SPEECH ASSISTANT
Final Prototype

Rahul Rangnekar | Ariel Chen | Justin Hwang | Anish Saha

Group 8: R A J A

05.03.2017

CS 160

CONTRIBUTIONS

Rahul Rangnekar	Product Manager	25%
Ariel Chen	Web Designer	25%
Justin Hwang	Alexa VUI Developer	25%
Anish Saha	Backend Developer	25%

PROBLEM AND SOLUTION OVERVIEW

Many students and professionals struggle with public speaking; presentations are heavily dependent on practice. However, simply reciting in front of a mirror is not necessarily always effective. As such, we created an Alexa skill that will help the user recite a speech--the skill gives feedback on incorrectly recited parts of the speech. We aimed to make speech recitation more interactive, with responses that tell you how many words differed and a repetition of what was spoken. Although the skill only comes equipped with a few famous speeches, the user can upload a custom speech using the web app that was developed to supplement the Alexa skill.

TASKS

1. Upload a speech [difficulty: *medium*]-- Users are able to upload a speech under 1,000 characters to the Amazon DynamoDB using the web interface.

   The user must upload a speech using a feature of the web app. To do this, the user should click the "+" button at the top right corner or clicking on an empty speech card. Then, the user must fill in the appropriate values in the form; this process is relatively intuitive and streamlined in our web app.

2. Search for a speech [difficulty: *easy*] -- Users are able to search for a built-in or user-uploaded speech using the web interface (speech.html) or the voice interface.

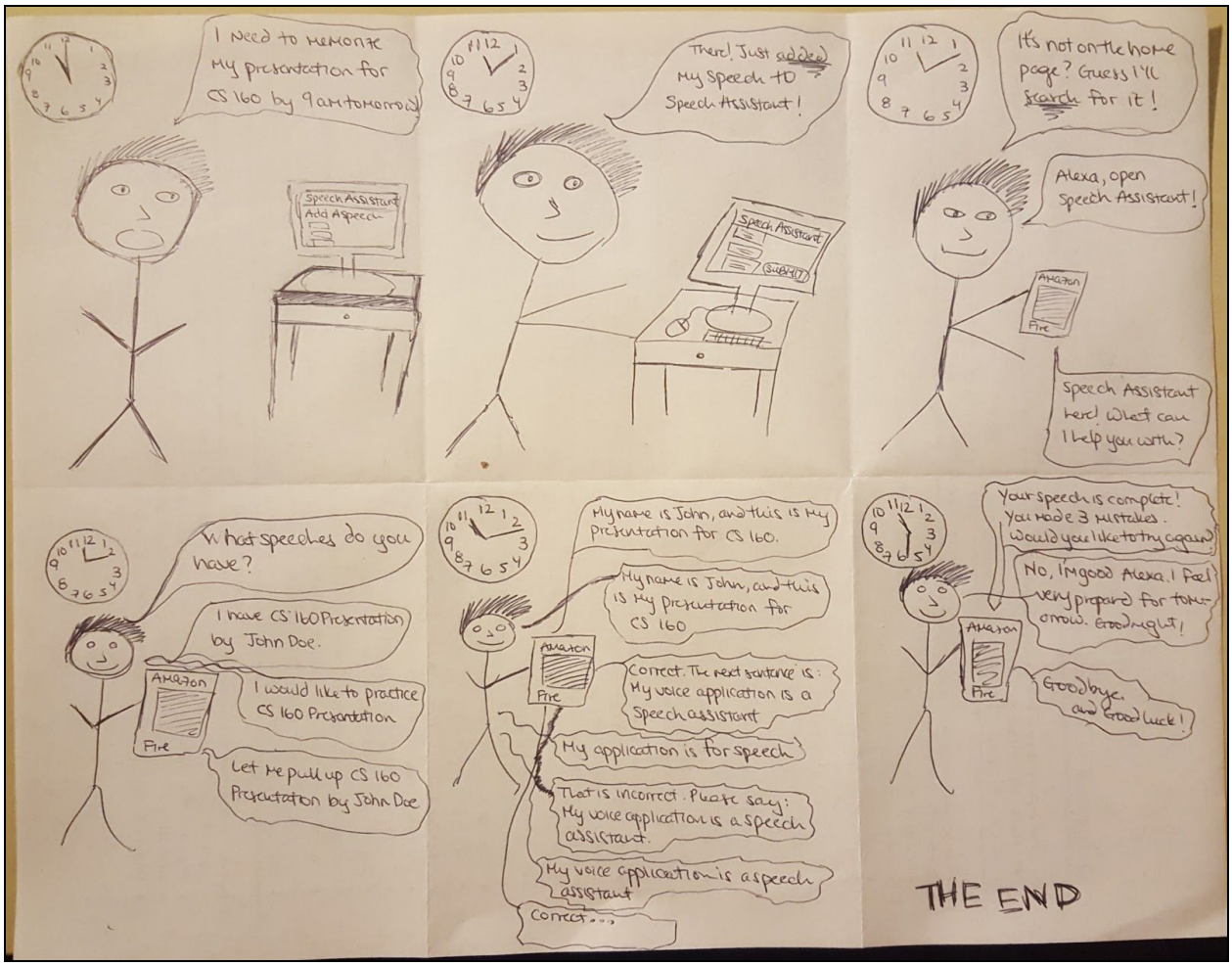
   To do this on the web interface, the user must click the magnifying glass and type in the appropriate term. On the voice interface, the user can ask for a list of speeches or a list of authors. For the list of authors, choosing an author name goes into the speeches specifically by that author.

3. Recite a speech [difficulty: *hard*] -- Users are able to recite a speech line-by-line with aid from the voice interface. Alexa dictates a sentence, then listens for the user to repeat it correctly. Users can read along to the speech using the web UI.

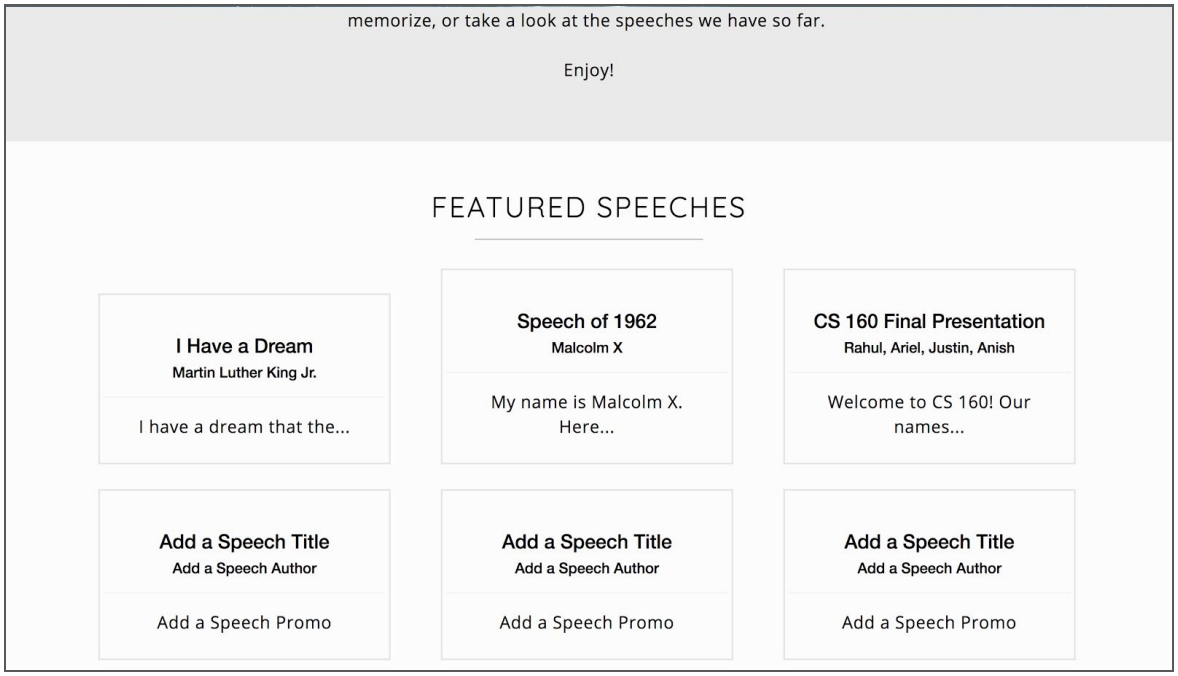
   The user must correctly use the Alexa skill to recite a speech. The user must listen to Alexa's instructions and get feedback for specific parts of the speech as necessary. This task is intuitive because the user simply needs to know the appropriate commands and follow the instructions he/she is given, which our Alexa skill is able to help with.

Storyboard Prototyping

The following storyboard was created as an elementary prototype of the voice interface in its early development stages.



REVISED INTERFACE DESIGN

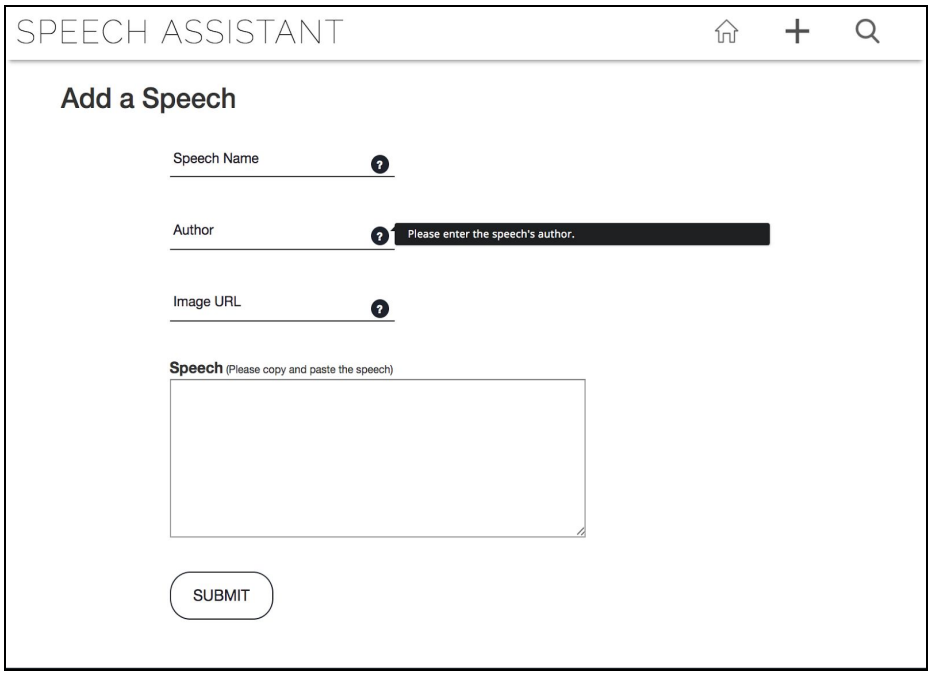


As one of the user reviews suggested, having two plus symbols both leading to the add speech page was redundant, so we altered the look of the “Featured Speeches” section significantly. However, we elected to keep the Add Speech functionality to empty speech cards, as the UI seemed very bare without it. Apart from this, we also increased the contrast between UI elements by using a darker gray bar for the footer and separators. This allowed certain parts of the web app to stand out a bit more. As for the VUI (Alexa skill), we made the dialogue seem more natural by including a wide variety of utterances for each intent.

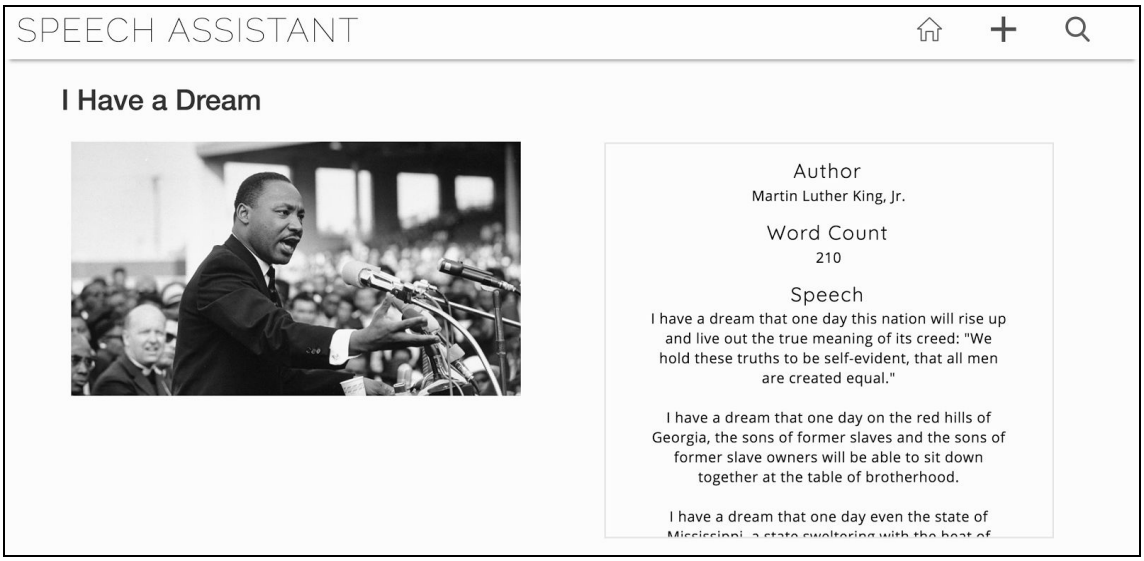
No part of our GUI remains unimplemented. However, we do have a one limitation - a user must search for the exact speech title or author they want from the database. For example, the search query “Malcolm X” brings up his “Speech at Ford Auditorium” but the search query “Malcolm” does not. The issue with redirecting to the proper speech page was resolved in our final prototype, alongside small bugfixes.

PROTOTYPE OVERVIEW

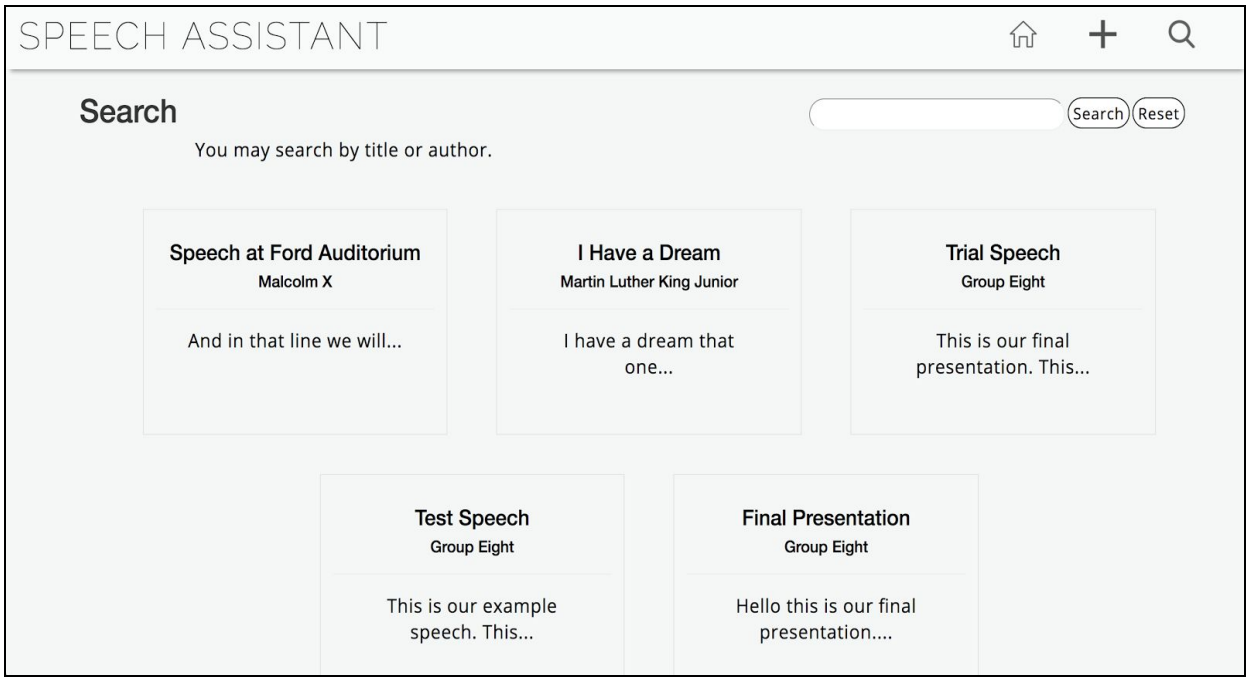
Final Prototype Screenshots



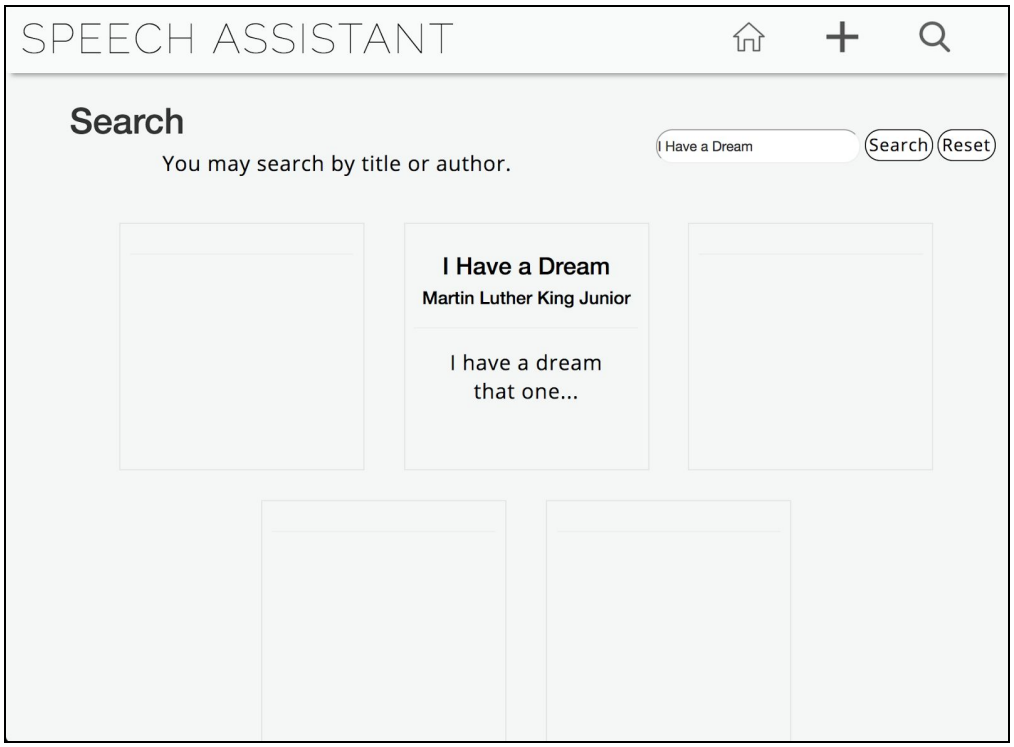
Our *Add a Speech* page was redesigned to be more visually pleasing. We included the ability to request more information upon click of the "?". We wanted to make our web application as minimalistic and intuitive as possible. However, we also wanted to ensure we met Nielsen's heuristic evaluation criteria. Our *Speech* page organizes the information entered in the *Add a Speech* page.



Our *Search* page is used to view and filter through the entire speech database through keywords (speech title or speech author).

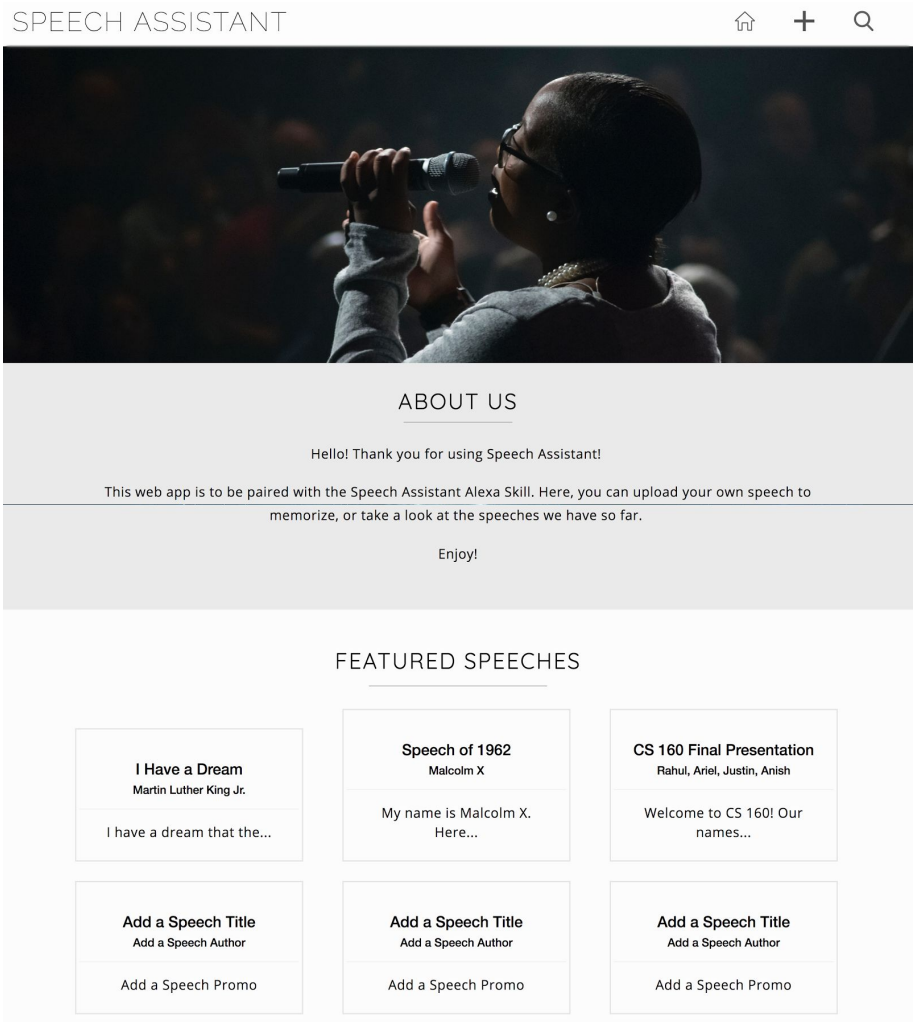


When we query for “I Have a Dream”, a single response appears.



What We Left Out

In our original design, we had included an instructions/tutorial page on how to use our GUI. However, we realized that we would only need that page if our design was unclear. Instead, we decided to iterate over several designs until we reached a simple, intuitive layout. The “About Us” section on our landing page, shown below, achieves such a layout. If the user needs help with the upload page, the hover buttons can serve this function.



Wizard-of-Oz Uses

The limitations of the Alexa software and hardware required a couple Wizard-of-Oz techniques to make the skill work more smoothly.

1. Hard coding of conversion from number to word

   In many speeches, the speech may include the words “one”, “two”, etc. However, when you recite back to Alexa a number, it will register as the number “1”. For example: saying “I have a dream that one day this nation will rise up” is processed as “I have a dream that 1 day this nation will rise up.” This interaction had to be hard coded to change “1” back to the word “one”, so that our word to word checker wouldn’t deem the user input as wrong.

2. Slots and Free Form Speech

   Another limitation of the Alexa setup is that free form speech can be caught as different intents occasionally. It is hard to create a schema where the titles are caught by one intent, each sentence of a speech are caught by another, and authors are caught by a third intent. Often times a speech will contain its title (ex. I Have a Dream) and sentences could be caught by the title Intent. To avoid this, the slots are pretty well enumerated: the title slot contains all the sample titles, the speech slot contains many sentences, and the author slot contains all the authors.

Prototype Video Link

https://www.youtube.com/watch?v=4Ac8o-TozpE