

# 인공지능 기반의 챗봇 시스템 기술 동향

박준호·윤경일·민성태 (㈜솔루게이트)

목 차	1. 서 론
	2. 챗봇 기술 동향
	3. 챗봇 적용 동향
	4. 인공지능 기반 챗봇 구축 방안
	5. 결 론

## 1. 서 론

최근 공공기관, 민간(금융)에서 고객 서비스를 위해 챗봇을 도입하고 있다. 챗봇 서비스는 별도의 플랫폼을 사용하지 않고 기존의 SNS(Social Network Service) 등과 연계하여 진행하고 있으며, 카카오톡을 통해 많은 서비스를 제공하고 있다.

공공기관은 챗봇 서비스를 통해 민원 업무를 콜센터 비중을 줄이기 위해 도입을 하고 있다. 금융기관은 고객의 서비스 만족도를 위해 챗봇을 도입하고 있으며, 이를 통해 상담사의 업무를 절감하고 있다. 또한 음성인식을 통한 가상상담(Virtual Agent) 서비스를 구축하고 있다.

챗봇은 사람과의 문자 대화를 통해 질문에 알맞은 답이나 각종 연관 정보를 제공하는 “인공지능(Artificial Intelligence)기반의 커뮤니케이션 소프트웨어” 로정의할 수 있다.

챗봇은 사람들이 필요로 하는 서비스와 데이

터를 적시에 찾아주는 역할로 정의하고 있으며, MS CEO 사티나 나델라는 “ 봇이 앱을 대체하고 디지털 개인비서가 새로운 메타 앱이 되고, 컴퓨터와 사람사이의 모든 상호작용에 AI(Artificial Intelligence)가 침투할 것”이라고 했다.

챗봇은 명령 방식에 따라 3가지로 구분이 된다. 표 1과 같이 챗봇, 음성인식 봇, 개인비서로 구분된다.

현재 국내에서는 챗봇 서비스를 가장 많이 이용하고 있다. 행안부의 민원 24, 대구시의 뚜벳, 강남구의 강남봇 등이 챗봇 서비스를 이용한 사례이다[4].

챗봇을 구축을 하기 위해 필요한 기술 요소는 서버(Back-end), 미들웨어(Middleware), 클라이언트(Front-end)로 구성이 된다.

서버(Back-end)는 사용자의 질의에 적절한 답변을 제공하기 위해 자연어 처리, 상황 인식, 대

〈표 1〉 명령에 방식에 따라 챗봇, 음성인식, 개인비서 구분

	챗봇	음성인식봇	개인비서
명령 방식	텍스트 메시지	음성	텍스트 메시지+음성 +검색 패턴, 위치, 사용패턴
핵심 서비스	Facebook Messenger, KiK BotShop	Amazon Echo, Apple Siri	Google Now, MS Cortana

화 처리, 지식 처리, 빅데이터 기술을 사용한다.

미들웨어(Middleware)는 막대한 메시지 트래픽을 처리하기 위해 하둡(Hadoop)과 같은 분산 컴퓨팅 환경에 적합한 기능을 가진 소프트웨어를 사용한다.

클라이언트(Front-end)는 사용자가 직접 화면으로 확인할 수 있는 모바일 메신저 앱 및 PC 기반의 메신저를 통해 사용하게 되며 하이브리드 앱 기술을 많이 사용한다. 최근에는 시나리오를 설계할 수 있는 기능을 클라이언트에 서 개발하여 제공하고 있다.

## 2. 챗봇 기술 동향

챗봇은 텍스트 기반으로 동작이 되는 서비스이다. 챗봇을 구축하기 위해서는 자연어 처리, 대화 처리, 지식처리 기술이 중요하다.

### 2.1 자연어 처리

인간이 사용하는 언어를 컴퓨터가 이해할 수 있도록 하기 위한 기술이 자연어 처리 기술이다 [3]. 이를 위해 기본적으로 NLU를 진행하고 이를 바탕으로 NLP를 진행한다. 컴퓨터가 사람의 용어를 이해하기 위해서는 인간과 인간의 언어로 이해하고 소통할 수 있어야 한다. 컴퓨터를 이용하여 인간의 언어를 분석하고 이해하는 기술이 자연어 처리 기술이다. 자연어 처리의 절차

는 일반적으로 형태소 분석기, 구문 분석, 의미 분석, 개체명 인식, 감정 분석, 화행 분석 등이 있다[1].

형태소 분석(Morphological Analysis)은 입력된 문자열을 분석하여 형태소 단위로 분석하는 과정이다. 형태소는 의미를 가지는 요소로서는 더 이상 분해할 수 없는 최소 단위의 언어 요소이다. 구문 분석(Syntactic analysis)은 문장이나 구절을 만드는 규칙인 구문에 따라 문장이 가지는 구문 구조를 분석하여, 문장을 구성하는 문자열들이 문장에서 어떤 역할을 하는지 결정한다. 형태소 분석기는 자체적으로 구현하여 사용하거나, 오픈 소스 기반의 코모란, 은전한잎 등을 사용한다. 한국어는 띄어쓰기에 문제가 있어 Spell Check는 반드시 필요하다. 또한 오픈소스의 경우에는 지속적인 업데이트를 보장받기가 어렵다.

### 2.2 대화 처리

대화 처리는 기본적으로 세 가지 관점에서 분류를 할 수 있다[2]. 첫번째로 발화의 관점이다. 대화를 주도하는 발화자의 의도를 인식하기 위해 현재 대화의 상황을 파악하여, 일련의 정보 수집을 거쳐 발화자의 의도, 욕구 등을 인식하는 방식이다. 두 번째는 사용자와 시스템 중 어느 한쪽이 대화의 주도권을 가지고 있는지의 문제를 판단하는 문제이다. 이는 질의 응답 시스템, 상담 시스템 등으로 분류된다. 세 번째, 대화에



(그림 1) 대화 처리 프로세스

사용되는 커뮤니케이션 방법에 대한 관점이다. 음성이나 제스처, 시선, 표정, 문자 등 다양한 정보 전달 방법에 대해 분석하는 것이다.

이에 일반적으로 대화 처리는 그림 1과 같이 입력문에 해당하는 사용자 발화를 이해하는 이해부, 응답문에 해당하는 시스템 발화의 생성부, 그리고 대화 관리의 세 부분으로 구성된다[1].

최근에는 그림 2와 같이 챗봇 시나리오 생성을 쉽게 작성하기 위해 JSON 기반으로 구성된 시나리오 편집기를 제공하고 있다.



(그림 2) 챗봇 시나리오

### 2.3 지식 처리

지식 처리는 사람이 챗봇에 질의한 사항에 답변을 할 수 있도록 구성된 것을 말한다. 지식이란 정보를 취득하고 분석하여 얻은 결과물에 대해 사람이 이해하는 것을 말한다.

인공지능에서 지식은 “경험이나 교육을 통해 얻어진 전문적인 이해와 체계화된 문제 해결 능력”, “어떤 주제나 분야에 대한 이론적 또는 실제적인 이해, 또는 현재 알려진 사실과 정보의 모음”을 의미한다. 문제 해결에 이런 지식을 사용하려면 컴퓨터가 쉽게 처리할 수 있도록 정형화된 형태로 표현이 되어야 한다.

이를 위해 개발된 것이 온톨로지(Ontology)이다. 온톨로지는 컴퓨터가 특정 영역의 지식을 공유하고 재사용할 수 있도록 해당 영역의 개념과 관계를 나타내는 어휘를 정의하고 이를 이용해 지식을 표현해 놓은 것을 말한다[1].

### 3. 챗봇 적용 동향

챗봇은 국민과 24시간, 실시간으로 소통하며 개인 맞춤형 서비스가 가능하여 공공부문의 민

원. 행정 분야에서 각광을 받으며 빠르게 확산되고 있다. 국내는 법제처, 대구시, 경기도 등 정부. 지자체뿐만 아니라 공공기관에서까지 챗봇 사업을 추진 중이다. 하지만 아직까지 공공부문의 챗봇은 서비스의 범위, 응답 수준 등이 제한적이며 대국민 활용도와 만족도가 높지 않은 실정이다.

### 3.1 법무부 : 인공지능 기반 법률비서

법무부는 국민들이 언제 어디서든 스마트폰과 PC를 이용하여 생활 법률 관련 궁금증을 해소할 수 있도록 서비스를 개발하여 운영중이다. 웹페이지(talk.laworder.go.kr)와 메신저 앱인 카카오톡 플러스친구를 통해 인공지능 기반의 생활 법률 챗봇 서비스를 제공하고 있다.

법무부 챗봇은 1세대와 2세대로 구분이 된다. 1세대 챗봇의 경우에는 자체 학습이 없었지만, 2세대 챗봇에는 자체학습 기능이 추가되었다(표 2 참고) [4].

### 3.2 행정안전부 : 정부 24(민원 24) 챗봇

행정안전부는 정부 24에 발생하는 민원 서비스 및 콜센터 문의 내용을 분석하여 이를 민원인에게 효율적으로 서비스를 할 수 있는 챗봇 서비

스를 도입했다. 민원인이 가장 많이 사용하는 민원 서식을 도출하여 이를 쉽게 이용할 수 있도록 구성했다. 본 사업은 1차 사전 설계를 통해 적용 가능성을 평가하고 2단계 사업에서 확산 및 적용하는 사업이다.

챗봇 구축은 대화 처리 방식을 이용하여 적용하였으며, 이를 위해 별도의 시나리오 편집기를 사용했다. 또한 다양한 질의에 대한 답변 제시를 위해 용어 사전 등을 추가로 구축하여 일부 지식 처리 기반도 구현한 사례이다.

### 3.3 싱가포르 : 공공기관 웹사이트 안내를 위한 챗봇(가상비서)

싱가포르는 민원인의 헬프 데스크에 대한 의존도를 줄이고, 자주 하는 질문에 대해서는 간편하게 답변을 얻을 수 있도록 하기 위해 챗봇 서비스를 도입했다. 싱가포르는 교육부, 토지국 등 30개 이상의 정부기관 웹 사이트에, 자연어 처리 엔진이 적용된 챗봇을 구축하여 보다 간단한 질의 응답 서비스를 구축했다. 이를 통해 특정한 질의를 위해 다른 기관의 웹 사이트를 방문할 필요 없이 어떤 사이트에서든 질문과 답변이 가능하다.

싱가포르는 챗봇 구축을 통해 정부에 대한 편

〈표 2〉 법무부 비버 발전 단계

구분	1세대	2세대
접근 방법	인터넷 버비 웹 사이트	카카오톡 추가
문답지식 분야	부동산임대차, 임금, 해고	상속분야 추가
문답지식 콘텐츠	38,894건	43,651건 추가
카드 뉴스 및 오디오 카툰 등	154건	53건 추가(범죄피해예방 등)
공공데이터 개방	없음	공공데이터포털 오픈
자체학습 기능	없음	있음

의성 및접근성이 향상되고, 기관 간 데이터 연계를 필요한 서비스를 여러 기관을 방문하지 않아도 끊임없이(Seamless) 수혜가 가능하다[4].

### 3.4 탈라(Talla) : 업무지원을 위한 지능형 비서

챗봇은 공공기관의 서비스이외에 내부의 개인 비서로 활용을 하고 있다.

탈라는 직원들의 업무생산성 향상을 위해 단순 업무지원에 인공지능 기반 챗봇 서비스이다. 이미지 인식, 자동거래, 자연어 처리 및 이해, 감속학습 등 인공지능 기술이 적용된 탈라(Talla)와 대화를 통해 직원들의 다양한 업무를 지원한다. 또한 일정관리, 고객관리, 문서 검색 및 관리 업무 이외에 예측, 자동화, 분류 등 고도화된 업무 수행도 가능하다. 탈라(Talla)는 똑똑한 기적

기반 업무지원으로, 단순 지원업무를 대체하고 기존 인력의 타 업무 수행 등 업무효율성 향상을 기대할 수 있다. 현재 탈라(Talla)는 복잡한 데이터 분류, 작업 자동화, 추세 예측에 사용되고 있다<표 3>[4].

## 4. 인공지능 기반 챗봇 구축 방안

인공지능 기반 챗봇 구축은 총 3단계로 진행된다. 1단계로 챗봇 서비스, 2단계 지능형 비서, 3단계 감성 비서이다. 최근에는 챗봇 서비스 이후 모델인 지능형 비서 구축으로 진행하고 있다. <표 4>는 챗봇 서비스 발전 단계를 정리한 표이다[4].

최근에는 1단계의 챗봇 서비스와 2단계의 지능형 비서를 구축하고 있다. 초기 개방형 형태로 서비스된 모델에서 내부 직원을 위한 폐쇄형 구

<표 3> 인공지능 탈라(Talla)의 주요 사용 분야

분 야	주요 내용
복잡한 데이터 분류	텍스트, 이미지, 문서, 데이터 파일 및 소프트웨어 코드를 안정적으로 인식하고 분류
작업 자동화	복잡한 논리흐름도(Logic Flow Chart)를 자체 생성하여 다양한 작업 상황에 대해 구체적 조치 가능
추세 예측	현재 정보를 기반으로 결과를 예측할 수 있는 분석 모델 자체 생성

<표 4> 챗봇 서비스 발전 단계

구분 단계	1단계 챗봇 서비스	2단계 지능형 비서	3단계 감성비서
제공방식	텍스트, 음성	텍스트, 음성, 시각자료	텍스트, 음성, 시각자료, 행동 인지
입력방식	폐쇄형, 개발형 일부	폐쇄형 일부, 개방형	개발형, 폐쇄형 복합 사용
주요기술	패턴 매칭, 키워드 및 연관어 추출 등	딥러닝,머신러닝, 자연어 처리, 기타 신기술 융합 등	감성인지기술, 데이터 정형화 기술
내용	- 학습된 내용에 대한 질의 응답. - 사용자와 단순한 형태의 소통 - 검색을 통한 결과 제공	- 사용자의 패턴, 상황을 고려한 개인 맞춤형 서비스 제공 - 간단한 업무 처리	- 감정 교류를 통한 서비스 및 각종 서비스에 대한 선제적 대응

조를 통해 서비스 만족도를 향상시키고 있다.

전산실의 ITSM(IT Service Management, IT 서비스 관리)의 경우 24시간 고객 서비스 대응을 위해 Client 장애처리 및 지원 서비스, Network 장애처리 및 지원, 소프트웨어 설치요청 및 지원 서비스, 보안 서비스 지원 및 지원 등에 대해 구축을 하고 있다. 이를 위해 사용자의 의도 구문을 파악하여 질문에 답변을 제공할 수 있는 모델을 요구하고 있다.

인공지능 기반 챗봇 구축은 총 3가지 형태로 구분된다.

첫번째로 자체 기술을 보유한 경우, 두번째는 오픈소스를 활용하여 구축하는 경우, 마지막 세번째는 클라우드 서비스를 활용하여 구축하는 방법이다.

인공지능 기반 챗봇을 자체적으로 구축하기 위해서는 기본적으로 자연어 처리 기술을 확보하고 있어야 한다. 또한, 대량의 데이터를 분석하기 위해 기존의 빅데이터 분석 기법을 적용하면 효과적으로 처리가 가능하다. 수집된 데이터에 대한 정보를 기반으로 분류하여 이를 사전에 분석을 하면 각 단어들 간의 연관 분석등이 가능하다[5].

자연어 처리 분석을 위해서는 형태소 분석기, 구문 분석기, 개체명 인식기는 보유를 해야 한다. 또한 딥러닝을 이용할 경우 BERT, n-gram, LSTM 등의 알고리즘을 자유롭게 사용이 가능해야 한다. 자체적인 기술을 보유한 경우에는 폐쇄적인 전산 환경에서 구축 시 매우 용이하다.

두번째 오픈소스 기반으로 구축하는 경우에는 서버(Back-end)는 수정이 불가하다. 사용자가 화면으로 보여지는 클라이언트(Front-end) 부문만 개발이 가능하다. 또한 오픈소스로 개발 후 상용으로 전환되는 경우가 있어 제품을 납품했을 경우에 문제가 발생할 수 있다. 제품의 안정성 및

업그레이드에 대한 명확한 일정이 없어 버그 발생 시 커뮤니티에서 지원을 하지 않을 경우 버그 해결에 많은 시간이 소요된다.

마지막으로 클라우드 환경에서 제공해주는 기능을 활용하여 챗봇을 구축하는 경우 기본적으로 1-2일 사이에 구축과 서비스 제공이 가능하다. AWS, MS Azure, Google 등에서 제공하며 별도의 서비스 금액이 발생한다. 클라우드 환경은 공개적인 서비스 도메인에서는 사용이 가능하다. 예로 들어 홈쇼핑, 여행사 등이 대표적이다.

최근에는 챗봇 서비스 만족도 개선을 위해 자동으로 학습할 수 있는 알고리즘을 개발하고 있다.

한국어는 교착어로 다양한 단어가 생성되며, 동음이의어 등이 발생하여 초기 구축이후 발생하는 단어에 대한 학습이 필요하다. 이를 개선하기 위한 자동학습 엔진을 보유한 업체가 생겨나고 있다.

마지막으로 명확한 답변을 위해 초기 질문에 대해 딥러닝 학습을 통해 각 서비스 도메인별 유사 질문을 자동으로 생성하여 학습하는 모델을 구축하고 있다.

## 5. 결 론

앞에서 언급한바와 같이 챗봇은 여러 다양한 기술을 통해 서비스되고 있다.

최근에는 시나리오 기반으로 챗봇을 설계하는 경우가 많아 지고 있다. 시나리오 모델은 설계 및 개발에 대한 이점은 분명히 있다.

하지만 시나리오의 개수가 많아지면 관리적인 측면과 유지보수 측면에 문제가 발생한다. 또한 사용자가 명확한 질문이 아닌 경우 시나리오에서는 답변을 제공하지 못하는 현상이 발생한다.

이를 개선하기 위해 사용자의 질의에 대한 답





(그림 3) 챗봇 아키텍처

변을 못 찾는 경우 질의 키워드 분석을 통해 해당되는 사이트로 연결하거나 미답변 리스트를 통해 별도의 학습을 진행한다. 챗봇이 제공하는 답변에 대해 사용자 만족도 분석을 통해 만족이 높은 답변에 대해 별도의 강화 학습을 진행한다.

챗봇은 다양한 서비스 모델에 적용하기 위해서는 그림 3과 같이 아키텍처가 정의되어야 한다.

챗봇 솔루션은 챗봇을 만들기 위한 빌더와 자연어 처리 기술이 구성되어야 한다.

딥러닝은 챗봇에서 가장 많이 사용하는 알고리즘으로 구성되며, 데이터는 다양한 형태로 연계 구성된다. 한국어 특성을 반영하여 지속적인 학습이 필요하며 이를 위한 자동 학습 모델이 설계되어야 한다.

마지막으로 이를 통해 고객 환경에 맞는 챗봇 서비스를 구현할 것으로 기대한다.

### 참 고 문 헌

- [1] 류기동, "인공지능 기반 컨택센터 시스템 연구", 서울과학기술대학교 IT정책전문대학원, 2019
- [2] 고창수, 이은희, 김효용, 이성우, 권정현, 정연주, 이현주, "인공지능 대화시스템 연구", 지식과 교양, 서울
- [3] 이건명, "튜닝 테스트에서 딥러닝까지 인공지능", 생능출판사
- [4] 서교리, "인공지능 기반 챗봇 서비스의 국내외 동향분석 및 발전 전망", 한국정보화진흥원, 2018-2호
- [5] Joonho park, Jinho Park, Nam Young Lee, "Design of proposal for Big Data Analysis Model", AWICE, 2016

저 자 약 력



**박 준 호**

이메일 : 108.joonho.park@gmail.com

- 2002년 동양대학교 이동통신공학과 (학사)
- 2015년 한국방송통신대학교 대학원 정보과학과 (석사)
- 2018년 숭실대학교 일반대학원 컴퓨터공학과 (박사수료)
- 2013년-2018년 대보정보통신(주) ICT1사업본부 과장
- 2018년-현재 (주)솔루게이트 전략기획실 실장
- 관심분야: 빅데이터 분석, 클라우드 설계, 소프트웨어 개발 방법론



**민 성 태**

이메일 : david@solugate.com

- 1991년 동국대학교 전자계산학과 (학사)
- 1991년-2000년 (주)LG 전자 / 정보화추진팀 팀장
- 2014년~현재 (주)솔루게이트 대표이사
- 관심분야: 지능형 컨택센터, 가상상담, 음성인식, 성문 인식



**윤 경 일**

이메일 : pofour1@gmail.com

- 2013년 신구대학 컴퓨터정보처리 (전문학사)
- 2018년 독학학위제 컴퓨터과학과 (학사)
- 2019년~현재 한국방송통신대학교 대학원 정보공학과 (석사과정)
- 관심분야: 임베디드 시스템, 자연어 처리, 네트워크 시스템