



An Architecture and Development Process for Integration of Alexa-based Voice Service and RESTful API

Chul-Jin Kim*

Dept. of Computer Systems and Engineering, Inha Technical College

ABSTRACT

Artificial Intelligence services such as Apple Siri, Amazon Alexa, Kakao I, and Naver Clova can be used to enhance convenience of service and extend service utilization through integrating with existing web service, mobile service, and IoT(Internet of Thing). By integrating the applied voice recognition technology with the artificial intelligence service, it can act as a voice secretary and it is possible to service learning knowledge of the voice service platform through interaction with the user. Because of the limitations of service extensions within the platform, services such as Amazon Alexa and Naver Clova are opening the platform so that a variety of voice services can be developed by users. In this study, we propose an architecture and development process that can integrate the open platform's voice recognition service with existing RESTful API services. Voice services are developed using Amazon's Alexa Skill Kit (ASK) and use Python's Flask-Ask Library to interface RESTful API services and voice services. In the case study, we experiment the integrating between voice service and RESTful API service through development of hotel reservation voice service. After the speech recognition for the reservation information, the reservation information is stored in the database through the RESTful API service. This verifies the suitability of the voice service and the RESTful API integrating structure.

© 2018 KKITS All rights reserved

KEYWORDS : Alexa, Voice service, Voice recognition, RESTful API, Integration

ARTICLE INFO: Received 26 April 2018, Revised 16 May 2018, Accepted 8 June 2018.

*Corresponding author is with the Department of
Computer Systems and Engineering, Inha Technical

College, 100 Inha-ro Nam-gu Incheon, 22212, KOREA.
E-mail address: cjkim@inhatec.ac.kr

1. 서론

음성인식 기술의 발전으로 인공지능 기반의 음성 서비스가 활발히 개발되고 있다. 현재 애플 시리(Apple Siri), 마이크로소프트 코타나(Microsoft Cortana), 아마존 알렉사(Amazon Alexa), 구글 어시스턴트(Google Assistant)가 대표적인 해외 음성 비서 서비스이며, 국내에서는 SK 텔레콤의 누구(SK Telecom NUGU), KT의 기가 지니(KT GiGA Genie), 삼성전자의 빅스비(Samsung Bixby), 카카오 아이(Kakao I), 네이버의 클로바(Naver Clova) 등이 서비스되고 있다. 이와 같이 음성인식을 통해 인공지능 서비스에 대한 연구개발이 활발하게 진행되고 있으며 음성 서비스를 통해 기존의 서비스들과의 연동을 통해 서비스의 편리성 향상 및 활용 범위를 확대하고자 하고 있다. 이에 본 연구에서는 음성인식 기술과 기존의 서비스를 연동하기 위한 아키텍처 및 개발 프로세스를 연구한다. 기존의 음성인식 기술 중에 공개된 플랫폼인 아마존의 알렉사를 활용하여 RESTful API와의 연동을 연구한다.

1장에서는 음성 인식 서비스와 기존 서비스 연동의 필요성을 설명하며, 2장에서는 관련 연구로서 기존의 음성인식 서비스와 음성인식 연구에 대해 파악한다. 3장에서는 아마존의 알렉사와 RESTful API를 연동하기 위한 아키텍처와 개발 프로세스를 제시한다. 4장에서는 제시한 아키텍처와 개발 프로세스를 검증하기 위해 호텔 예약 음성 서비스를 개발하여 실험하며, 5장에서 결론 및 향후 연구 과제를 제시한다.

2. 관련연구

2.1 아마존 알렉사

아마존에서는 사용자의 음성을 인식하여 상호작

용할 수 있는 인공지능 음성 서비스인 알렉사(Alexa)[1]를 제공한다. 알렉사 음성서비스에서 사용자가 활용할 수 있는 음성 서비스를 스킬(Skill)이라고 하며 개발하기 위한 알렉사 스킬 키트(ASK, Alexa Skills Kit)[2]을 제공한다. 알렉사에서는 스킬 키트의 개발 템플릿을 제공한다.

커스텀 스킬(Custom Skill) : 사용자에게 의해 정의된 음성 서비스를 알렉사를 통해 수행하는 스킬로서, 음성 서비스인 인텐트(Intent)와 어터너스(Utterance)로 구성된 스킬

스마트 홈 스킬(Smart Home Skill) : 클라우드 기반의 가정용 디바이스를 음성으로 제어하기 위한 스킬로서 아마존 웹 서비스 람다(AWS Lambda)를 이용하여 가정용 디바이스를 제어

플리시 브리핑(Flash Briefing) : 뉴스 서비스를 음성으로 제공하기 위한 스킬

음성비서 서비스인 애플 시리(Siri, Speech Interpretation and Recognition Interface)[3], 마이크로소프트 코타나(Cortana)[4, 5]와 다르게 아마존 알렉사의 가장 큰 차별점은 개방형 플랫폼을 지향한다는 것이다.

2.2 인공지능 개인비서 애플 시리와 울프람 알파

연구[6]은 애플 시리의 핵심 엔진 기술에 대해 분석한다. 애플 시리는 음성인식을 위해 음성 신호를 텍스트로 변환해 주는 STT(Speech to Text) 기술과 지식검색을 위해 울프람 알파(Wolfram Alpha)[7, 8] 기술을 적용하고 있다. 울프람 알파는 지식 검색을 위해 수학에 기반을 두고 있는 지식 검색엔진이다. 애플 시리의 음성인식은 뉘앙스 커뮤니케이션(Nuance Communications)의 음성인식 엔진을 사용하며 지식검색은 울프람의 SMP(Symbolic Manipulation Program)[9]을 기반으로 한다.

시리는 음성인식 기술을 통해 아이폰의 다양한 기능을 작동하기 위한 인터페이스 역할을 하며 올 프람 알파 기술을 통해서도 지능형 서비스로 다양한 웹 서비스와의 인터페이스 역할을 한다. 시리의 이러한 지능형 서비스 기술은 현재 비공개형 서비스로서 사용자 의해 음성 검색 서비스를 추가 개발할 수 없다. 이에 반해 아마존의 알렉사는 웹 서비스와의 연동을 통해 음성 서비스를 개발할 수 있다.

2.3 철도예약서비스를 위한 VoiceXML 기반의 음성인식 구현에 관한 연구

연구[10]은 VoiceXML[11]을 이용하여 철도 예약 서비스를 위한 음성인식 연구이다. 인터넷을 통해 철도 예약의 입력된 음성 신호를 VoiceXML를 이용하여 변환하고 음성인식 시스템에 의해 처리된 결과값을 VoiceXML로 변환하여 사용자에게 음성으로 변환하기 위한 방안을 제시한다. 연구[10]은 전화 환경에서의 음성신호에 대한 인식 및 처리를 위한 방안으로 현재 VoiceXML의 호환성, 음성 문법에 따른 모호성, 응답속도, 등의 문제점으로 인해 응용 어플리케이션 개발에 적용하기에 해결해야할 문제점들을 내포하고 있다.

3. 음성서비스와 RESTful API 연동 아키텍처 및 개발 프로세스

본 연구는 음성인식 서비스와 웹 서비스를 연동하기 위한 연구로서 음성인식 서비스와 RESTful API를 연동하기 위한 아키텍처 및 개발 프로세스를 제안한다. 본 연구에서는 특화된 음성인식 서비스를 위해 알렉사 스킬을 기반으로 하며 RESTful API 서비스와 연동을 통해 인식정보를 관리할 수 있다.

3.1 음성 서비스와 RESTful API 연동 아키텍처

음성인식 서비스와 웹 서비스인 RESTful API를 통합하기 위한 기본 구조는 <그림 1>에서와 같이 오디오 디바이스로부터 요청 음성을 전달하면 RESTful API는 해당 음성을 알렉사로 전달한다. 알렉사는 음성을 인식한 후 정보로 전환하여 RESTful API에게 반환하여 오디오 디바이스로 음성 서비스한다. 이와 같이 비즈니스 처리와 음성인식 처리를 연동하기 위해 웹 서비스와 음성 서비스를 연동할 수 있다.

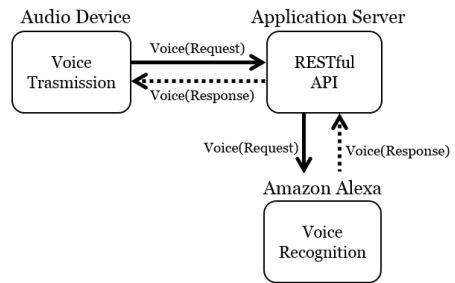


그림 1. 음성 인식 서비스와 RESTful API의 통합 구조
Figure 1. Integration Structure of Voice Recognition Service and RESTful API

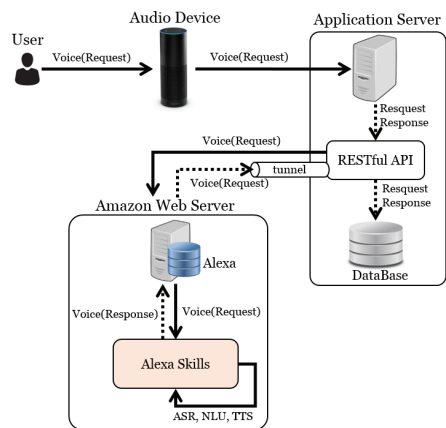


그림 2. 알렉사와 RESTful API 통합 아키텍처
Figure 2. Integration Architecture of Alexa and RESTful API

음성서비스와 RESTful API 연동을 위한 상세 아키텍처는 <그림 2>와 같다. 요청하는 음성에 대해 RESTful API는 음성 인식을 위해 아마존 알렉사 서비스를 이용한다. 아마존 웹 서버로 전송된 음성은 개인화된 서비스인 스킬을 호출하여 음성 인식을 수행한다.

자동음성인식(ASR, Automated Speech Recognition), 자연어이해(NLU, Natural Language Understanding), 텍스트음성변환(TTS, Text To Speech) 기능은 알렉사 서비스에서 제공되는 서비스로서 요청된 음성 에 대해 음성인식 및 자연어이해, 그리고 음성변환 처리를 수행한 후 응답음성을 웹 서비스인 RESTful API로 전달한다. 웹 서비스를 아마존 람다(Amazon Lambda)를 이용할 수 있으나 본 연구에서는 외부 웹 서비스와의 연동 구조를 제안한다. 알렉사로부터 음성을 외부의 RESTful API를 호출하여 전달하기 위해서는 방화벽을 통과하기 위한 터널링 서비스를 활용해야 한다.

스킬은 음성으로 질의하고 응답하는 기능을 제공하는 음성 서비스로서, RESTful API로부터 음성 서비스를 전달받고 음성 서비스를 제공하기 위한 스킬의 처리 과정은 <그림 3>에서와 같이 RESTful API로부터 음성을 스킬로 전달하면 스킬은 ASR를 통해 음성인식을 수행하고 인식된 결과는 NLU를 통해 자연어를 인식할 수 있도록 처리한다. 일반적으로 NLU는 처리 결과를 인텐트(Intent)로 결과를 반환한다. 인텐트는 개인화된 스킬로 전달하여 응용 서비스를 수행하여 결과를 텍스트로 반환하면 TTS에 의해 음성으로 반환하여 RESTful API로 반환한다.

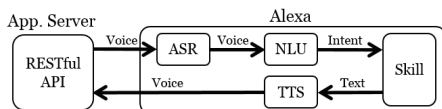


그림 3. 요청 음성의 인식 프로세스
Figure 3. Recognition Process of Request Voice

3.2 음성 서비스와 RESTful API 연동 개발 프로세스

알렉사의 음성 인식 서비스와 RESTful API 연동 아키텍처를 기반으로 개발 프로세스를 제안한다. RESTful API는 파이썬을 기반의 알렉사와 연동하기 위한 라이브러리인 플라스크 에스크(Flask-ask) [12] 라이브러리를 이용하여 RESTful API를 구현한다.

음성 서비스와 RESTful API 연동 개발을 위한 개발 프로세스는 다음과 같다.

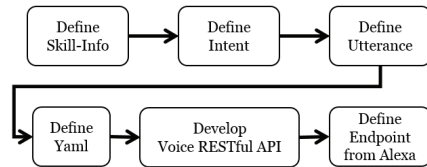


그림 4. 알렉사와 RESTful API 통합 개발 프로세스
Figure 4. Development Process for Integration of Alexa and RESTful API

<그림 4>의 개발 프로세스에서 음성 서비스 관련 단계는 아마존의 개발자 포탈[13]에서 정의할 수 있다. 아마존 개발자 포탈에서 알렉사 스킬을 개발하기 위한 도구인 알렉사 스킬 키트(ASK)을 이용하여 개발한다. 본 연구에서는 알렉사 스킬 키트 기능 중에 커스텀 스킬을 이용하여 RESTful API와 연동한다.

- 스킬 기본 정보 정의

스킬 기본 정보는 스킬 명과 스킬을 호출할 수 있는 호출명(Invocation Name)으로 구성된다. 호출명은 알렉사 서비스를 이용 시 스킬을 구동시키는 역할을 하므로 유일한 호출명을 정의해야 한다.

-인텐트 정의

인텐트는 사용자에게 의해 요청한 음성을 처리하기 위한 구조를 나타내며, 인텐트 명과 슬롯(slot)으

로 구성된다. 슬롯은 인텐트 명이 나타내는 대화 문장의 데이터를 나타내는 것으로 인텐트 명에 해당하는 대화 문장은 어터런스에 정의한다.

```
{ "intents": [
  { "intent": INTENT_NAME,
    "slots": [
      { "name": DATA_NAME,
        "type": AMAZON_DATE_TYPE },
    ]
  }
]}
```

그림 5. 인텐트 구조
Figure 5. Intent Structure

<그림 5>에서와 같이 “intent”에 인텐트 명을 정의하며 “slots”에는 어터런스 문장과 매핑될 데이터(데이터명, 데이터 타입)를 정의한다. 슬롯 데이터 타입의 종류는 아마존에서 정의하는 데이터 타입으로 정의하며, 정의되는 데이터 타입에 따라 음성으로 입력된 데이터가 숫자 인지, 주소 인지, 성명 인지를 알렉사가 인지할 수 있다. 인텐트의 구조는 JSON 구조를 사용하여 정의한다.

- 어터런스 정의

어터런스는 대화 문장을 의미하며 <그림 6>에서와 같이 인텐트 명에 해당하는 대화 문장을 정의한다. 인텐트 명은 <그림 5>에서 정의한 것 같이 인텐트에서 정의한 인텐트 명과 맵핑된다.

```
INTENT_NAME Conversation Sentence
```

그림 6. 어터런스 구조
Figure 6. Utterance Structure

어터런스에서 정의한 대화 문장에서 사용자에게 의해 요청된 음성 데이터가 포함되면 <그림 7>에서와 같이 인텐트의 데이터와 어터런스의 문장이 결합되어 대화 문장이 구성된다.

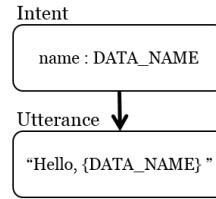


그림 7. 인텐트와 어터런스 맵핑
Figure 7. Mapping of Intent and Utterance

-Yaml 정의

진자(Jinja)[14] 기반의 템플릿을 이용하여 RESTful API 서비스 소스 코드에 대화 문장을 분리하기 위해 YAML(YAML Ain' t Markup Language)[15]를 이용할 수 있다. 알렉사 스킵 킷을 이용하여 어터런스를 정의할 경우 데이터에 관련된 정보를 인텐트와 어터런스에서 상호 참조하며 정의하면 되지만, RESTful API의 서비스 코드에서 발생하는 대화 문장을 정의할 때 Yaml에 분리해서 정의하면 관리 측면에서 효율적일 것이다.

- 음성 RESTful API 서비스 개발

음성 RESTful API 는 웹 서비스를 통해 알렉사와 연동하여 음성 처리를 하기 위한 서비스이다. 알렉사와의 연동을 위해 파이썬 기반의 플라스크 에스크 라이브러리를 사용하여 알렉사의 음성 서비스를 이용한다.

```
@ask.launch
def startService():

@ask.intent(intent_name)
def service():
```

그림 8. 음성 RESTful API의 실행 및 인텐트 서비스
Figure 8. Launch and Intent Service of Voice RESTful API

<그림 8>에서와 같이 @ask.launch 은 사용자에게 의해 스킵이 구동되었을 때 웹 서비스가 구동할 수 있도록 정의한다(startService()). @ask.intent(intent_name)

은 사용자에게 의해 요청된 음성의 구조가 intent_name 과 일치할 경우 처리하기 위한 서비스를 정의한다. 알렉사와의 상호 대화를 @ask.intent를 이용하여 개발할 수 있다.

- 알렉사 엔드포인트 정의

알렉사에서 음성 처리 결과를 제공할 때 아마존 웹 서비스의 람다(AWS Lambda)를 이용할 수 있으나 본 연구에서는 외부 서비스인 RESTful API를 이용한다. 알렉사에서 외부 웹 서비스를 접근하기 위해서 <그림 2>의 아키텍처에서와 같이 터널링 서비스를 통해 알렉사에서 음성 서비스를 제공할 수 있도록 해야 한다. 터널링 서비스를 통해 방화벽이나 NAT(Network Address Translation) 내부의 웹 서비스를 노출시키기 위한 서비스를 이용하여 알렉사에서 접근할 수 있도록 해야 한다. 본 연구에서는 ngrok[16] 터널링 서비스를 이용하여 외부에서 접근할 수 있도록 설정하였으며 알렉사의 엔드포인트를 설정한다.

4. 실험 및 평가

본 실험에서는 알렉사 기반의 음성 인식 서비스와 RESTful API 연동을 위한 아키텍처에 대해 호텔 예약 음성서비스 사례를 개발하여 검증하며, 개발 과정은 3장에서 제시한 개발 프로세스를 따른다.

호텔 예약 음성 서비스를 위한 인텐트는 <그림 9>와 같이 정의한다. 3개의 인텐트를 정의하였으며 “ReserveIntent” 는 사용자에게 의해 음성으로 응답하는 예약 정보에 대한 구조를 정의한다. 예약자 성명(rName), 예약날짜(rDate), 예약기간(rPeriod)에 대해 정의하며 데이터의 타입은 알렉사에서 음성 인식할 수 있도록 데이터 타입으로 정의한다. “AMAZON.DATE” 와 같은 데이터 타입은 아마존 서비스의 시스템 날짜와 연동하여 음성 데이터를

인식한다. “today” 라고 음성 입력하여 오늘 날짜를 얻어와 데이터 처리나 음성 서비스 한다.

```
{ "intents": [
  { "intent": "YesIntent" },
  { "intent": "ReserveIntent",
    "slots": [
      { "name": "rName",
        "type": "AMAZON.US_FIRST_NAME" },
      { "name": "rDate",
        "type": "AMAZON.DATE" },
      { "name": "rPeriod",
        "type": "AMAZON.DURATION" }
    ]
  },
  { "intent": "CompleteIntent" }
]
```

그림 9. 호텔 예약 음성 서비스를 위한 인텐트
Figure 9. Intents for Hotel Reservation Voice Service

정의된 인텐트에 맵핑되는 음성 응답 표현인 어터런스를 <그림 10>과 같이 정의한다. 3개의 인텐트에 대해 사용자에게 의해서 전달되는 음성의 표현 문장을 정의한다. 사용자에게 의해 예약하기 위한 3개의 음성 데이터를 전송 받을 경우 “ReserveIntent” 인텐트로 인식한다. 3개의 예약 정보에 대해 음성 인식하여 데이터를 응답 문장과 결합하여 음성 서비스 한다.

```
YesIntent yes
ReserveIntent {rName} {rDate} {rPeriod}
ReserveIntent {rName} {rDate} and {rPeriod}
CompleteIntent Thank you for Reservation
```

그림 10. 호텔 예약 음성 서비스를 위한 어터런스
Figure 10. Utterances for Hotel Reservation Voice Service

사용자에게 음성 서비스될 문장은 yaml로 정의하며 RESTful API 서비스 코드와 분리하여 정의한다. 어터런스에 정의할 수도 있으나 어터런스는 인텐트와의 맵핑을 위한 문장을 정의할 경우에 활용하며 일반 인텐트와 분리하는 음성 서비스 문장을 정의할 때 웹 서비스 코드와 분리하여 정의할 수 있다.

```
start : Hello. It is a Korean Hotel. Would you like to
       make a reservation ?
reserve : What is your name, reservation date and
         reservation period ?
complete : Your reservation is complete. Thank you.
```

그림 11. 호텔 예약 음성 서비스를 위한 Yaml
Figure 11. yaml for Hotel Reservation Voice Service

RESTful API 서비스 코드는 <그림 12>와 같이 파이썬의 플라스크를 활용하여 호텔 예약 음성 서비스를 개발하였다. 알렉사와 연동하기 위해 플라스크 에스크를 포함하였으며, 본 호텔 예약 음성 서비스를 시작하면 “@ask.launch” 에서 정의한 “reserve_start()” 함수가 수행된다.

```
@ask.launch
def reserve_start():
    start_msg = render_template('start')
    return question(start_msg)
...
@ask.intent("ReserveIntent")
def reserve_info(rName, rDate, rPeriod):
    reserve_info_msg =
    = "Reservation information of {0} ₩
    is reserved for {1} days in {2}.". ₩
    format(rName, rPeriod, rDate)
    db.insert(rName, rDate, rPeriod)
    return statement(reserve_info_msg)
```

그림 12. 호텔 예약 음성 서비스를 위한 웹 서비스
Figure 12. Web Service for Hotel Reservation Voice Service

사용자에게 제공되는 음성 서비스는 <그림 11>의 yaml에서 정의한 “start” 문장이 음성 변환이 되어 서비스된다. 사용자에게 의해 예약 음성 정보가 입력되면 3개의 데이터인 예약자 성명(rName), 예약날짜(rDate), 예약기간(rPeriod)의 구조로 입력되면 “ReserveIntent” 인텐트와 맵핑되는 “reserve_info()” 함수가 실행된다. 본 함수는 예약 음성 정보와 문장을 구성하여 음성 서비스하며 데이터베이스에 저장하는 기능을 수행한다. 입력 받은 음성을 인식하여 데이터베이스에 저장할 수 있다.

지금까지 개발 프로세스에 의해 알렉사 스킵 킷 기반의 인텐트, 어터런스를 정의하였으며, 웹 서비

스 개발을 위해 파이썬 플라스크 에스크를 이용하여 호텔 예약 음성서비스를 개발하였다.

<그림 13>과 같이 로컬에 RESTful API 웹 서비스를 구동하며 알렉사에서 음성을 웹 서비스로 전달하기 위해 터널링 서비스를 <그림 14>와 같이 구동한다. 포워딩 주소는 알렉사에서 접근하기 위한 URL 주소가 되며 알렉사 스킵 킷 정의의 시 설정해 주어야 한다.

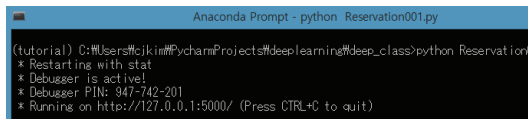


그림 13. 호텔 예약 음성 서비스의 실행

Figure 13. Running of Hotel Reservation Voice Service

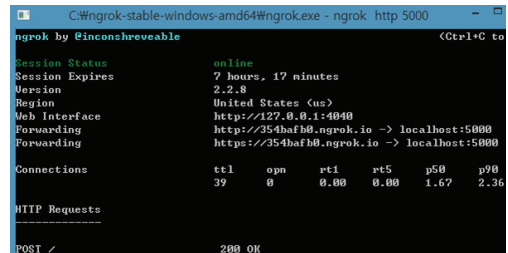


그림 14. 알렉사에 의해 웹서비스 접근을 위한 터널링 서비스

Figure 14. Tunneling Service for Accessing Web Service by Alexa

알렉사 및 RESTful API 서비스가 구동된 상태에서 알렉사를 호출하여 호텔 예약 음성 서비스를 제공 받을 수 있다.

호텔 예약 음성 서비스 실행 명령어는 “Alexa, start” + 호텔 예약 스킵명” 와 같이 본 호텔 예약 음성 서비스의 스킵명은 “Korean Hotel Reservation” 이며 <그림 15>와 같이 호텔 예약 음성 서비스 스킵을 시작할 수 있다. 음성에 의한 상호 작용은 인텐트와 어터런스에 의해 정의된 구조와 문장에 의해 처리된다. <그림 15>는 알렉사 시뮬레이터(Alexa Simulator)를 이용하여 음성 대화

를 문장으로 제공한다.

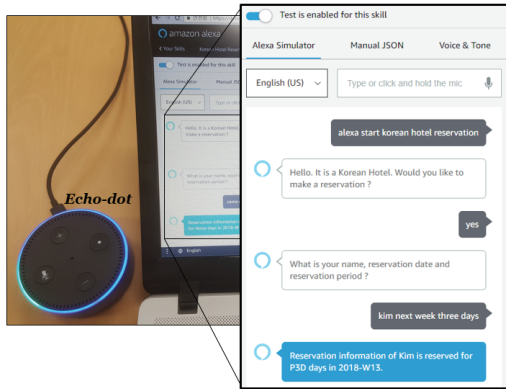


그림 15. 에코닷을 이용한 호텔 예약 음성 서비스 결과
Figure 15. Result of Hotel Reservation Voice Service using Echo-dot

호텔 예약 정보는 <그림 15>에서와 같이 사용자에게 의해 3개의 데이터가 음성 입력되면 “ReserveIntent”로 인식하고 해당 음성 데이터에 대해 알렉사는 어터런스에 정의된 문장 구조로 데이터를 인식하여 웹 서비스로 전달한다. 전달된 웹 서비스에서는 사용자가 확인할 수 있는 문장으로 구성하여 음성 서비스한다. <그림 16>과 같이 예약 정보에 대해 음성 인식하여 <그림 17>과 같이 데이터베이스에 저장된다. 이와 같이 음성으로 입력된 정보는 알렉사의 NLU 기술을 통해 정확히 인식되며 <그림 17>에서와 같이 음성 정보가 데이터베이스에 저장됨을 확인할 수 있다.

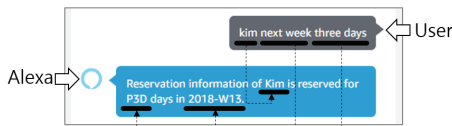


그림 16. 호텔 예약 정보
Figure 16. Hotel Reservation Information

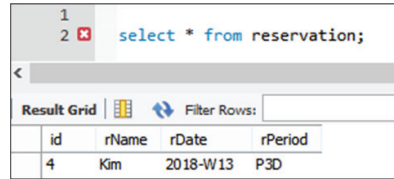


그림 17. 호텔 예약 정보의 DB 입력 결과
Figure 17. DB Insert Result of Hotel Reservation Information

지금까지 알렉사 스킬 킷과 파이썬 플라스크 에스크 기반으로 호텔 예약 음성 서비스를 개발하여 본 연구에서 제안한 알렉사 음성 서비스와 RESTful API 연동 아키텍처 및 개발 프로세스를 검증하였다. 또한 입력되는 음성을 데이터로 변환하여 웹 서비스의 비즈니스 로직과 연동되는 구조를 확인하였다.

5. 결론

본 연구에서는 아마존의 알렉사 음성 서비스와 RESTful API 서비스를 연동하기 위한 아키텍처와 개발 프로세스를 제안하였다. 음성 서비스는 알렉사 스킬 킷을 이용하여 개발하며 연동하기 위한 RESTful API 서비스는 파이썬의 플라스크 에스크 라이브러리를 활용하여 개발할 수 있다. 음성 서비스의 음성 구조는 알렉사의 인텐트와 어터런트를 이용하여 정의할 수 있으며 이러한 인텐트는 RESTful API 에서 참조하여 서비스될 수 있다. 사례연구에서는 음성 서비스와 RESTful API 연동 아키텍처를 확인하기 위해 호텔 예약 음성 서비스를 개발하였다. 음성으로 예약된 정보가 데이터로 추출이 되며 이러한 예약 정보가 RESTful API 서비스를 통해 데이터베이스에 저장될 수 있음을 확인하였다.

향후에는 알렉사의 음성 서비스와 객체인식 기술을 연동하여 시각장애인을 위한 객체 인식 음성 서비스를 연구하고자 한다.

References

- [1] Amazon Alexa, <https://developer.amazon.com/alexa>, retrieved Jan. 2018.
- [2] Alexa Skills Kit, <https://developer.amazon.com/alexa-skills-kit>, retrieved Jan. 2018.
- [3] Apple Siri, <https://www.apple.com/kr/ios/siri>, retrieved Jan. 2018.
- [4] MicroSoft Cortana, <https://developer.microsoft.com/en-us/cortana>, retrieved Jan. 2018.
- [5] Y. W. Kim, and K. H. Lee, *Microsoft artificial secretary cotana*, The Korean Society Of Computer And Information, Vol. 25, No. 1, pp. 7-10, Jun. 2017.
- [6] S. S. Kang, *AI assistant Apple's Siri and Wolfram Alpha*, Communications of the Korean Institute of Information Scientists and Engineers, Vol. 35, No. 8, pp. 42-50, Aug. 2017.
- [7] Wolfram Alpha Service, <http://www.wolframalpha.com>, retrieved Feb. 2018.
- [8] Wolfram Alpha, https://en.wikipedia.org/wiki/Wolfram_Alpha, retrieved Feb. 2018.
- [9] C. A. Cole, and S. Wolfram, *SMP: A symbolic manipulation program*, Proceedings of the fourth ACM symposium on Symbolic and algebraic computation (SIGSAM), Snowbird, Utah, 1981.
- [10] B. S. Kim, and S. H. Kim, *A study on realization of speech recognition system based on voiceXML for railroad reservation service*, Journal of the Korean Society for Railway, Vol. 14, No. 2, pp. 130-136, Apr. 2011.
- [11] VoiceXML, <http://www.voicexml.org>, retrieved Jan. 2018.
- [12] Flask-Ask: A New Python Framework for Rapid Alexa Skills Kit Development, <https://developer.amazon.com/blogs/post/Tx14R0IYYGH3SKT/Flask-Ask-A-New-Python-Framework-for-Rapid-Alexa-Skills-Kit-Development>, retrieved Feb. 2018.
- [13] Amazon Developer Portal, <https://developer.amazon.com>, retrived Feb. 2018.
- [14] Jinja, <http://jinja.pocoo.org/>, retrived Mar. 2018.
- [15] YAML Ain't Markup Language (YAML) Version 1.1, <http://yaml.org/spec/1.1/>, retrived Mar. 2018.
- [16] ngrok, <https://ngrok.com/>, retrived Mar. 2018.

알렉사 기반의 음성서비스와 RESTfull API 연동을 위한 아키텍처 및 개발 프로세스

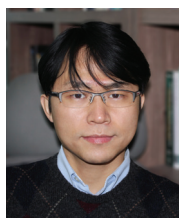
김철진

인하공전 컴퓨터시스템과

요 약

인공 지능 음성 서비스인 애플 시리, 아마존 알렉사, 카카오 아이, 네이버 클로바 등은 기존의 웹 서비스, 모바일 서비스, 사물 인터넷 등과 연동을 통해 서비스의 편리성을 향상시키고 서비스의 활용 범위를 확대 시킬 수 있다. 적용된 음성 인식 기술에 인공지능 서비스를 결합하여 음성 비서로서의 역할을 할 수 있도록 하고 있으며, 사용자와의 상호 대화를 통해 음성 서비스 플랫폼의 학습 지식을 서비스할 수 있도록 하고 있다. 플랫폼 내에서 서비스 확장이 한계가 있으므로 아마존 알렉사나 네이버 클로바 등과 같은 서비스는 플랫폼을 공개하여 사용자에게 의해 다양한 음성

서비스가 개발될 수 있도록 하고 있다. 이에, 본 연구에서는 오픈 플랫폼의 음성 인식 서비스와 기존의 RESTful API 서비스를 연동할 수 있는 아키텍처와 개발 프로세스를 제안한다. 음성 서비스는 아마존의 알렉사 개발 도구인 스킬킷(ASK, Alexa Skill Kit)을 이용하여 개발하며 RESTful API 서비스와 음성 서비스 연동을 위해 파이썬의 플라스크 에스크 라이브러리를 이용한다. 사례 연구에서는 호텔 예약 음성 서비스 개발을 통해 음성 서비스와 RESTful API 서비스의 연동을 실험한다. 예약 정보에 대해 음성 인식 후 RESTful API 서비스를 통해 예약 정보가 데이터베이스에 저장되도록 한다. 이를 통해 음성 서비스와 RESTful API 연동 구조의 적합성을 검증한다.



Chul Jin Kim received his B.E. degree in Computer Science from the Kyonggi University in 1996. He received both M.S. and Ph. D. degree in Computer Science from the Soongsil

University in 1998 and 2004. He worked as a senior researcher in SAMSUNG Electronics from 2004 to 2009. He has been a professor in the Department of Computer Systems and Engineering at Inha Technical College since 2009. His research focus is software engineering, object-oriented & component technique, software architecture, customization, and embedded software.

E-mail address: cjkim@inhac.ac.kr