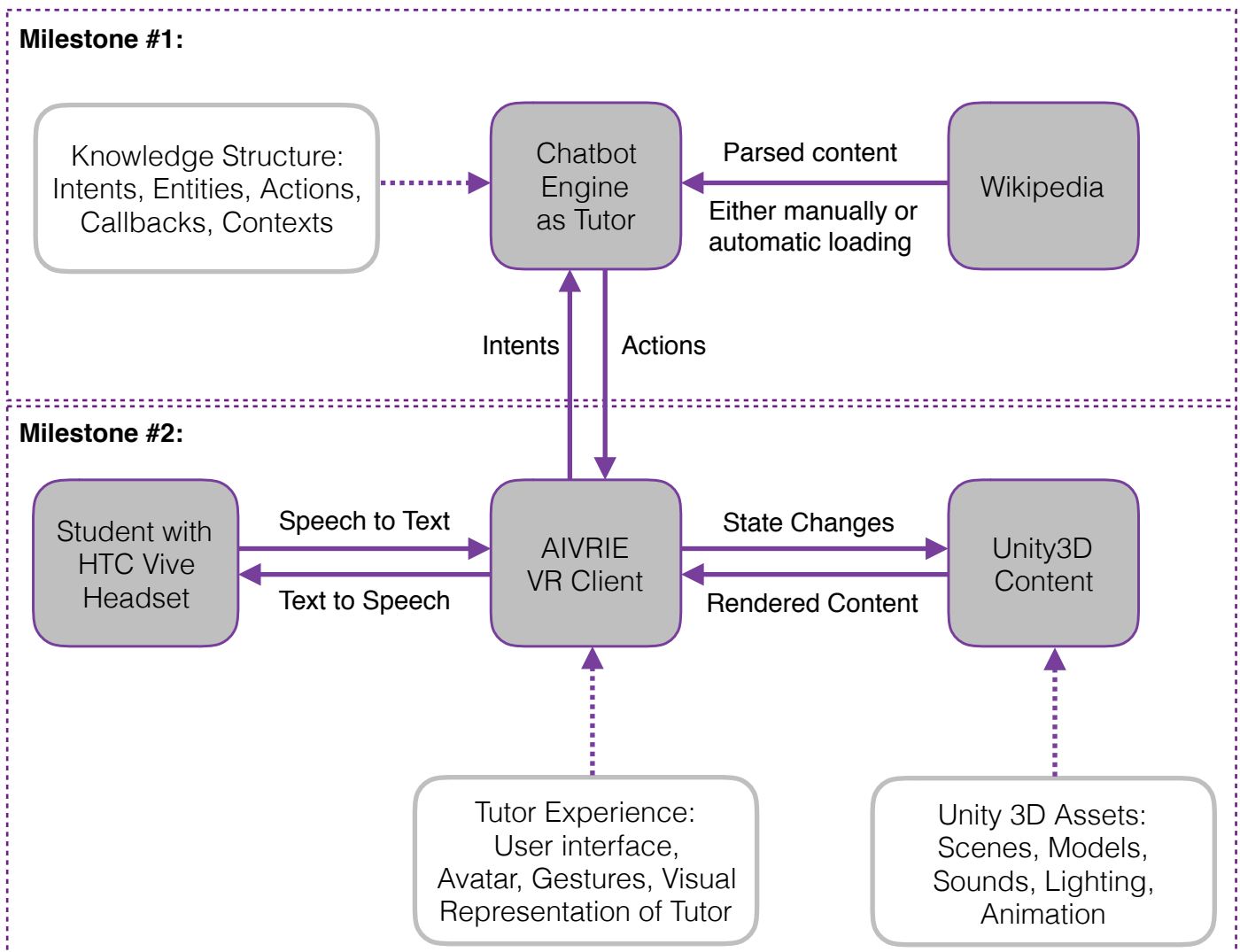
## Implementation Plan

### Milestone #1

For the first milestone, the deliverable is a functional chatbot trained on Wikipedia content for the Serengeti. The chatbot will be testable using either FaceBook Messenger, Slack or Kik. It will include lectures and a question/answer session.

### Milestone #2

For the second milestone, I would propose either creating a YouTube video of the VR experience or a web-based version of the VR world for testing. A web based experience is preferable, but it doesn't always work well with complex scenes or third party libraries (like chatbots). If a web version is not viable, a YouTube video of the experience will be submitted.



## Final Deliverable

The final deliverable will be a 3 or 4 minute video demonstrating the VR experience with the HTC Vive. A polished presentation with cutscenes, background music, and HD format.

Similar to the video I submitted for the class [Computational Photography](#).

## Schedule

The weekly schedule is as follows, covering nine weeks:

Schedule	Tasks
Week 1	Final decision between chatbot engines, Wit.ai, API.ai, or IBM Bluemix.
Week 2	Manually download content into Chatbot from Wikipedia
Week 3, MileStone #1	Connect Chatbot to Messenger, automate wikipedia import into Chatbot
Week 4	Wire up VR with Chatbot, enable speech to text and text to voice
Week 5	Generate first few scenes in Unity - trees, weather, water, animals
Week 6, MileStone #2	Provide either sample video of experience or WebGL demonstration of level
Week 7	Generate Aivrie avatar with basic animations
Week 8	Add music, sound effects, menus, polish the experience.
Week 9	Write Paper, create final presentation video

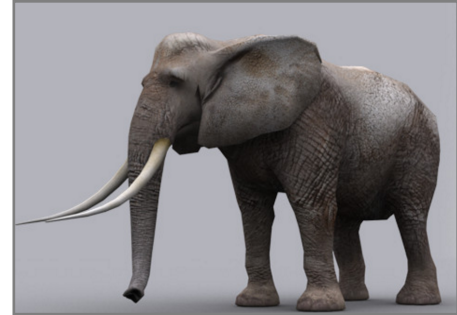
## Content Description

I will focus on content for the Serengeti. Specifically with the article [Serengeti National Park](#) and related Wikipedia pages; such as elephants, zebras, or weather patterns.

In the image to the right, sourced from Wikipedia Commons, is an image of a Serengeti watering hole. Notable content includes trees, grasslands, a clear sky and a herd of grazing elephants. In Virtual Reality I will recreate scenes similar to this one. With trees waving in the breeze, water reflecting sunlight, grasslands with patches of dirt.



Elephants will be life size and like like. Wearing an HTC Vive, it will be possible to walk, touch, and interact with the elephants. Elephant herds will be driven by an AI to graze, flee or fight. To the right, one elephant model from the asset library in Unity3D. Note the level of detail and realism.



### Tools Used

There are two primary tools to bring this application together. First is the chatbot engine, the second is Unity3D.

ToolSet	Source
Chatbot	One of <a href="#">API.ai</a> , <a href="#">Wit.ai</a> , or IBM Watson <a href="#">Bluemix</a>
3D Game Engine	<a href="#">Unity3D</a>
PC Hardware	Windows 10, Nvidia <a href="#">GTX 1080</a> GPU
VR Gear	<a href="#">HTC Vive</a> by Valve
Speech to Text	<a href="#">Google Speech Api</a> or built into Chatbot, depending on accuracy
Text to Speech	<a href="#">Ispeech.org</a> , IVONA, or IBM Watson
Wikipedia Parser	Python Library: <a href="#">Wikipedia</a>
AIVRIE VR Avatar	To be defined, many options
World Creator, StateMachine	<a href="#">GAIA</a> , <a href="#">Playmaker</a>
Virtual Reality Toolset	<a href="#">Steam VR</a> , <a href="#">Lab Rendered</a> , <a href="#">VRTK</a>
Water, Weather, Sky	<a href="#">Aguas</a> , <a href="#">UniStorm</a> , <a href="#">Azure Sky</a>
User Interface	<a href="#">Curved User Interface</a>
Models	<a href="#">African Big Pack</a> , <a href="#">Nature Starter Kit</a>

Additional Unity packages will likely be used, for both development and finalized content. Many asset packages are drag and drop for installation in VR. For example, it is possible to add human eye tracking to the avatar. When the instructor is speaking to the student, it would appear the avatar is look and talking directly with the user.

### **FallBack Plan and Technical Risk**

There is significant technical risk for this project. With so many moving parts, everything has to come together to provide a cohesive education experience.

#### **ChatBots**

The chatbot libraries of Wit.ai or API.ai have not been pushed to this extent for education. They will likely have hundreds of rules. Without stress testing the system, it is not possible to determine how they will handle the project.

Likewise, the engine providers could change policies before the project is complete, impacting the final result. For example, Google purchased API two weeks ago, this week Samsung acquired a chatbot competitor named Viv.

In the event that the chatbots engines cannot handle the necessary level of complexity, the project will have to be scaled back, pivoted or a different chatbot approach used.

#### **Voice to Text and Text to Speech**

The human to computer audio component needs to operate in real time, in a streaming manner. In the event that either of these systems is not suffice, similar functionality will have to be programmed into the hand controllers. Text pops would need to be used versus voice.

#### **Unity**

VR development in Unity is a relatively new technology and it is not without its issues. Although this technology is quickly improving, it is not as seamless as it could be.

Various packages in Unity may not migrate well to Virtual Reality. VR requires a frame rate of 90 fps, which is not sustainable for sophisticated meshes. Each scene is limited to about 2 million vertices. If any Unity package cannot deliver the required results, alternatives will need to be found, but this will take time. Over two dozen packages are to be deployed in the final product; there are many moving parts.

### **Wikipedia Parser**

It is my intention to automate the loading of Wikipedia into the Chatbot engine via a python script. Initial testing of the API has generated some concerns as it appears not all data is being captured. Alternatives may be required, or the data may need to be loaded manually. The batch loading mechanism in the Chatbot engines is untested. It may not offer sufficient control to structure a classroom.

### **Team Members**

I will be the only Team Member, it is a solo project. I have 5 years experience building with Unity 3D and own a library of 1000+ asset packages. I am also fluent in Python and Machine Learning libraries such as TensorFlow.

### **Project FallBack**

My biggest concern and unknown is the chatbot training. Without testing it in detail, there is simply no way to determine if the project can scale to the size of a few Wikipedia pages or more. If the chatbots are not sophisticated enough to handle an online class, I would propose pivot the project into training the chatbots in an automated fashion for education. If I went this route, I would create a python program to parse Wikipedia content and upload it both into Wit.ai and API.ai. Possibly turning it into a web-based service.

It is clear that likely both Google and Facebook will be a significant force in this domain, providing a tool that can adapt content to either at scale would be invaluable to many companies and stake holders. It would be the first step in creating cognitive tutors in an automated fashion.

### **Conclusion**

This project is aggressive and ambitious. I think it is doable in the time frame, but it may exceed the 100 hour budget. My ultimate goal is to build a product that can be released in the HTC Vive app stores and Steam VR. As per the discussion in the forums, the instructor explained any content created is own by the student. As a result, I for see no issues in releasing the final product to the public as a commercial endeavour.

### Referencesx

Benjamin Bloom. (1984) The Two Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One tutoring., *American Educational Research Association.*

Davis B, Summers M. (2014). Applying Dale's Cone of Experience to increase learning and retention. *QScience Proceedings (Engineering Leaders Conference)*

Gates, Bill. (Apr 25, 2016) Educational Software with Artificial Intelligence.  
*The Verve*

Thompson, Clive. (2011). How Kkan Academy is Changing the Rules of Education. *Wired Digital.*