Eliza.
A Virtual Assistant for the Ubuntu Platform

Lincoln Karuhanga

ABSTRACT

This project seeks to start work on an open-source, community driven, open and transparent virtual assistant primarily targeting the Ubuntu Operating System, with the goal of further adding value to the platform and offering consumers variety.

Key areas of contribution will include exploration of existing research into natural language processing and speech recognition on embedded and primarily offline systems and attempting to consolidate and layout the best performing strategies currently available, improving speech based action resolution for the virtual assistant, including making use of the current application context where possible as well providing a clean intuitive method of interacting with the assistant.

OVERALL OBJECTIVE

To develop a proper structure for and build a Virtual Assistant Primarily targeting the Ubuntu Operating System.

CONTACT

Lincoln Karuhanga
Bachelor of Science in Computer Science
karuhanga475@gmail.com
+256 780 127241
https://www.linkedin.com/in/karuhanga/

BACKGROUND AND SITUATIONAL CONTEXT

Technology exists to make our lives easier and to this end, several tools are constantly being developed and experimented with. A key part of this ecosystem is variety. Variety gives consumers a chance to decide what works for them out of the range of available products.

After decades of dominance on the personal computer scene, the Windows Operating System is increasingly seeing more competition from the previously little known Ubuntu Operating System, a Linux Operating System distribution. A key factor in this growth has been the increasing availability of matching and often better and more open, decentralized tools that offer functionality at par with that offered by the Microsoft Operating System.

An increasingly useful tool on our smart devices is the Virtual Assistant which the Ubuntu Operating System conspicuously lacks.

The purpose of the project is therefore to start work on an open-source, community driven, open and transparent virtual assistant primarily targeting the Ubuntu Operating System, with the goal of further adding value to the platform and offering consumers variety.

METHODS AND MATERIALS

- A passive search and information structure build up approach to action resolution. This will allow the application to return context specific results, for example, requesting a music track to be played will trigger a search in only areas known to have music.
- Triggering some commands on a shortcut based approach. This fits well into the scaling approach and is application specific, which handles the issue of these actions and commands being unique for each application.
- Phoneme based speech recognition
- Speech model generation from training datasets
- The primary data collection tool will be the ODK toolkit.
- Utilising context when resolving actions for commands

CORE ISSUES THE PROJECT IS ADDRESSING

There have been attempts at implementing such a project before, none of which has managed to achieve widespread adoption, for several reasons, the standout ones being;

- Inaccuracy in speech recognition for multiple dialects
- Generic approach to application control
- Attempting to tackle several different issues concurrently, especially in regards to attempting the generic control of multiple applications in a similar manner
- A distinct lack of focus on a single platform
- Limited menu access methods and approaches generally applicable to all applications in the Gnome desktop environment

PROJECT TOOLS

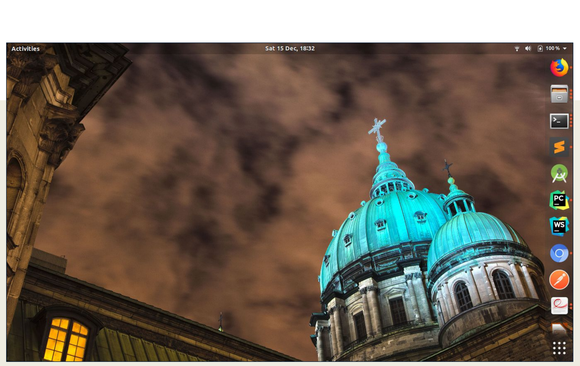
- ODK Server: http://cs2-216003532-karuhanga-odk.appspot.com/
- Github Repositories: https://github.com/Karuhanga/eliza https://github.com/Karuhanga/Elizet
- Documentation: https://drive.google.com/drive/folders/1fB-_SF52nH7X6Wis2Dw245f3MArxWoEy?usp=sharing

EXTERNAL TOOLS

- Kaldi- speech recognition
- PocketSphinx- speech recognition
- pydub- audio processing
- pyaudio- audio device access
- sklearn- a bit of machine learning
- Mozilla Deep Speech
- Google cloud Speech API

SOLUTIONS AND PROJECT OBJECTIVES

- Efficient action resolution and execution based on evaluated input.
- Develop automated menu access methods and approaches generally applicable to all applications in the Gnome desktop environment.
- Understand speech recognition impediments for offline systems and find appropriate solutions to each, ultimately implementing a scalable approach.
- Structure project in a way that allows multiple contributors to gradually add multiple application support
- Understand why previous projects have not achieved widespread use and remedy these issues



PROGRESS TO DATE

- Primary research into current speech recognition standard practice, with primary focus on offline first methods and projects
- Project breakdown and planning complete
- Prototype implementation with proof of concepts for primary features complete

CITATIONS AND REFERENCES

- http://publications.idiap.ch/downloads/papers/2012/Povey_ASRU2011_2011.pdf
- https://arxiv.org/abs/1603.03185
- https://www.researchgate.net/publication/314938892_Comparing_Speech_Recognition_Systems_Microsoft_API_Google_API_And_CMU_Sphinx