# **Industry Capstone Project**

## Problem definition

A wake word is a phrase that causes an amazon echo or a google assistant device to begin recording an end user’s request so it can be sent to the cloud for processing. When the device detects its wake word, it records the next spoken request and sends a recording of the user’s request to amazon web services or google cloud depending on the device used. Wake words are critical to the design of low-power machine learning process data to wake up the device for full processing.

While typical virtual assistant systems tend to classify wake words as speech commands, this process can generate false alarms from the limited number of negative training samples in the dataset. So, we want to design a machine learning model that can process a custom wake word.

## Why

A custom wake word can help deepen user engagement with associated brands by creating lasting impressions. By saying the custom wake word multiple times during the day, users won't forget which car, TV, app they’re using. This will create the first connection between the customers and the brand when they use the voice assistant. These wake words will be key to sonic branding.

## Success

If the machine learning model is able to respond with the question asked after the wake word or able to verify if it heard the question accurately or not.

## Audience

This problem is being solved for Spotify. Having a custom wake word for spotify not only helps users access their music but also helps the company build a relationship with customers that will help retain them for a long time.

## What

For this used case i will utilize public domain datasets such as :

1. Mozilla common Voice: <https://commonvoice.mozilla.org/en/datasets>
2. Google speech commands dataset:

<https://ai.googleblog.com/2017/08/launching-speech-commands-dataset.html>

The process of MLE for wake word detection involves audio preprocessing, data augmentation, and model training and evaluation. These will form a pipeline for producing deployable models from raw audio data.

For modeling, the suitable ML models that might give us results include Convolutional neural network and recurrent neural network. ROC curves can be used to evaluate the accuracy of the model which compares false alarms per hour and false rejection rate.

## Questions

1. This project needs a machine learning deployment pipeline because we want to implement the custom wake word in a live environment.
2. The product will be able to respond to the questions imposed after the wake work or able to verify if it heard the question accurately or not.
3. Need to deep-dive into CNN and RNN machine learning models for this project.