Now Presenting: Presentalk

By Graham Hughes, Chanan Walia, Drew Kim, Ryan Hoque, and Kyle Kranen



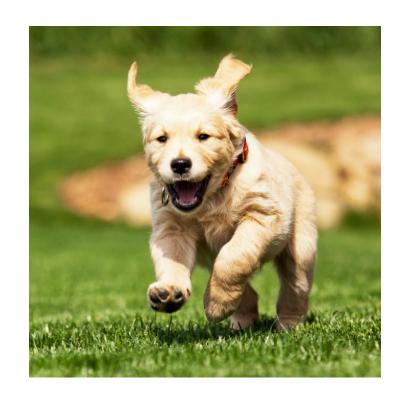
Setting The Scene:

- You're presenting
- You're separated from your computer
- You don't want to break the visual continuity of the presentation

The Solution? Presentalk.

Overview

- Web-based,
 voice-controlled
 presentations
- Allows upload of a PDF file of your presentation
- Gives mic access and uses voice commands to control the presentation



Voice Control

- Uses the Python SpeechRecognition library
 - Defaults to Google Speech Recognition API
 - o Backup is IBM Watson Speech to Text
- Thanks to multithreading, microphone is always active in the background
- Commands require a trigger like "slide" to avoid accidentally changing slides
- Can go forward, back, or to a specific slide number.



More Advanced Functionality

- We've implemented computer vision using the Google Cloud Vision API so that
 we can jump to slides with certain pictures on them. Remember the dog on the
 Overview slide? Let's go to the slide with the dog.
- Say you're a teacher and a student asks what Donald Trump looks like. We've
 integrated Google CSE (Custom Search Engine) API to help you pull images off
 the web right into your presentation. Show me Donald Trump.

More tech, more commands!

- The PDFMiner and PyPDF2 packages scrape text and images from the slides to power the ML image search and text search, e.g. go to the slide titled Voice Control
- Alternatively, search entire slides for text. E.g. Search for 'computer vision.'
- Flask framework connects JavaScript web infrastructure with the Python back-end
- Regular expressions allow flexibility in vocal commands
- Cat on the right too small? Zoom in on the picture of the cat.



Why is Presentalk significant?

- Creates a cleaner visual presentation experience.
- Provides additional functionality over normal presentation software.
- Essentially an intelligent aid for presentation
- Universally applicable as a tool to any PDF

Presentalk.

Because Presentation is Everything.

APIs, Libraries, etc.

- Google Speech Recognition Library
- PortAudio & PyAudio
- IBM Watson Speech to Text Library
- Google Cloud Vision API
- PDFMiner, PyPDF2, XPDF
- Flask Framework
- Google CSE API
- Google Cloud Speech API

Questions?