Creative Integrated Design, Fall 2016

Project Kalexa

with Amazon Web Service

D조

최영진 김준혁 이상훈

윤석찬(AWS, Tech Evangelist)

Table of Contents

[1. Abstract 3](#_Toc469517282)

[2. Introduction 3](#_Toc469517283)

[A. 프로젝트의 배경 및 중요성 3](#_Toc469517284)

[B. 프로젝트 개요 4](#_Toc469517285)

[3. Background Study 4](#_Toc469517286)

[A. 관련 접근 방법 / 기술 장단점 분석 4](#_Toc469517287)

[B. 프로젝트 개발환경 5](#_Toc469517288)

[4. Goal/Problem & Requirement 5](#_Toc469517289)

[5. Project Architecture 5](#_Toc469517290)

[A. Architecture Diagram 5](#_Toc469517291)

[B. Architecture Description 5](#_Toc469517292)

[6. Implementation Spec 5](#_Toc469517293)

[A. Input/Output Interface 5](#_Toc469517294)

[B. Inter Module Communication Interface 5](#_Toc469517295)

[C. Modules 5](#_Toc469517296)

[7. Solution 5](#_Toc469517297)

[A. Implementation Details 5](#_Toc469517298)

[B. Implementation Issues 5](#_Toc469517299)

[8. Results 5](#_Toc469517300)

[A. Experiments 5](#_Toc469517301)

[B. Result Analysis and Discussion 5](#_Toc469517302)

[C. Limit 5](#_Toc469517303)

[D. Future Plan 5](#_Toc469517304)

[9. Division & Assignment of Work 5](#_Toc469517305)

[10. Demo Plan 5](#_Toc469517306)

[11. Conclusion 6](#_Toc469517307)

[[Appendix] User Manual 6](#_Toc469517308)

# Abstract

본 프로젝트에서는 Amazon Alexa에서 동작하는 음성 인식 기반의 응용 프로그램을 제작하였다. Alexa는 아마존에서 개발한 음성 인식 시스템으로, 간단한 음성 비서 기능에 더하여 사용자가 다양한 응용 프로그램을 추가할 수 있도록 되어 있다.

그러나 Alexa는 한국어를 지원하지 않아 한국어로 된 질문을 하거나 응답을 받을 수가 없다. 본 프로젝트는 이 점에 착안하여 Alexa에서 한국어 관련 서비스를 제공하는 것을 목표로 하였다. 한국어 번역 기능, 한국어 학습 기능, K-POP 컨텐츠 제공 등을 주요 요소로 정하여, 어플리케이션을 디자인하고 구현하였다.

해당 프로그램을 제작하기 위하여 AWS(아마존 웹 서비스)내의 다양한 서비스를 활용하였다. 주요 로직을 실행하는 Lambda, 파일 저장소의 역할을 하는 S3, 데이터 저장소의 역할을 하는 dynamoDB 등을 융합하여 서버를 구현하였다. 그리고 네이버의 기계번역 API, 음성합성 API를 사용하여 외국인들이 한국어 음성을 들을 수 있도록 하였다. 또한 멜론의 차트, 앨범, 아티스트 등의 API를 활용하여 K-POP 관련 컨텐츠를 제공하였다. 이 과정에서 여러 프레임워크들이 쓰였으며, 정확도 및 속도를 개선하기 위해 다양한 기법을 적용하였다.

# Introduction

## 프로젝트의 배경 및 중요성

최근 K-POP 스타들의 많은 해외 진출과 소셜 미디어의 발달로 인해 해외에서의 K-POP에 대한 열기가 뜨겁다. 해외 팬들이 다양한 한국어 컨텐츠를 접하면서 한국어 학습에 대한 수요 또한 증가하고 있는 추세이다. 이러한 측면에서 외국인들에게 K-POP에 대한 정보를 제공해 줄 뿐만 아니라 한국어를 간단하게 학습할 수 있고 올바른 발음을 들려주는 기능을 제공하는 것은 충분히 필요가치가 있다.

따라서 본 프로젝트의 목표는 크게 한국어 학습 제공, K-POP 정보 제공의 두 가지로 나뉜다. 현재 Amazon Alexa가 한국어 인식/발음을 모두 지원하지 않고 있는 점이 한국어 기반의 정보를 제공하는 프로젝트에 가장 큰 한계와 문제점이다. 이 제약 조건 속에서 어떤 방식으로 문제를 해결할지 설계하고, 완성도 있는 결과물을 구현하는 것이 본 프로젝트의 일차적 목표이다. 더 나아가서는 K-POP 음악 추천과 같은 심도 있는 기능을 제공하고, 정량적으로 측정 가능한 지표들을 개선하는 작업을 하고자 한다.

## 프로젝트 개요

본 프로젝트에서 제공하려는 기능들을 정리하면 다음과 같다.

* 한국어 학습
  + 한국어 번역 및 원어민 발음 제공
  + 한국어 퀴즈
* K-POP 정보
  + K-POP 차트 정보 제공
  + 가사, 아티스트 및 그에 관련된 정보 제공
  + 음악 재생
* 음악 추천
  + 감정 기반 추천
  + 사용자 데이터 기반 추천

# Background Study

## 관련 접근 방법 / 기술 장단점 분석

### Amazon Alexa



Figure 1 Amazon Echo

아마존 알렉사(Alexa)는 2015년 출시된 일종의 음성 비서이다. 에코(Echo)라는 원통형 스피커 형태로 출시되었는데, 이 안에는 음성 인식 프로그램인 Alexa가 들어 있다. Alexa는 시각, 날씨 알림 등 기본 기능은 물론이고 주요 뉴스나 특정 정보를 브리핑하는 기능을 제공한다. 아마존은 장차 Alexa가 집 및 사무실을 자동화하는 데 핵심적인 역할을 할 것이라고 기대하고 있다.

Alexa는 현재 한국어를 지원하지 않는다. Apple 사의 Siri 등 다른 음성 인식 인공지능 역시 한국어를 지원하는 경우를 찾기 어렵다. 본 프로젝트에서는 Alexa를 타 서비스들과 연계하여 한국어 관련 기능을 제공하는 것을 목적으로 한다. 한국어를 인식하고 한국어를 구사하는 것은 가능하지도 않거니와 본 프로젝트의 범위에서 벗어난다. 그러나 영어 전용 프로그램인 Alexa에 한국어 번역 질의 응답이나 K-POP 등 인기 있는 한국 컨텐츠를 제공하는 기능을 추가하여 Alexa의 가용 범위를 늘리는 데 의의가 있다.

### Amazon Web Service

Amazon은 세계적인 전자 상거래 회사이나, 지난 2006년 Amazon Web Service(AWS)라는 이름으로 클라우드 컴퓨팅 서비스를 시작하였다. AWS는 Amazon.com에서 축적된 기술과 저렴한 가격을 바탕으로 급속도로 성장하였고, 전 세계적으로 가장 널리 이용되는 클라우드 서비스가 되었다. 현재는 Amazon.com이 AWS의 가장 큰 고객이라는 주객이 전도된 듯한 말까지 나오는 상황이다.

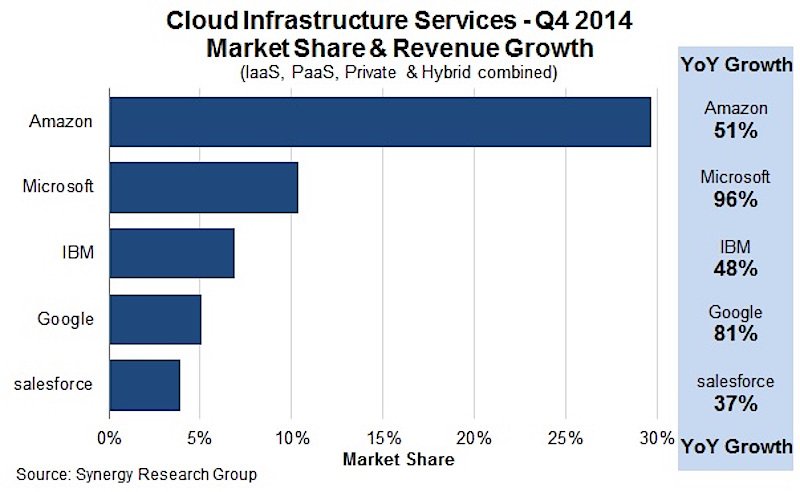


Figure 2 Cloud Service 시장 점유율

본 프로젝트는 AWS 코리아와 진행하는 것이기도 하고, Alexa가 아마존의 서비스이기 때문에 AWS와의 연동이 용이하다는 장점 때문에 AWS의 솔루션들을 사용하기로 결정하였다.

### AWS Lambda

아마존 웹 서비스의 람다(Lambda)는 컴퓨터과학에서 익명 함수(Anonymous Function)을 지칭하는 용어에서 유래하였다. 프로그래밍 언어에서 익명 함수를 복잡한 선언 과정 없이 사용하듯이, 프로그래머가 개발 환경에 신경쓰지 않고 로직에만 신경쓸 수 있도록 한다. 예를 들어, 간단한 HTTP 동작을 처리하기 위해서도 일반적으로 서버를 구축하고 기본적인 웹 개발 환경을 갖추어야 한다. 그러나 아마존 람다는 해당 기능만 구현하면 서버를 관리할 필요 없이 AWS의 환경에서 실행이 가능하다. 기능에 맞는 환경을 편리하게 선택 가능하며, 호출하는 시점만 서버 자원을 사용하므로 가격 면에서도 합리적이다.

본 프로젝트는 기능 자체는 많으나 외부 인터페이스와의 연결이 큰 부분을 차지하고 있어 코드 자체의 규모는 그리 크지 않을 것으로 예상된다. 또한 서버 머신이나 운영 체제 등에 종속될 경우도 없다고 판단되기에, 간편하게 로직을 구현하고 실행할 수 있는 람다가 적합하다고 판단하였다. EC2 등 서버를 사용하면 서버 환경 설정 및 관리에 불필요한 시간이 들어가고, 비용도 더욱 발생하는 데 반해 유리한 점이 많다고 할 수 있다.

### AWS S3

AWS S3은 Simple Storage Service의 약자로, 빠르고 안정적인 저장 공간을 제공하는 솔루션이다. 또한 서비스가 성장함에 따라 필연적으로 저장되는 데이터도 늘어나는데, 이를 자동으로 처리해 줌으로서 Scalablity에 영향을 준다. 본 서비스에서는 한국어 재생을 위한 음성 MP3 파일의 저장에 필수적이다.

### AWS DynamoDB

DynamoDB는 NoSQL 기반 데이터베이스이다. 본 프로젝트의 경우 복잡한 테이블 구조가 필요하지 않아, 빠르게 구현할 수 있는 DynamoDB를 선택하였다. 유저 정보, 곡 정보 등이 저장되게 된다.

### 네이버 기계 번역 / 음성 합성 API

국내 유명 포털인 네이버는 개발자 저변을 확대하면서 여러 공개 API들을 제공하고 있다. 본 프로젝트에서는 그 중 기계 번역 및 음성 합성 API를 사용할 것이다. 기계 번역은 사람의 손을 거치지 않아 아직까지는 번역의 질이 떨어지나, Alexa와 같은 실시간 서비스에서 즉시 응답을 줄 수 있다는 장점이 있다. 또한 네이버의 기계 번역은 한국어 품질이 수준급이라고 알려져 있다. Alexa에서 한국어 번역 질의가 들어왔을 때, 이를 한국어 응답으로 되돌려주기 위해 사용한다.

음성 합성 API란 TTS(Text to Speech) 서비스라고도 불리는데, 텍스트를 음성으로 변환한다. Alexa는 영어 응답을 지정할 수 있으나 한국어는 발음 자체가 불가능하다. 음성 합성 사용은 이를 해결하기 위한 아이디어로, 음성 합성 API에서 반환된 MP3 파일은 언어와 관계 없이 재생이 가능하기에 Alexa가 한국어를 말하도록 할 수 있다.

## 프로젝트 개발환경

# Goal/Problem & Requirement

# Project Architecture

## Architecture Diagram

## Architecture Description

# Implementation Spec

## Input/Output Interface

## Inter Module Communication Interface

## Modules

# Solution

## Implementation Details

## Implementation Issues

# Results

## Experiments

## Result Analysis and Discussion

## Limit

## Future Plan

# Division & Assignment of Work

# Demo Plan

# Conclusion

# [Appendix] User Manual

df