JKIICE Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 25, No. 7: 877~883, Jul. 2021

RCMS에 활용하기 위한 인공지능 기반 챗봇 시스템

김용국[1] · 김수진[2] · 정회경[3*]

Artificial intelligence-based chatbot system for use in RCMS

Yongkuk Kim[1] · Sujin Kim[2] · Hoekyung Jung[3*]

[1]Senior Principal Researcher, Digital Innovation Department, KEIT, Daejeon, 35262 Korea
[2]Graduate Student, Department of Computer Engineering, Paichai University, Daejeon, 35345 Korea
[3*]Professor, Department of Computer Engineering, Paichai University, Daejeon, 35345 Korea

요 약

인공지능 기술은 제조 로봇, 인공지능 스피커, 로봇 청소기 등 산업 및 스마트홈 분야에서 다양하게 사용되고 있다. 본 논문에서는 RCMS(Real-time Cash Management System)에서 활용하기 위한 인공지능 기반 1:1 챗봇(chatbot) 시스템을 설계 및 구현하였다. 구현한 RCMS 챗봇은 기존 온라인 게시판의 1만 3천 5백여건의 질의응답을 기반으로 연구비 사용, 시스템 사용법 등 9개 영역에 총 210개의 질의시나리오로 구축하였다. 챗봇은 부족한 상담인원 문제를 해소하고, 근무시간 이후에 연구자의 문의에 대응하여 사용자의 만족도를 제고 할 것으로 예상되며, 연구자의 상담 문의가 가장 많았던 사용비목에 대한 추천 서비스는 상담건수를 감소시켜 다른 상담문의에 대한 답변의 질적 수준 향상이 기대된다.

ABSTRACT

Artificial intelligence technology is widely used in industrial and smart home fields such as manufacturing robots, artificial intelligence speakers, and robot vacuum cleaners. In this paper, we designed and implemented a 1:1 chatbot system based on artificial intelligence for use in RCMS (Real-time Cash Management System). The RCMS chatbot implemented in this paper was constructed with a total of 210 query scenarios in nine areas, including research expenses and system usage, based on 13,500 questions and answers from existing online bulletin boards. The chatbot is expected to solve the problem of insufficient number of counselors and to increase user satisfaction by responding to the researcher's inquiries after working hours, and the recommendation service for the cost of use, which had the most inquiries from researchers, reduces the number of consultations. It is expected to improve the quality of answers to other counseling inquiries.

키워드 : 연구비관리 시스템, 인공지능, 질의어, 챗봇

Keywords : Artificial intelligence, Chatbot, Query language, Real-time cash management system

Received 3 May 2021, **Revised** 5 May 2021, **Accepted** 3 June 2021

*** Corresponding Author** Hoekyung Jung(E-mail:hkjung@pcu.ac.kr, Tel:+82-42-520-5640)
Professor, Department of Computer Engineering, Paichai University, Daejeon, 35345 Korea

Open Access **http://doi.org/10.6109/jkiice.2021.25.7.877** print ISSN: 2234-4772 online ISSN: 2288-4165

ⓒThis is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(http://creativecommons.org/li-censes/ by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright ⓒ The Korea Institute of Information and Communication Engineering.

Ⅰ. 서 론

최근 인공지능 기술이 다양하게 활용됨에 따라 본 논문에서는 연구비관리시스템이 보유한 관리 빅데이터와 최근의 인공지능 기술을 활용하여 콜센터에서 처리한 질의응답 데이터를 학습시켜 인공지능 기반 챗봇을 제안한다. 연구비관리시스템의 1:1 챗봇 서비스의 구현은 시스템의 Q&A 게시판 문의 내역과 상담원의 온라인 상담 기록을 활용하여 질의 유형을 분류하고 이를 시나리오로 관리하며 재학습이 가능한 형태의 인공지능 기반의 챗봇 서비스를 구현한다. 이를 활용하여 야간 및 공휴일 등 근무외 시간에 연구자의 문의에 응답할 수 있을 것이다. 본 논문을 통해 기존 연구비관리시스템에 빅데이터를 활용한 연구비관리 챗봇 기능이 추가된 인공지능 기반의 연구비관리시스템은 연구비관리시스템의 활용성 및 국가연구개발예산의 효율성을 높이는데 기여할 것으로 판단된다.

Ⅱ. 챗봇(Chatbot)

인공지능 활용분야 중에서 가장 광범위하게 사용되고 있는 서비스가 챗봇이다. 챗봇은 인간이 하던 고객대응 업무를 인공지능 컴퓨터가 대신하는 것으로, 가장 핵심적인 기능은 고객과 대화하는 능력이다. 채팅(Chatting)과 로봇(Robot)의 합성어로 인간과의 커뮤니케이션을 통해서 질문에 적합한 답을 제시하거나 또는 명령을 받아서 처리하는 인공지능 기반의 소프트웨어이다[1].

정부 및 공공기관은 국민에 대한 정부 대응성을 높여서 국민에게 더 나은 행정서비스를 제공하기 위해서 다양한 형태의 챗봇 서비스를 개발하여 제공하고 있다[2].

챗봇은 자연어 처리, 대화 처리 및 지식 처리 기술이 중요하다. 컴퓨터가 인간의 언어를 이해할 수 있도록 하는 기술이 자연어 처리 기술이다. 대화 처리 기술은 대화를 주도하는 발화자의 의도 인식을 통해 대화의 상황을 파악하고, 대화의 주도권을 판단하며, 대화에 사용하는 정보 전달 방법에 대해 분석하는 것이다[3]. 마지막으로 지식 처리 기술은 질의 내용에 대한 답변이 정형화된 형태로 만들어져, 이를 이용해서 지식을 표현하는 기술이다[4].

자연어 처리는 형태소 분석 및 구문 분석 등의 절차를 거친다. 문자열의 의미를 가지는 최소 단위의 언어 요소인 형태소로 나누어 분석하는 과정이 형태소 분석이고, 구문 분석은 문자열들이 어떤 역할을 하는지 알기 위해 문장이나 구절을 만드는 규칙에 따라 구문 구조를 분석하는 절차이다[5-6]. 자연어를 텍스트 분석하는데 딥러닝을 이용하는 기술은 챗봇 외에도 온라인 게임에서 욕설을 탐지하는데 사용되기도 한다[7]. 질의 의도를 분류하여 질의-응답 형태의 시나리오로 대응하는데, 시나리오 구조는 상담 챗봇에서 상당히 중요한 요소 중 하나이다[8].

Ⅲ. 챗봇 시스템 구축

연구비관리시스템의 사용방법과 연구비집행 처리에 대해 FAQ, 게시판 및 콜센터 상담사를 통해 고객 문의에 대응하고 있다. 이는 긴급한 문의에 대해 실시간 대응이 어렵고, 근무시간 외 대응이 불가능하며, 연구비규정의 변화 등 다양한 문의사항에 대한 숙련된 상담사 확보가 필요하다. 본 장에서는 이러한 문제점을 개선하기 위해서 인공지능 기술을 적용한 챗봇 기능을 갖춘 지능형 연구비관리시스템을 구현한다.

3.1. 지능형 연구비관리시스템 개선 범위

기존의 연구비관리시스템을 유지하고, 인공지능 기반의 챗봇 기능을 추가하여 사용자의 편의성을 향상시킨다. 현재 게시판 등록 및 답변 형태에서 인공지능 기반의 챗봇을 도입하여 실시간 고객의 문의에 자동 응답 시스템의 대응범위를 초과하는 고객문의에 대응하고자 한다.

챗봇은 기존 1:1 문의 게시판에 적용하여 사용자가 챗봇을 사용하거나 1:1 문의를 선택하여 사용하도록 한다. 챗봇을 통한 상담기록을 활용하여 재학습이 가능한 형태로 챗봇 모형을 구현한다. 또한, 질의의도 및 답변, 시나리오 관리 및 상담 통계 확인이 가능한 관리자 시스템도 구축한다.

1:1 문의 챗봇은 기계학습, 지식관리, 언어 전처리 분석을 수행하는 관리도구 부분과 채팅 어플리케이션과 답변을 API 형태로 제공하는 챗봇 서비스 부분으로 구성된다.

3.2. RCMS 챗봇 시스템

3.2.1. 사용자 질의 데이터 분석을 통한 대화흐름(Chatflow) 구성

챗봇을 구축하기 위해서 먼저 해야 할 일은 기존 사용자의 질의 데이터를 분석하는 것이다. 홈페이지의 FAQ 데이터 194개와 1:1문의 데이터 13,500개를 대상으로 질의내용 및 질의유형을 분석하였다. 1:1문의 데이터의 경우에 여러 의도를 가진 긴 문장의 글이 대부분이어서 하나의 의도를 가진 문장으로 요약하는데, 제목을 중심으로 진행한다.

정리된 데이터를 분류하기 위해서 필요한 대표영역(주제영역)은 RCMS 시스템의 메뉴 카테고리를 기반으로 도출하였다. 도출된 대표영역은 9개이며, 연구비 현황, 연구비 사용, 연구비 취소, 연구비 정산, 환경설정, 시스템 관련, 마이페이지, 비목증빙정보, 용어정의이다.

데이터를 도출한 9개의 대표영역으로 분류하고, 대표영역 안에서 다시 질의의도(intent)로 상세 분류하였다. 분류의 정합성을 높이기 위해서 1차 분류 후에 다시 2차 분류를 하였고, 1차 분류 결과 질의 의도는 총 365개였지만, 2차 분류에서 비슷한 질의의도를 통합하여 최종적으로 201개의 질의의도를 도출했다. 1차 분류 결과 도출한 질의의도에 대해서 1:1 문의에 대한 답변을 담당하는 콜센터 상담사와의 검토를 진행하였고, 일부 질의의도를 추가하기도 하였다. 가장 많은 질의의도를 가진 대표영역은 용어정의로 46개이며, 마이페이지 대표영역이 5개의 질의의도로 가장 적었다.

분류된 대표영역 및 질의의도별로 시나리오를 구성하였다. 시나리오는 대표질의-응답 세트로 구성되는데, 대표질의는 응답과의 호응이 좋고, 질의의도가 잘 표현되는 것으로 작성하였고, 응답은 제공 필요성이 높고 어떤 형태의 질의와도 호응이 잘 되는 내용으로 사용자의 가독성을 고려하여 작성하였다.

다음은 질의확장을 진행하였는데, 기존 1:1 문의 데이터를 참고하여 사용자가 주로 발화할 만한 다양한 형태의 질의를 질의의도별로 확장한 후 질의 내 주요 단어를 명사 단위로 추출하여 유사단어(Entity)를 구축하였다. 마지막으로 유사단어와 질의의도별 질의응답을 기반으로 대화흐름을 구성하였다.

이렇게 만들어진 대화흐름에 기반하여 챗봇은 사용자의 질의를 받았을 때 질의에 적합한 답변을 제시한다. 대화흐름은 새로운 대표영역이나 질의의도가 발생할 경우 지속적으로 추가되어져야 한다. 그림 1은 기존 데이터를 기반으로 대화흐름을 만드는 과정이다.

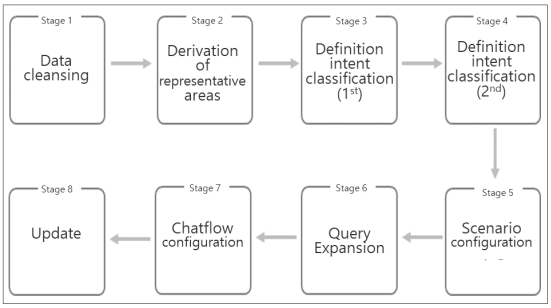

Fig. 1 Chatflow Creation Process

3.2.2. 사용자 질의 의도 분석

챗봇이 실제 사용자의 질의에 응답하기 위해서는 그림 2와 같이 사용자의 질의를 분석하여 질의 의도를 파악해야 한다. 사용자 의도의 파악은 구문 분석, 의도 분석 및 의도 분류의 순으로 진행된다.

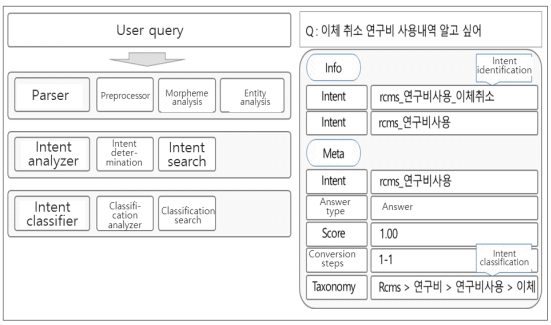

Fig. 2 Flowchart for user query intention classification

그림 2에서 사용자 질의가 들어오면 구문분석기가 전처리, 형태소 분석, 개체 분석을 수행하고, 이어서 의도분석기가 의도를 판별하며, 마지막으로 의도분류기가 도출된 질의의도를 분류한 것 중에서 적합한 질의의도 분류항목을 가져온다.

고객의 질의 의도를 이해하기 위해서 제일 처음으로 질의에 적용되는 프로세스는 구문 분석을 위한 형태소 분석이며, 사용된 형태소 분석기는 KoBRICK이다. 이 형태소 분석기를 활용하여, 문장 내 단어의 형태소를 분해하고 품사(Part-of-Speech)를 부착하였다.

그림 3을 보면 사용자의 질의가 입력되면, 전처리, 형

태소 분석 및 품사태깅, 후처리를 통해 질의어를 분석해 낸다. ‘RCMS의 담당자와 상담시간은 어떻게 되나요?’ 라는 질의를 받으면, 먼저 어절 단위인 ‘RCMS’, ‘담당자와’, ‘상담시간은’, ‘어떻게’, ‘되나요’로 나눈 후, 이 어절을 다시 단어와 품사 형태로 분류하는데, ‘RCMS’, ‘담당자’, ‘상담’, ‘시간’, ‘어떻’, ‘되’로 나누고 여기에 품사 ‘의’, ‘와’, ‘은’, ‘게’, ‘나요’가 붙는 형태이다.

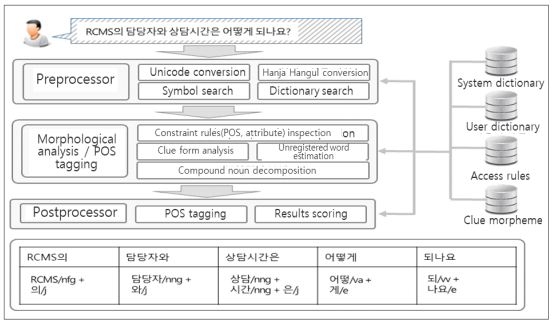

Fig. 3 Morpheme analysis in chatbot

정확한 질의어 분석을 위해서는 사용자의 오류도 정상의 질의로 받아들일 수 있어야 한다. 일상생활에서 빈번하게 발생하는 한영 변환 입력 및 오타를 형태소 분석 시 자동으로 고쳐서 분석하고, 사투리도 동의어 및 유의어 사전에 사전 등록하여 질의를 분석한다. 이런 절차를 통해 사용자의 질의에 대한 정확한 답변을 제공할 수 있다. 그림 4는 한영변환과 오타교정을 통해 정확한 답변을 제공한 예시이다. ‘회원가입’에 대해 질의하기 위해 사용자가 입력하였으나 자판이 영문으로 설정되어 ‘ghldnjsrkdlq’라고 등록하더라도 한영변환을 통해 ‘회원가입’으로 질의를 인식하고 이에 적절한 답변을 하는 것이다. 마찬가지로 ‘세세목’에 대한 사용자의 질의 문장에 ‘새새목’이라고 입력하여도 원래의 의미로 인식한다.

또한, 다양한 언어의 자연어 이해를 위해서 형태소 분석 단계에서 한자 및 영어, 숫자에 대한 인식을 하고난 후 정규화를 수행한다. 숫자는 시스템 개체를 통해 일반적인 숫자, 날짜 시간 등 다양한 형태의 숫자 타입을 인식하고, 다국어의 경우 다국어 API를 활용하여 한글로 번역하여 질의 의도를 분석한다.

사용자의 질의의도를 분석하기 위해서 Grammar(Rule) 기반 질의의도 이해 모델과 기계학습 기반 질의의도 이해 모델을 결합하여 이용한다. 사용자의 질의문을 NLU (Natural Language Understanding)기반으로 질의의도를 판별한다.

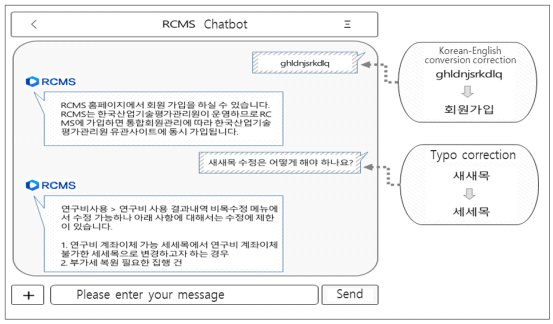

Fig. 4 Korean-english conversion and typo correction

Grammar 기반 질의의도 이해 모델은 정확도가 높지만 생산성이 떨어진다. 그래서 기계학습 기반 질의의도 이해 모델을 같이 사용하는데, 그림 5와 같이 등록되어 있는 질의의도의 질의문들과 유사도를 계산하여 사용자 질의의 질의의도를 결정하는 방식으로 개념적으로는 카테고리 자동분류와 동일하다. 원본 데이터와 학습 데이터를 분석한 후 사용자의 질의와 유사한 top-k의 질의의도의 질의문을 추출하여 사용자의 질의의도를 판별한다.

또 다른 방법은 기존의 FAQ와 1:1문의 게시판에 있는 다양한 질문과 답변을 지능형 검색엔진을 바탕으로 색인하고, 사용자가 질의하면 사용자의 질의에 가장 유사한 질문을 알고리즘에 의해 찾아서 매핑된 결과를 제공한다. 답변 결과에는 기계학습 기반 질의의도 이해 모델과 비교하여 유사율을 같이 보여주는데, 그림 6과 같이 매칭/미매칭 리스트도 같이 제공된다.

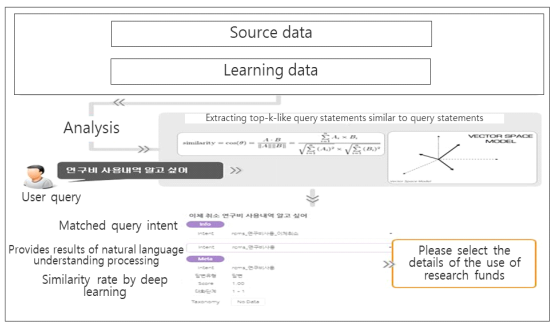

Fig. 5 Query inquiry understanding model based on machine learning

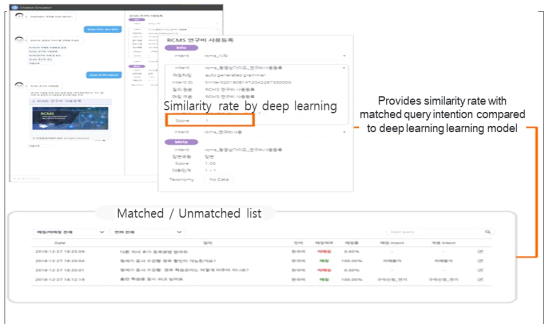

Fig. 6 Query inquiry understanding model based on searching previous Q&A

이렇게 자동 생성한 Grammar 기반의 질의의도 분석 방법과 기계학습 기반의 질의의도 분석 방법을 결합한 결과와 개별적인 색인·검색 기반 질의의도 분석 방법에 의한 결과를 비교하여 사용자의 질의의도를 효과적으로 파악한다.

3.2.3. 사용자 편의성 향상 기능

챗봇을 운영하기 위해서는 사용자의 질의의도를 분석하여 결과를 제공하기 위해 지속적인 질의의도, 대화흐름 등을 유지관리하고, 형태소 분석을 위한 다양한 용어사전을 업데이트해야 한다. 또한, 챗봇 자체가 사용자, 상담사 및 관리자 등에게 편의성을 제공하여야 한다. 질의-답변의 구성을 기본으로 하되 질의에 대한 답변이 여러 가지인 경우에는 사용자에게 구분자를 제공하여 선택하도록 한다. '연구비 사용등록'을 질의의도로 판별하면 챗봇은 질의자에게 증빙우선등록, 비세목우선등록, 참여연구원 직접지급등록, 현물사용등록 중에서 답변을 원하는 분류를 선택하도록 하고, 선택된 분류에 대한 답변을 제공한다.

RCMS 챗봇의 형태소 분석을 위해 KoBRICK라는 형태소분석기를 사용했는데, 이 분석기에는 관리기능을 통해 관리자가 직접 단어를 추가 및 삭제할 수 있다. 예를 들어 유사단어를 관리하는데, '사용등록'이라는 단어에 '집행등록'을 유사단어로 등록하면 질의문에 사용등록과 동일한 답변을 할 수 있다.

챗봇이 사용자의 질의내용을 이해하지 못하는 경우에는 사용자 불편을 해소하기 위해 상담사와 연결하도록 하여 사용자의 편의성을 갖췄다. 이 때, 상담사는 사용자-챗봇간 대화내용을 확인하여 사용자의 질의의도를 정확히 파악하여 답변할 수 있도록 하였다. 그림 7은 답변 실패 시 상담사 연결 업무절차도이다.

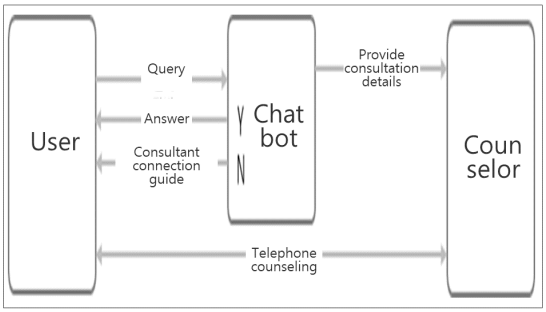

Fig. 7 Procedure for connecting a counselor when no-matching cases

RCMS Chatbot 시스템의 관리운영 편의성을 위해서 웹 기반의 관리기능을 구현하였다. HTML 5 기반 시각화 통계를 통해 챗봇의 운영현황 확인이 가능하고, 운영현황 통계분석정보를 엑셀형태로 다운받을 수도 있다. 운영현황은 지식현황, 질의응답 매칭현황, 사용 순위 현황의 3가지이며, 지식현황은 대화흐름 및 질의의도 등의 구축건수이다. 질의응답 매칭현황은 매칭율과 함께 질의 수 및 처리량도 함께 보여준다. 마지막으로 질의의도 및 질의어 사용 순위를 제공하는데, RCMS 사용자의 관심도를 확인할 수 있어 질의의도나 대화흐름의 세분화, 다양화 등을 판단하는 참고자료가 될 것이다.

3.2.4. RCMS 챗봇 시스템 적용

RCMS 시스템에 구현된 챗봇은 모든 사용자에게 서비스되는 것이 아니라 RCMS를 통해서 과제의 연구비를 사용하고 있는 연구자가 수행기관로그인을 통해서만 사용할 수 있다. 그림 8은 RCMS에서 제공하고 있는 챗봇에 실제 질의를 하고 답변을 제공받은 화면이다.

3.3. 고찰 및 향후연구

RCMS 챗봇은 시간의 구애 없이 일정 수준의 고객 문의에 대해 실시간 대응이 가능하다. 현재 RCMS 챗봇은 연구자용 서비스로 구현하였다. 이는 챗봇을 구성할 때 사용한 데이터가 연구자의 1:1 문의 및 답변 정보만을 이용했기 때문이다. 관리기관 담당자 등 모든 사용자의 상담 로봇이 되기 위해서는 Intent를 지속적으로 개발하여 유연한 대화흐름이 되도록 해야 하며, 신규 용어에 대한 사전 등록도 빠짐없는 수행이 필요하다. 챗봇의 답변 오류로 인해 발생 가능한 문제에 대비하기 위해 상담

내역에 대한 주기적 분석과 대화흐름의 개선 또한 필요하다.

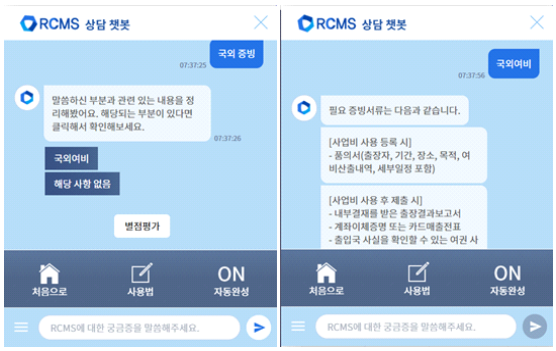

Fig. 8 Screenshot of RCMS chatbot

Ⅳ. 결 론

본 논문에서는 연구비관리시스템의 개선방안으로 인공지능 기술 기반의 챗봇 서비스를 제안하고 구현하였다. 본 논문에서 구현한 인공지능 기반 서비스는 24시간 365일 연구자의 상담에 대응할 수 있는 RCMS 챗봇 서비스이다.

서비스 구현을 위해서 인공지능 기반의 모형의 설계 절차에 따라 RCMS 챗봇은 상담내역 데이터를 통해 사용자의 질의의도를 분석하고 대화흐름 방식으로 답변을 할 수 있도록 질의의도를 구현하였다. 본 논문에서 구현한 인공지능의 모형 중 RCMS가 범부처 통합연구비관리시스템이 된 이후 콜센터로의 전화문의가 점점 증가하여 최근에는 매일 900여건에 달하고 있다. 본 논문에서 구현한 챗봇의 사용이 활성화되고, 사용자의 문의와 답변이 계속 누적된다면 더 많은 intent를 생성하고, 더 나아가 챗봇의 답변의 질적 수준이 성장할 것이다.

본 논문에서는 국가연구개발사업의 연구비 관리를 목적으로 운영하고 있는 연구비관리시스템(RCMS)의 사용자 편의성 향상 및 효율성 증진을 위해 인공지능 기반 챗봇 기술을 도입하였다. 이것은 국가연구개발사업 관리를 위한 정보시스템에 인공지능 기술을 적용하는 초기의 시도이지만 제안된 향후 보완해 나가면 더 많은 효율성이 있을 것으로 사료된다.

REFERENCES

[1] S. O. Yoon, “A Study on the Main Issues of Artificial Intelligence-based Public Services,” *Korea Public Management Review*, vol. 32, no. 2, pp. 83-104, Jun. 2018.

[2] K. R. Seo, D. S. Kang, J. K. Eo, S. J. Park, J. M. Kim, and H. J. Kim, “A Study on Development Model of Public Intelligent Virtual Assistant Service based on Artificial Intelligence,” *Proceedings of Symposium of the Korean Institute of Communications and Information Sciences*, pp. 890-891, Jun. 2017.

[3] J. h. Park, K. I. Yoon, and S. T. Min, “AI-based Chatbot System Technology Trend,” *Korea Information Processing Society Review*, vol. 26, no. 2, pp. 107-117, Jul. 2019.

[4] J. T. Kim, H. G. Lee, and H. S. Kim, “Effective Generative Chatbot Model Trainable with a Small Dialogue Corpus,” *Journal of the Korea Information Science Society*, vol. 46, no. 3, pp. 246-252, Mar. 2019.

[5] H. Y. Song and Y. S. Lee, “Informational Chatbot System about COVID-19 based on Natural Language Processing,” *Proceedings of the Korean Information and Communication Society Women's ICT Conference*, pp. 23-25, Aug. 2020.

[6] J. W. Lee, I. Y. Yeo, and H. K. Jung, “Document Analysis based Main Requisite Extraction System,” *Journal of the Korea Institute of Information and Communication Engineering*, vol. 23, no. 4, pp. 401-406, Apr. 2019.

[7] J. W. Lee, G. Wu, and H. K. Jung, “Deep learning Document Analysis System Based on Keyworkd Frequency and Section Centrality Analysis,” *Journal of Information and Communication Convergence Engineering*, vol. 19, no. 1, pp. 48-53, Mar. 2021.

[8] J. W. Lee, H. J. Yang, and J. G. Kim, “Developing Scenario for Implementation of Counseling Chatbot and Verifying Usefulness,” *Journal of the Korea Contents Association*, vol. 19, no. 4, pp. 12-29, Apr. 2019.

김용국(Yongkuk Kim)

2001년 홍익대학교 컴퓨터공학과(공학사)
2006년 홍익대학교 컴퓨터공학과(공학석사)
2021년 배재대학교 컴퓨터공학과(공학박사)
2000년~현재 한국산업기술평가관리원 재직
※관심분야 : 빅데이터, 인공지능 연구관리

김수진(Sujin Kim)

1994년 계명대학교 전자계산학과(공학사)
1996년 계명대학교 전자계산학과(공학석사)
2020년~현재 배재대학교 컴퓨터공학과(박사과정)
2021년~현재 베스핀글로벌 재직
※관심분야 : 정보보안, 딥러닝, 빅데이터

정회경(Hoekyung Jung)

1985년 광운대학교 컴퓨터공학과(공학사)
1987년 광운대학교 컴퓨터공학과(공학석사)
1993년 광운대학교 컴퓨터공학과(공학박사)
1994년~현재 배재대학교 컴퓨터공학과 교수
※관심분야 : Machine learning, Big data, Embedded system, U-Healthcare, IoT