

# Watson Assistant를 이용한 신조어 챗봇 시스템 구현\*

고준희, 남윤한, 박현중, 최성진, 백세인, 이용규  
동국대학교 컴퓨터공학과-서울  
e-mail: junheee.ko@gmail.com

## Implementation of a Neologism Chatbot System Using Watson Assistant

Jun Hee Ko, Yun Han Nam, Hyeon Jong Park, Seong Jin Choi,  
Se In Baek, Yong Kyu Lee

Department of Computer Science and Engineering, Dongguk University-Seoul

### 요 약

인공지능의 중요성과 가치가 대두됨에 따라 이를 활용하여 다양한 챗봇들이 개발되고 있다. 그러나 신조어를 의미를 제공하는 챗봇 시스템의 연구가 부족하다. 본 논문에서는 Watson Assistant를 활용하여 신조어의 의미를 제공하는 신조어 챗봇 시스템을 구현하고자 한다. Watson Assistant는 자연어 형식으로 된 질문에 응답하는 것을 지원하는 API다. 웹 크롤링을 통해 신조어 데이터를 수집 및 저장하고, Watson Assistant를 활용하여 사용자의 의도를 파악하여 신조어의 의미를 추출한다. 제안한 시스템을 통해 현실 세계에서 끊임없이 양산되는 신조어의 의미를 빠르게 파악할 것이라고 기대한다.

### 1. 서론

대화형 에이전트(Conversational Agent)는 사용자와 시스템이 정보를 주고받을 수 있는 시스템으로, 자연어 처리를 통해 자연스러운 대화를 수행할 수 있다. 과거에는 이러한 자연어 처리 시스템을 구현하기 위해 언어 처리 모듈을 직접 구성했다[1]. 현재 개발자들이 대화형 에이전트를 쉽게 구현할 수 있도록, 여러 기업들은 언어 처리 기능을 자동화한 API를 배포하고 있다.

챗봇(Chatbot)은 대화형 에이전트의 한 예로써, 정해진 응답 규칙을 기반으로 사용자의 질문에 응답하는 시스템이다. 주제에 맞는 간단한 질의 응답형 챗봇에서부터 사용자의 패턴을 분석하는 서비스까지 점차 그 활용 범위가 넓어지고 있다. 이에 인공지능(AI, Artificial Intelligence) 기술까지 결합하는 연구가 최근에 활발히 진행 중이다[2].

인공지능은 학습, 문제 해결, 패턴 인식 등과 같이 인간 능력과 관련된 인지 문제를 해결할 수 있는 컴퓨터 공학 분야이다. 인공지능은 로봇 공학이나 미래의 모습을 내포하고 있지만, 공상 과학 소설의 비현실적인 로봇을 넘어 현실적인 첨단 컴퓨터 공학이 되고 있다. 빅데이터를 기계가 스스로 학습하면서 인공지능의 지능이 혁신적으로 발전하고 있다[3]. 이러한 인공지능을 활용한 언어 처리 기능을 지원하는 기술로 IBM의 Watson Assistant가 있다. Watson Assistant는 앱과 사용자 간 자연어 대화 플로우의 상호작용을 가능하게 하는 API다. 이를 활용하여, 사용

자의 의도(Intent)를 파악하여 의도에 맞는 데이터를 제공할 수 있다.

인공지능과 결합된 챗봇 서비스는 인간과 인간이 대화하듯이 인간과 컴퓨터 간에 대화가 가능하도록 돕는다. 따라서 IT 기기 사용에 능숙하지 않은 사람도 챗봇과 대화를 함으로써 필요한 정보를 얻을 수 있다. 누구나 쉽게 접근 가능한 신조어 챗봇을 이용해 끊임없이 만들어지는 신조어의 의미를 단 시간 내에 파악하기 위해 신조어 챗봇이 필요하다.

본 논문은 Watson Assistant와 웹 크롤러, Node.js를 활용하여 신조어 검색 및 학습을 손쉽게 할 수 있도록 돕는 신조어 챗봇 시스템을 제안하고자 한다.

### 2. 관련 연구

#### 2.1 챗봇(Chatbot)

챗봇(Chatbot)[4]은 음성이나 문자를 사용하는 인간과의 대화를 통해 특정한 작업을 수행하는 컴퓨터 프로그램이다. 또한 대화형 인터페이스에서 정해진 규칙이나 인공지능을 기반으로 사용자와 상호작용을 하는 서비스이다.

최근 챗봇 시스템을 이용해 사회적 문제를 해결하려는 연구가 진행되고 있다. 뉴스 미디어 산업의 위기를 타개하고자 챗봇을 활용한 사례가 있다[5]. 또한 챗봇에 크라우드 소싱(Crowdsourcing)을 접목하여 질의응답 시스템을 구현한 사례도 있다[6].

챗봇의 동작 방식은 세 가지가 있다. 인간의 언어에 자연어 처리 기술을 적용한 언어 이해 방식이 있고, 입력한

\* 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 SW중심대학지원사업의 연구결과로 수행되었음(2016-0-00017)

은 말의 특정 단어를 파악하여 결과를 출력하는 검색 방식이 있다. 또한, 각본에 따라 사용자의 입력을 받고 결과를 출력하는 각본 방식이 있다.

본 논문에서는 언어 이해 방식과 검색 방식, 각본 방식을 모두 결합한 신조어 챗봇 시스템을 구현하고자 한다.

## 2.2 Watson Assistant

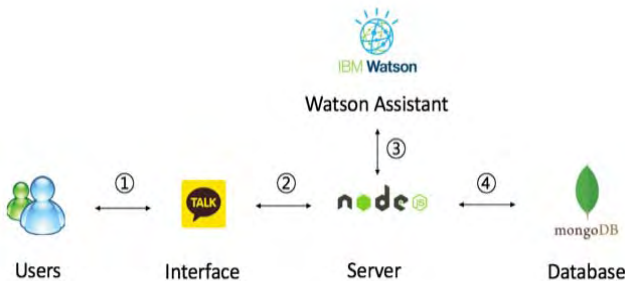
Watson Assistant[7]는 자연어 형식으로 입력된 질문에 응답할 수 있는 인공지능 컴퓨터 시스템이며, IBM에 의해 개발되었다. 현재 금융, 방송, 의학, 교육, 쇼핑등의 다양한 분야에 적용되고 있다. 음성 인식 및 이미지 인식, 데이터 인사이트의 기능을 하는 총 28개의 API를 제공하고 있으며, 지속적으로 확장되고 있다. 신조어 챗봇에 사용자의 요청이 들어왔을 때, 요청의 컨텍스트(Context)를 정확하게 파악하도록 보장해야 한다.

본 논문에서는 사용자의 의도를 파악하는 Watson Assistant를 활용하여 신조어 챗봇 시스템을 구현한다.

## 3. Watson Assistant를 활용한 신조어 챗봇

본 논문에서는 Watson Assistant를 활용한 신조어 챗봇 시스템을 구현한다. 신조어 챗봇 시스템은 사용자의 질의를 파악하여 데이터베이스에 저장되어 있는 해당 신조어의 의미를 출력한다.

본 논문에서 제안하는 신조어 챗봇 시스템의 전체 구성도는 (그림 1)과 같다.



(그림 1) 신조어 챗봇 시스템의 전체 구성도

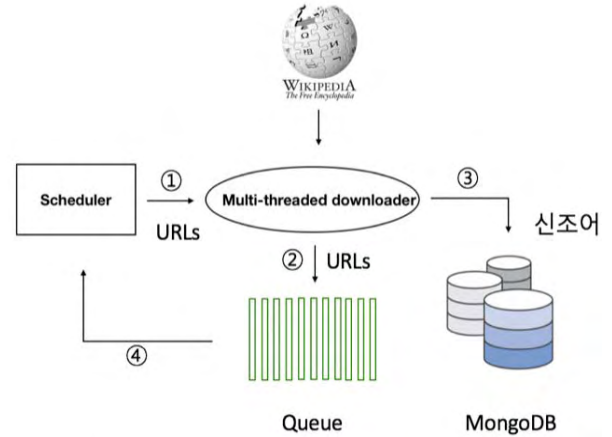
(그림 1)의 ①에서는 사용자가 어플리케이션(카카오톡)과 상호작용한다. ②에서는 어플리케이션으로부터 서버가 사용자의 입력을 전달받는다. ③에서는 서버가 전달받은 텍스트에 대해서 자연어 처리를 하기 위해 Watson Assistant API를 호출한다. 호출의 결과로 사용자의 의도와 핵심어를 전달받는다. ④에서는 서버가 핵심어에 대한 정보를 데이터베이스에서 가져와 사용자에게 출력한다.

### 3.1 신조어 데이터 수집 및 저장

본 논문에서 제안하는 신조어 챗봇 시스템은 웹 크롤링을 이용하여 데이터베이스 최신성을 보장한다. (그림 2)는 웹 크롤링 방식을 이용하여 데이터베이스를 최신의 데

이터로 갱신하는 과정을 나타낸 것이다.

웹 크롤링은 Multi-thread downloader를 기반으로 동작한다. Multi-thread란 하나의 독립적인 실행 프로그램 내에서 여러 개의 세부 실행 단위가 병행적으로 처리되는 처리 방식을 의미한다.



(그림 2) 데이터베이스 갱신을 위한 웹 크롤러 동작 과정

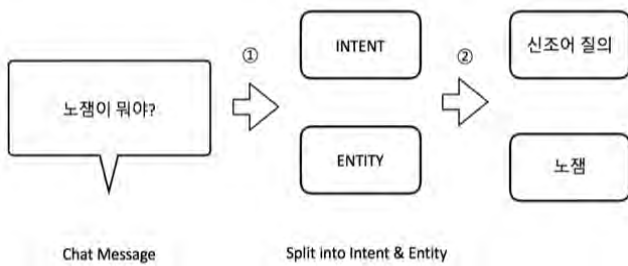
(그림 2)의 ①에서는 주기적으로 약속된 시간에 Scheduler가 동작하여 Thread를 동작시킨다. ②에서는 Thread가 ①에서 전달 받은 URL을 하나의 Queue에 저장하는 작업을 한다. ③에서는 Thread가 Queue에 저장된 URL을 추출하여 해당하는 웹의 데이터를 파싱하여 변화된 데이터를 확인한다. 변화된 데이터가 발견되면 이를 데이터베이스에 갱신한다. ④에서는 Queue에 저장된 URL이 없으면 다음 Scheduler의 동작을 기다리게 된다.

본 논문의 신조어 챗봇 시스템은 데이터베이스가 최신의 데이터를 가지는 것을 보장하기 위해 웹 크롤러를 사용한다. 웹 크롤러를 통해 위키백과에 새로운 신조어의 개념과 의미가 추가되면 이를 데이터베이스에 갱신하는 작업을 수행한다. 이를 통해, 사용자는 신조어 챗봇 시스템에 질의를 하였을 때, 최신의 데이터를 제공받을 수 있다.

### 3.2 사용자 의도 및 핵심어 파악

본 논문에서 제안하는 신조어 챗봇 시스템에서는 사용자의 요청이 들어오면 사용자의 의도(Intent)와 핵심어(Entity)를 파악하고자 Watson Assistant API를 사용한다. Intent란 챗봇과 대화하는 사용자의 의도를 말하고, Entity란 입력문장에 포함되어 있는 중요 키워드를 의미한다. 사용자의 요청에서 추출된 Intent와 Entity를 통해 시스템이 결과를 출력하게 된다.

사용자의 입력 문장에서 Intent와 Entity를 파악하는 방법을 도식화 한 것은 (그림 3)과 같다. Chat Message는 사용자로부터 시스템이 수신한 메시지를 나타낸다. ①에서는 API 호출을 통해 Intent와 Entity를 추출한다. ②에서는 추출된 결과를 시스템에 전달한다.



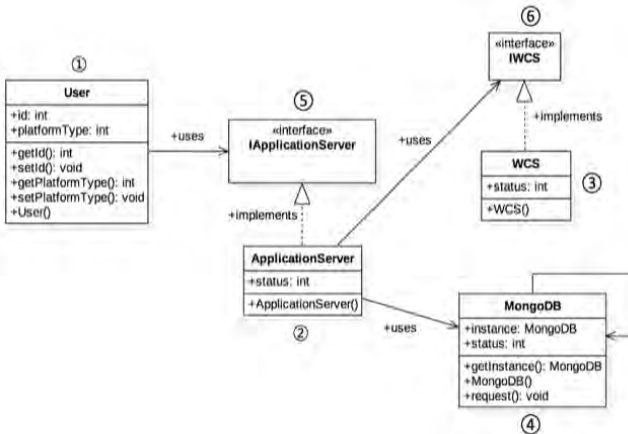
(그림 3) 질의로부터 Intent / Entity 추출

Watson Assistant API를 통해 어플리케이션에 자연어 처리 기능을 추가하여 신조어 챗봇과 일반 사용자와의 상호작용을 가능하게 한다. 자연어 처리 기술은 콘텐츠로부터 개념, 엔티티, 키워드, 카테고리, 감성등과 같은 메타데이터를 추출하여 텍스트를 분석하는 것이다.

#### 4. 설계 및 구현 결과

##### 4.1 시스템 설계

본 논문에서 제안하는 신조어 챗봇 시스템의 논리적인 구조인 클래스 다이어그램(Class Diagram)은 (그림 4)와 같다.



(그림 4) 전체 시스템 클래스 다이어그램

클래스 다이어그램은 신조어 챗봇 시스템의 사용자를 나타내는 ①User 클래스, 서버를 나타내는 ②ApplicationServer 클래스, Watson Assistant를 나타내는 ③WCS(Watson Conversation Service) 클래스, 데이터베이스를 나타내는 ④MongoDB 클래스로 구성된다. ⑤IApplicationServer와 ⑥IWCS는 각각 ApplicationServer 클래스와 WCS 클래스의 인터페이스 역할을 한다.

User 클래스로부터 ApplicationServer 클래스에 입력이 전달되면, ApplicationServer는 WCS에 전달된 데이터의 Intent와 Entity를 요청한다. WCS로부터 결과를 받으면 Database로부터 해당 데이터를 추출하여 User에게 제공한다.

##### 4.2 데이터베이스 설계

본 논문에서는 제안하는 신조어 챗봇 시스템은 사용자의 질의에 따른 신조어의 의미를 출력해야 한다. 따라서 시스템은 미리 구조화된 형태의 데이터로써 신조어의 의미에 대해 저장된 목록을 가지고 있어야 한다. 효율적인 데이터 목록 관리를 위해 데이터베이스 설계가 필요하다. 데이터베이스로는 NoSQL기반의 MongoDB를 사용한다.

MongoDB는 JSON과 같은 동적 스키마형 문서들을 선호함으로써 전통적인 관계형 데이터베이스와는 차별화된다. MongoDB는 데이터 단순 검색 및 추가 작업이 간단해서, 빅데이터를 활용한 웹 어플리케이션(Web Application)에 자주 활용된다. 일반적인 관계형 데이터베이스에서 사용하는 행(Row)이라는 개념보다 유연한 모델인 Document를 사용하는 데이터베이스이다. 복잡한 계층관계를 하나의 레코드(Record)에 입력 및 표현할 수 있다.

본 논문에서 제안하는 시스템 내부의 데이터베이스에 저장하는 신조어 Document의 JSON 형식은 (그림 5)와 같다. ①은 데이터들을 구분할 수 있는 고유 키 값이다. ②는 신조어의 이름, ③은 신조어의 의미, ④는 신조어의 예시를 나타낸다. ⑤는 Document가 생성된 날짜, ⑥은 수정된 날짜이다.

```

1 {
2   ① "_id" : ObjectId("5b9e724f5c4feb398cb98fa1"),
3   ② "name" : "뿌먹",
4   ③ "meaning" : "뿌러서 먹다",
5   ④ "example" : "담수육은 뿌먹해야 맛있지!",
6   ⑤ "createdAt" : "2018-03-10T11:54:57.882+0000",
7   ⑥ "updatedAt" : "2018-03-10T11:54:57.882+0000"
8 }
9

```

(그림 5) '뿌먹' 신조어 Document의 JSON 형식

##### 4.3 구현 결과

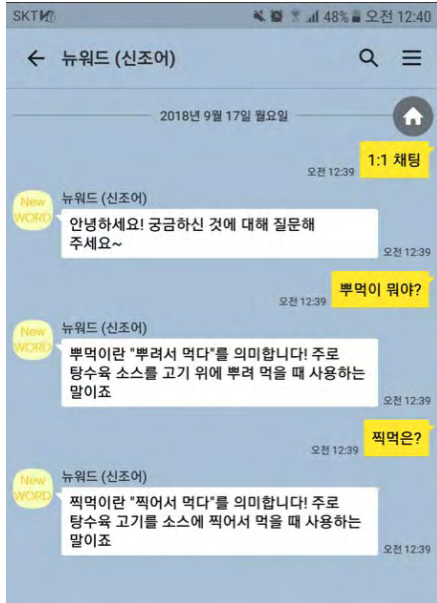
본 논문에서는 웹 크롤러, Watson Assistant, Node.js를 활용하여 신조어 챗봇 시스템을 구현하였다.

(그림 6)은 '뿌먹'과 '찍먹'이라는 신조어를 사용자가 질의한 결과를 출력한 화면이다. 신조어 챗봇 시스템을 통해 사용자가 요청한 신조어에 대한 의미를 출력해주는 것을 확인할 수 있다.

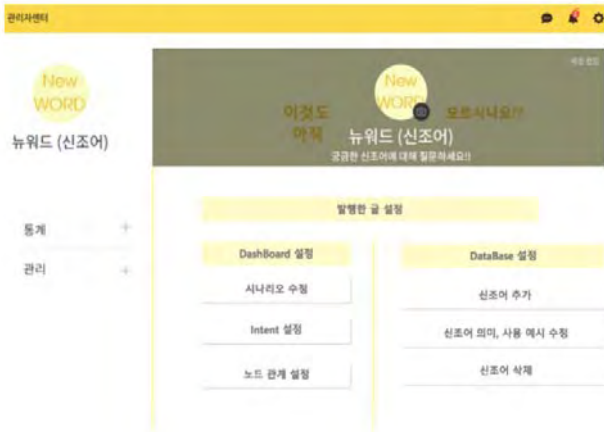
(그림 7)은 신조어 챗봇과 데이터베이스를 관리하기 위한 관리자 웹 화면이다. 데이터베이스에 웹 크롤링을 통해서 데이터를 추가할 수도 있지만, 관리자 웹을 통해 직접 신조어를 추가 할 수 있게 함으로써 유동적인 데이터 관리를 가능하게 하였다. 또한 사용자의 요청을 분석한 결과를 통계치로 나타내고, 사용자의 요구를 분석할 수 있는 시스템을 구축하였다.

참고문헌

- [1] 홍금원, 이연수, 김민정, 이승욱, 이주영, 임해창, “한국어 모바일 대화형 에이전트 시스템”, 한국컴퓨터정보학회 논문지, Vol. 13, No. 6, pp. 263-271, 2008.11.
- [2] 이대근, 나승유, “중소상인을 위한 AI 챗봇 플랫폼의 설계 및 구현”, 한국전자통신학회 논문지, Vol. 13, No. 3, pp. 561-570, 2018.06.
- [3] 최예림, 김관호, “인공지능 개요 및 적용 사례”, ie 매거진, Vol. 23, No. 2, pp. 23-29, 2016.06.
- [4] 채터봇, <https://ko.wikipedia.org/wiki/채터봇>
- [5] 오세욱, “미디어로서의 봇”, 한국언론정보학회 한국언론정보학보 Vol. 79, pp. 70-103, 2016.10.
- [6] 김보경, 김현아, 심채은, 정진우, “소셜 챗봇을 활용한 클라우드 소싱 기반 질의응답 시스템의 설계 및 구현”, 한국정보기술학회 논문지, Vol. 15, No. 11, pp. 125-138, 2017.11.
- [7] IBM Watson Assistant, <https://www.ibm.com/watson>



(그림 6) 신조어 ‘뿌먹’, ‘찍먹’의 검색 및 결과 화면



(그림 7) 관리자 웹 화면

5. 결론

본 논문에서는 Watson Assistant 기반의 신조어 챗봇 시스템을 구현하였다. 신조어 검색 및 학습을 쉽게 할 수 있는 방법으로 웹 크롤러, Watson Assistant API, Node.js를 활용한 챗봇 시스템을 구현하였다.

웹 크롤러는 위키백과의 신조어 목록을 모니터링하고 새로운 신조어가 추가되면, 데이터베이스에 신조어를 갱신함으로써 데이터베이스의 최신성을 보장한다. 사용자가 신조어 챗봇 시스템에 신조어를 질의하면, 시스템은 사용자의 의도를 정확하게 파악하기 위해 Watson Assistant를 통해 컨텍스트를 분석한다. 그리고 Node.js로 구성된 서버를 통해 결과값을 출력하여 사용자에게 제공한다.

본 논문에서 제안하는 시스템을 통해 신조어의 의미를 쉽게 파악할 것이라고 기대한다. 향후 신조어 외에 다른 언어를 지원하는 챗봇 시스템을 구현할 예정이다.