

# Supporting Activities of Daily Life for Seniors via Ai Assistant

Research by: Steven Howell | Mentor: Guoliang Fan | Institution: OSU Visual Computing and Image Processing Lab



## Goal

The goal of this project was to improve human to machine communication with individuals who suffer from a decreased mental function as a result of dementia. The first step was to understand complex and fuzzy inputs made by my user and route those to assistants actions. The second step was to make communication more fluid however that proved to be more difficult than first anticipated and with the creation of chatgpt dozens of new methods have been opened that were not available at the proposal stage of this project. Therefore the first part of the project was focused on and mostly accomplished.

## Methods

To create the NLP model I used SparkNLP to fine tune a new model based off a pre trained transformer. I first trained an (Named Entity Recognition) NER model to classify individual words into a few categories. These are requests, location, actions, objects and O. These are to more accurately determine actions and connect behaviors with the intended environment. The second part is sentence classification sentences are classified into activity requests, location requests, do actions, voice controls, camera controls, shut down, and navigation. Activity requests links of the location and informs the user of the activities at the location. Do action checks the NER result for an object and then plays a video for the object if its available. Navigation uses the NER location data and the current location information to give navigation instructions to the requested location. Video controls manipulate the video to pause play rewind ect.

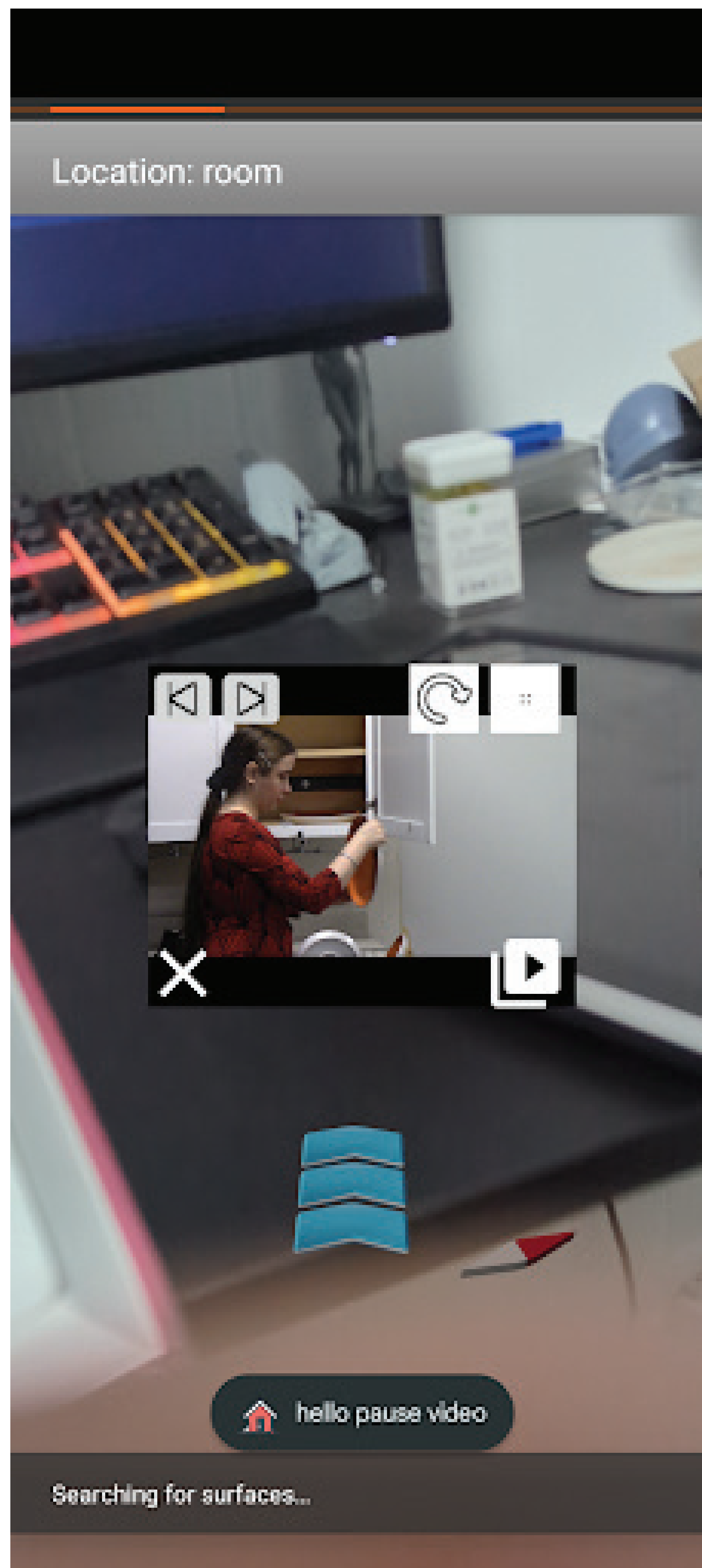
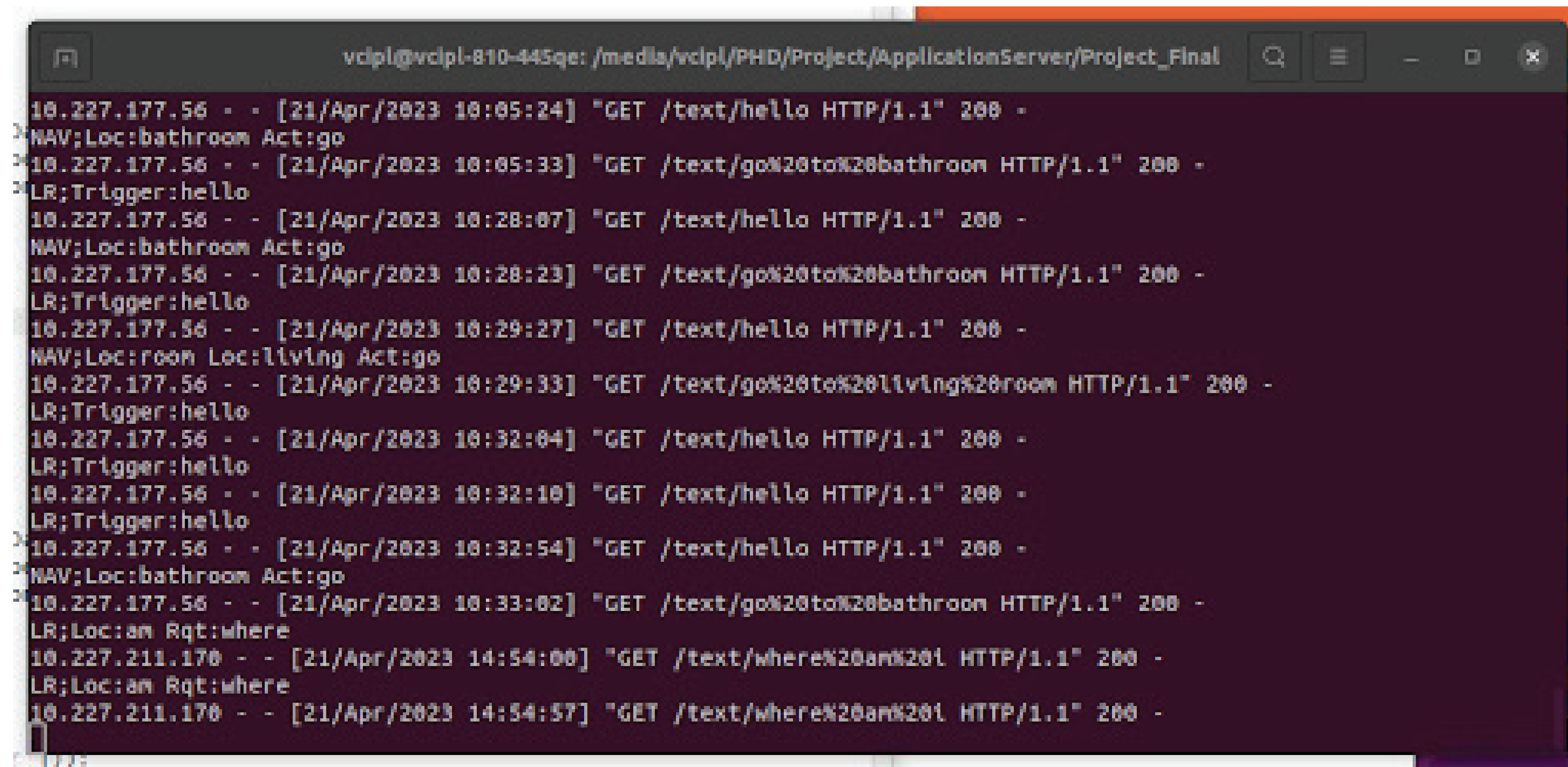


Figure 2. Flow Chart of the Ai's input and output process



## NLP

Currently, we can categorize sentences into tasks that are linked to the functions on the app and rapidly narrow down the behavior that should be performed. It also improves the app's ability to deal with fuzzy sentences from dementia patients. Then, each word is classified and assigned a variable that is used to perform tasks. For example, if you say take me to the kitchen the sentence is classified as a navigation task, and the kitchen is classified as a location which is then used to process the request.

## Actions

**Activity Request** -Then the current detected location is pulled and the activities are read out from the database. If there is only one result it asks if you want to play the video. The trigger word is overridden and the user can speak directly to say yes or no. If pos is detected the video is played. If neg is detected the video is not played.  
**Location request**- if the current location is given.  
**Shut down**- the app shuts down and returns to the home menu.  
**Pause/resume/rewind/fast forward Video**- If an act is detected the act is compared to the predefined control keywords and then passes a command to the video.  
**Activity Call** - If an obj is detected then the object is pulled and passed to the database if the obj is in the database then the locations are checked. If the current location is the same as the obj location then the video plays. If it is not the same then a prompt is given to move to the correct location. If the obj is not in the database then a prompt is given saying the object does not exist.