Journal of Digital Convergence
Vol. 18. No. 11, pp. 215-226, 2020

ISSN 2713-6434 / eISSN 2713-6442
https://doi.org/10.14400/JDC.2020.18.11.215

사용자 친화적인 대화형 챗봇 구축을 위한 개발방법론에 관한 연구

현영근[1], 임정택[2], 한정현[1], 채우리[1], 이기현[1], 고진덕[1], 조영희[3], 이주연[4*]
[1]아주대학교 산업공학과 학생, [2]SK(주) C&C 연구원,
[3]아주대학교 산학협력단 연구원, [4]아주대학교 산업공학과 교수

A Study on the Development Methodology for User-Friendly Interactive Chatbot

Young Geun Hyun[1], Jung Teak Lim[2], Jeong Hyeon Han[1], Uri Chae[1], Gi-Hyun Lee[1], Jin Deuk Ko[1], Young Hee Cho[3], Joo Yeoun Lee[4*]
[1]Student, Division of Industrial Engineering, Ajou University,
[2]Researcher, SK Holdings Co., Ltd.
[3]Researcher, Ajou University Industry-Academic Cooperation Foundation
[4]Professor, Division of Industrial Engineering, Ajou University

요 약 챗봇이 비즈니스의 중요한 인터페이스 창구로 떠오르고 있다. 이러한 변화는 챗봇 관련 연구가 자연어처리(Natural Language Processing)기법에서 자연어이해(Natural Language Understanding) 그리고 자연어생성(Natural Language Generation)으로 지속적으로 발전했기 때문이다. 하지만, 챗봇을 개발하는 과정에서 도메인 지식을 이끌어내고, 사용자 친화적인 대화형 인터페이스로 개발하는 방법론적 연구는 미약한 것이 현실이다. 본 논문에서는 챗봇 개발의 프로세스적 기준을 제시하기 위해 이전 논문에서 제시한 방법론을 바탕으로 실제 프로젝트에 적용하며 개발방법론을 개선하였다. 결론적으로 가장 핵심적인 단계인 테스트 단계의 생산성을 33.3% 향상하였으며, 그 반복횟수도 37.5%로 단축하였다. 이러한 결과를 바탕으로 “3 Phase and 17 Tasks 개발방법론”을 제시하였으며, 이것은 챗봇 개발의 시행착오를 획기적으로 개선할 것으로 기대한다.

주제어 : 챗봇, 개발방법론, 챗봇 라이프사이클, 대화형 챗봇, 인공지능

Abstract Chatbot is emerging as an important interface window for business. This change is due to the continued development of chatbot-related research from NLP to NLU and NLG. However, the reality is that the methodological study of drawing domain knowledge and developing it into a user-friendly interactive interface is weak in the process of developing chatbot. In this paper, in order to present the process criteria of chatbot development, we applied it to the actual project based on the methodology presented in the previous paper and improved the development methodology. In conclusion, the productivity of the test phase, which is the most important step, was improved by 33.3%, and the number of iterations was reduced to 37.5%. Based on these results, the "3 Phase and 17 Tasks Development Methodology" was presented, which is expected to dramatically improve the trial and error of the chatbot development.

Key Words : Chatbot, Development Methodology, Chatbot Lifecycle, Interactive Chatbot, AI

*This paper was studied with the support of the 2019 Ajou University Academic Research and Development Fund (S-2019-G0001-00522)
*Corresponding Author : Joo-Yeoun Lee(jooyeoun325@ajou.ac.kr)
Received August 13, 2020 Revised September 8, 2020
Accepted November 20, 2020 Published November 28, 2020

1. 서론

최근 챗봇 기반 메신저 형태의 인터페이스가 주목받고 있다. 챗봇은 'Chat'과 'Robot'의 합성어로 텍스트를 활용한 대화형 인터페이스를 통해 사용자와 상호작용하는 소프트웨어 에이전트를 일컫는다[1~4]. 사용 환경에 따라 인터페이스 에이전트(Interface Agent), 대화형 에이전트(Conversational Agent), 가상 도우미(Virtual Assistant) 또는 가상 동반자(Virtual Companion)와 같이 다양한 용어들로 사용되어 왔으나, 최근에는 메신저 상에서 인공지능을 기반으로 대화를 나누는 시스템을 주로 챗봇이라 지칭한다[5].

최초의 챗봇은 1966년 미국 MIT 인공지능 연구소에서 개발한 'ELIZA'이였으며, 정해진 질문에 정형화된 대답만 제공하는 한계점이 있었다. 하지만 최근에는 머신러닝 및 딥러닝 등의 인공지능 기술을 접목하면서 사용자 질문에 대한 의도파악이 가능하게 되었고, 이에 따라 더 복잡한 일을 수행할 수 있게 되었다. 또한 모바일 메신저의 사용자가 급증하면서 챗봇 서비스가 누구에게나 익숙한 인터페이스 환경이 되었고, 이에 따라 기존 메신저 업체들은 챗봇 시장에서 확고한 입지를 다지기 위해 각자의 메신저 플랫폼 API를 제공함은 물론 인공지능 기술이 적용된 메신저 플랫폼을 개발 및 공급하고 있다[5~7].

2. 챗봇 개발유형 및 적용사례

2.1 챗봇 개발유형

챗봇 서비스의 발전과 동시에, 대화 알고리즘을 직접 설계할 필요 없이 누구나 편리하게 챗봇을 제작할 수 있도록 빌더의 보급이 늘어나고 있다. 대표적으로 IBM의 왓슨(Watson) 및 구글의 다이얼로그 플로우(DialogFlow)가 있으며, 국내 기업으로는 네이버 클라우드 플랫폼의 네이버 챗봇 및 카카오의 아이 오픈 빌더(Kakao I Open Builder) 등이 있다[5][8].

챗봇의 개발유형 측면에서 살펴본다면, 인공지능(AI) 기반의 챗봇과 룰(Rule) 기반의 챗봇으로 구분될 수 있다. 인공지능 기반의 챗봇은 주어진 데이터를 통해 학습을 한 후 사용자의 명령을 처리하는 방식이며, 따라서 정해진 범위 외 새로운 형태의 요청도 학습하여 처리할 수 있다는 장점이 있다. 하지만 학습 데이터가 부족하거나 적절하지 않은 경우 많은 오류가 발생하여 룰 기반 챗봇보다 더 낮은 결과를 얻을 수 있다는 단점이 있다. 인공지능 기반의 챗봇는 왓슨(Watson), 다이얼로그 플로우(DialogFlow)가 대표적인 예이다. 반면에 룰 기반의 챗봇은 미리 세워둔 대화 규칙에 따라 사용자의 요청을 처리한다. 따라서 새로운 형태의 요청은 처리할 수 없지만 오류가 발생할 위험이 적다는 것이 장점이 있으며, 대표적인 예로 ALICE와 ELIZA가 있다[9]. Fig. 1에서는 인공지능 기반의 챗봇과 룰 기반 챗봇의 차이점을 비교하였다.[10]

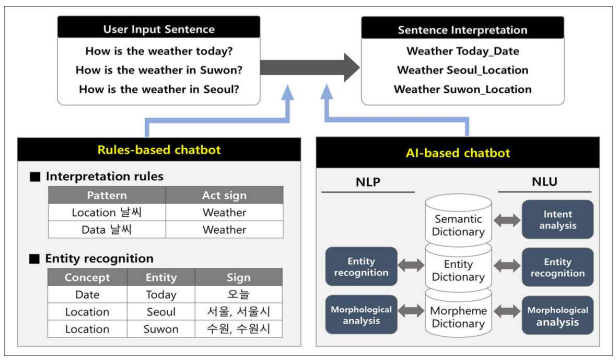


Fig. 1. Comparison Between Rule-based Chatbot and AI-based Chatbot

2.2 챗봇 도입사례

2.2.1 금융업 적용현황

챗봇이 가장 활발하게 도입하는 분야는 금융산업으로, 핀테크 챗봇과 함께 AI뱅킹 시대로 발전하고 있다[11]. 국내의 경우 은행, 증권사 및 보험사를 중심으로 고객 문의응대를 목적으로 한 챗봇이 주로 도입되고 있다. 예를 들어 대신증권의 경우 "벤자민"이라는 챗봇은 빅데이터를 활용해 생활정보(뉴스, 날씨) 뿐만 아니라 종목, 시황 및 펀드상품 추천 등을 제공하고 있다. 이 외에도 우리은행, 동부화재, 신한카드, KB 손해보험, NH농협 등이 상품안내, 고객 문의응대, 고객 맞춤형 카드추천 등의 서비스를 제공하고 있다. 해외의 경우 금융서비스 안내뿐만 아니라 금융비서 기능도 제공하고 있다. 예를 들어 Bank of America는 Erica라는 챗봇을 이용해 부채 상환관리 및 잔액확인 등의 서비스를 포함한 자산관리 기능을 제공하고 있으며, 페이스북 플랫폼을 활용한 Cleo는 은행계좌 및 신용카드 데이터를 학습하여 최적의 금융상품을 찾아주는 서비스를 제공 중에 있다[7].

2.2.2 유통업 적용현황

국내의 경우, 오픈마켓을 운영하는 대형 이커머스 업체 또는 대기업 중심으로 개인화된 상품을 추천하는 챗봇을 도입하고 있으나, 아직 비교적 단순한 특정 업무에서 개인 맞춤형 상품 추천기능을 제공하고 있다. 사례를 살펴보면 현대닷컴, 롯데닷컴, 11번가, 신세계백화점, CJ, 인터파크, 네이버 등에서 자체적으로 또는 카카오톡, 네이버 등의 플랫폼을 활용하여 상품검색, 주문, 배송 조회 등 업무를 제공하고 있고, 구매이력 및 개인 선호분석을 통해 상품추천을 하거나 선호 브랜드에 대한 쇼핑 정보를 제공하고 있다. 반면에, 해외의 경우는 보다 다양한 분야/기능들을 탑재한 챗봇이 운영되고 있으며, 매출을 향상시키기 위해 게임을 유도하거나 질문을 통해 사용자의 취향을 파악하는 기능을 탑재하기도 한다. 오픈마켓의 경우, AI 챗봇이 아닌 실제 판매자와 1:1 상담을 연결해주는 시스템을 운영하고 있으며, 중국 알리바바 타오바오의 아리왕왕(阿里旺旺), 인도의 플립카트(Flipkart)의 인앱 메시지 플랫폼 핑(Ping) 등이 대표적이다[7]

2.2.3 서비스업 적용현황

음식 배달서비스 업체들 또한 음식주문의 자동화를 위해 챗봇을 도입하고자 노력하고 있다. '딜리버리 히어로'는 해외에서 페이스북 메신저를 통해 '헝그리봇'이라는 챗봇을 시범운영 중에 있으며, '배달의 민족'은 네이버와 협력하여 인공지능을 적용할 예정이다. 방송, 택배 관련 분야와 여행, 숙박과 관련된 분야에서도 챗봇을 도입하여 운영하고 있다. 예를 들어 '여기어때'의 알프레도 챗봇은 테마별 숙소 추천, 사용자가 원하는 조건에 맞는 숙소 추천, 고객민원 응대 기능을 제공하고 있다. 또한 아시아나 항공 챗봇은 예약, 출국/도착 정보 확인, 수하물 정보, 탑승 절차, 마일리지, 인터넷 및 모바일 체크인 기능을 제공하고 있다. 이러한 챗봇들의 공통된 특징은, 다른 분야와 마찬가지로 대부분 고객문의 응대 및 상품 추천을 목적으로 한 챗봇이며, 난이도가 낮은 단순한 고객응대 서비스를 24시간 제공하기 위해 운영되고 있다. 해외 사례도 크게 다르지 않으며, 우버의 경우 기존 우버서비스 사용을 지원하는 목적으로, 대부분의 챗봇은 서비스 이용과 관련된 자주 묻는 질문에 대한 답변을 제공하는 기능을 주요 목적으로 하고 있다. SkyChat은 항공기, 탑승구, 좌석 정보, 화물 안내 등 항공기 이용 지원을 위한 챗봇 서비스를 제공하고 있고, Burner은 HomeAway, VRBO, Airbnb 등 객실 임대 서비스 자동응답시스템을 제공하고 있다[7].

2.2.4 헬스케어 적용현황

해외의 많은 시장조사기관이 챗봇이 도입되는 주요 분야로 헬스케어 분야를 꼽았지만, 국내의 의료분야 및 헬스케어 챗봇은 치료와 관련된 상용화된 서비스가 없는 상황이며, 단순 문의 중심으로 챗봇이 개발되고 있다. 이렇게 헬스케어 분야에 AI 도입이 난항을 겪는 주된 이유는 의료 및 제약 분야는 현재 원격의료가 법적으로 불가능하며, 원격의료에 대한 이해관계자들의 의견대립으로 인해 관련 법 제정이 어렵기 때문이다.

국내 챗봇 도입사례를 살펴보면, 경희의료원은 진료과 추천, 증상문의, 검사 안내, 치료 후 관리, 질환별 관리법 안내, 처방약 설명 등 병원 방문 전후 상담을 제공하고 있으며, 연세 의료원은 '웰니스 동반자 서비스'라는 명칭으로 의료 상담 서비스를 제공하고 있다. 해외 챗봇의 경우 크게 치료서비스 지원 및 건강관리를 지원하는 챗봇으로 분류할 수 있으며, 단순 문의응대뿐만 아니라 유료서비스로 구독료를 지불해야 이용 가능한 서비스도 있다. 건강관리 챗봇의 주요기능은 처방받은 약품에 대한 투약/복용 알람 그리고 증상에 따른 의료기관 정보의 제공이며, 치료서비스 챗봇의 주요기능은 심리치료 중심의 대화서비스를 제공하고 있다. 챗봇 제품을 구체적으로 살펴보면, Baidu라는 챗봇은 Melody를 통해 환자의 증상정보를 수집하고 의사에게 전달하며, 병원방문 예약서비스 및 건강정보를 제공한다. Florence라는 챗봇은 증상확인, 건강 관련 조언, 약 복용 알림, 의료기관 정보제공 등 개인 건강관리 서비스를 제공하고 있다. 이 외에도 X2AI, Woebot 등은 심리치료, 우울증, 기분 장애 관리 방법, 심리교육 및 인지치료 기술 등을 제공하고 있으며, Fitmeal는 식단 관리와 음식 보고서 등을 통한 다이어트 서비스를 제공하고 있다[7].

2.2.5 생활 유틸리티 적용현황

금융 서비스의 대체, 커머스의 상품 추천 그리고 헬스케어의 치료/건강 관리 등 각 산업별 비즈니스의 핵심적인 업무 외에도, 일상의 편의를 향상시키는 챗봇도 있으며 일기예보, 스케줄 관리, 지하철 도선도 등이 그 대표적인 사례이다. 예를 들어 젠틀파이의 날씨 챗봇인 '날보'는 페이스북 메신저 기반으로 간단한 날씨 정보를 제공하고 있다[12].

3. 관련 연구

챗봇과 관련한 연구의 주요방향성은 구문분석, 형태소 분석 및 추론기술 등의 자연어처리 (NLP: Natural Language Processing), 자연어 이해(NLU: Natural Language Understanding) 및 자연어 생성 (NLG: Natural Language Generation) 등 기술적 측면에만 집중되는 경향이 있으며, 사용자 경험적 측면을 고려한 챗봇 개발방법론에 대한 연구는 그 사례를 찾기 어려웠다.

하지만, 사용자 경험적 측면의 챗봇 개발방법론은 매우 중요하다. 그 이유는 챗봇 대화에서 정확도를 향상시키는데 개발방법론이 핵심적인 역할을 할 수 있으며, 나아가 챗봇을 통해 구현하고자 하는 비즈니스의 목표를 달성하는데 기여할 수 있기 때문이다. 그 단적인 예로, Microsoft 사의 'Tay' 라는 챗봇은 대화모델 설계에 집중하지 않아 사용자에게 성 차별, 인종 차별적인 응답을 내놓는 등의 문제로 하루 만에 서비스가 중단되는 상황도 발생하였다[11].

이것이 의미하는 것은, 기술적 연구에 힘입어 챗봇이 실생활에서 사용될 수 있는 수준으로 향상되긴 했지만, 아직은 인간처럼 가치를 판단하거나 유창한 대화를 이어갈 수 있는 수준에는 이르지 못한 것이 현실이고, 그럼에 따라 부정적 사용자 경험으로 이어질 수 있는 챗봇의 현실적 문제점을 보완할 '방법'이 필요하다는 것을 나타내는 것이다[2].

3.1 챗봇 개발방법론 사례

챗봇 개발방법론과 관련하여, 처음으로 그 기준을 제시한 것은 Chatbots Magazine(2018)에서 공개한 'Best Practice of Chatbot Development Methodology'이며 Fig 2.와 같다.

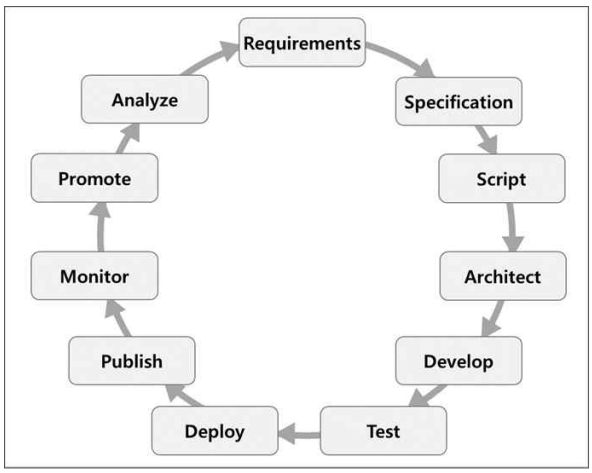


Fig. 2. Best Practice of Chatbot Development Methodology In Chatbots Magazine

챗봇 개발 수명주기에는 요구사항 수집 및 설계와 같은 단계가 포함되어 있으며, 전통적 소프트웨어 프로그램 개발과 유사하다. 하지만, 챗봇 개발에는 차별화된 몇 가지 고유 단계가 있으며, 그 이유는 챗봇이 기본적으로 대화식 인터페이스이기 때문이다[2]. 또한 챗봇은 사용자가 활용할 수 있는 수많은 메시징 응용 프로그램이 있기 때문에 확장된 테스트가 필요하며, 또한 챗봇을 배포(Deploy)한 후 대화 로그를 지속적으로 모니터링하여 챗봇의 유용성과 대화흐름을 개선할 필요가 있다[13].

Chatbots Magazine에서 제시하는 개발방법론의 각 단계별 활동은 Table 1.과 같다.

Table 1. Activity at Each Stage

Requirements	•Gathering market requirements •Definition about target customer, the goal of chatbot and benefits
Specification	•Develop a product spec for the bot •Identifying the features and functionality
Script	•This step is unique process. •Building conversational scripts that represent user interactions. •Scripting for user desired task.
Architect	•Creating the engineering design including both the front-end and back-end components.
Develop	•Making the conversational interface, •Developers should iterate much more coding and testing than traditional software development.
Test	•The code must be tested not just in the emulator, but also in the actual messaging platform
Deploy	•Deploying to a hosted environment. •The hosted environment must be stable and needs its own monitoring and DevOps support.
Publish	•Submitting to the various app stores for approval.
Monitor	•Indicating that systems are well, but the bot may still be unresponsive to certain user conversations.
Promote	•Introducing the distributed chatbot to new users.
Analyze	•Analyzing conversation logs and usage metrics.

3.2 기존 챗봇 개발방법론 사례의 개선

Gillian Cameron et al.(Lessons from Knowledge Engineering Methodologies for Chatbot Design and Development, 2018)은 Chatbots Magazine의 개발방법론이 일상생활에서의 다양한 상황과 업무적 특성을 적용하는데 한계가 있음을 파악하였고, 이를 개선하기 위해 특정한 도메인의 지식을 반영하는 '지식수집(Knowledge Gathering) 단계'와 '사용성 향상을 위한 유용성 테스트(Usability Testing) 단계'를 추가한 개선된 개발방법론을 제시하였으며, Fig 3.과 같다.

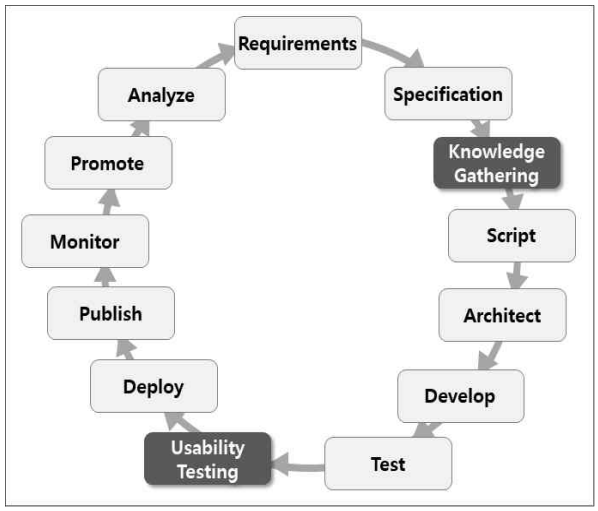


Fig. 3. Proposed Chatbot Development Life-Cycle

'지식수집(Knowledge Gathering) 단계'에서의 핵심은, 사용자와의 효과적인 대화식 인터페이스를 구성하기 위해 도메인(비즈니스) 지식이 얼마나 대화모델에 반영되도록 구성하는가가 중요하며, 보다 유용하고 강력한 챗봇을 구축하기 위해 도메인 전문가의 지식과 각 지식 간의 관계표현이 대화모델에 포함될 수 있도록 챗봇을 구축해야 한다는 개념이다. 따라서, 챗봇 대화모델의 지식기반 스키마를 작성하기 위해, 추가된 지식수집(Knowledge Gathering) 단계에서는 다양한 유형의 지식을 검색하기 위한 전문가와의 인터뷰를 반복적으로 수행하여야 하며, 인터뷰 이후에는 도메인 전문가로부터 검증(지식 검색 및 구조화)받기 위한 프로세스 또한 진행을 권장하고 있다[1].

'유용성 테스트(Usability Testing) 단계'에서의 핵심은, 지식 기반의 챗봇 스크립트를 생성한 후 도메인 전문가에게 다시 전달하여 스크립트가 지정된 목표를 달성할 수 있다고 생각될 때까지 유효성 검사를 수행한다는 개념이다. 즉, 사용자의 질의에 대해 챗봇이 이해하지 못했다면, 챗봇이 할 수 있는 것에 유도할 수 있도록 대화모델을 가져가야 하며, 또한 적절한 유머와 감정을 사용하는 스크립트를 적용하여 사용자의 경험을 향상시킬 수 있도록 전문가의 지속적 테스트를 통해 대화 흐름을 개선하는 것을 목표로 한다[1].

결론적으로, 첫 번째 챗봇 개발방법론 사례(Chatbots Magazine)에서 개선된 Gillian Cameron et al.의 개발방법론은, 새롭게 추가된 '지식수집(Knowledge Gathering) 단계'와 '유용성 테스트(Usability Testing) 단계'를 기존 Task와 상호 연결을 통해, 대화모델의 스크립트가 지식기반으로 작성될 수 있도록 개발 프로세스를 유도하는 것이 핵심이라고 할 수 있다.

4. 연구방법론

본 연구는 IBM Waston Assistant를 활용하여 약 15개월 동안 H사, S사 그리고 M사를 대상으로 한 실제 프로젝트에 적용했던 사례를 기반으로 진행되었다. 앞서 살펴본 바와 같이, 챗봇 개발방법론에 대한 기준 및 가이드가 없었던 관계로, Gillian Cameron et al.에 의해 제시된 개발방법론을 기준으로 프로젝트를 진행하였으며, 이를 토대로 첫 번째 프로젝트와 세 번째 프로젝트의 사업수행 기간 및 대화모델의 정확도를 비교 분석하여 효율적인 개발방법론에 대해 연구하였다.

먼저, 챗봇 프로젝트 수행경험의 부족과 고객의 챗봇에 대한 요건 불명확으로, 리스크 관리(Risk Hedge)를 위해 기존 Waterfall 방식의 절차적인 개발방법론 보다는 각 단계를 병렬적으로 진행하여 보다 민첩하게 대응할 수 있는 애자일(Agile) 방법론을 기준으로 프로젝트를 수행하였다.

애자일 방법론은 진행되는 프로젝트의 특성과 구성원의 참여도 등 실제 프로젝트 여건에 따라 그 유형을 선택할 수 있으며, Scrum, XP (eXtreme Programming), DSDM (Dynamic System Development), FDD (Feature Driven Development), ASD(Adaptive Software Development), Lean Software Development, Crystal Clear 등이 그 예이다[14]. 본 연구에서는 일반적으로 사용되는 방법론인 Scrum을 적용하였다.

Gillian Cameron et al.의 챗봇 개발방법론을 기반으로 Scrum 방법론을 적용한 구체적 방법은 Fig. 4와 같다. 먼저, Scrum의 첫 번째 단계인 Product Backlog는 Gillian Cameron et al.의 챗봇 개발방법론 중 Requirement단계, Specification단계 및 Knowledge Gathering단계에서 도출된 요구사항을 반영하였으며, Script단계부터 Test단계까지의 총4개 단계는 Scrum의 Sprint Backlog, Sprint, Daily Scrum Meeting 그리고 Shippable Product를 반복하여 적용하였다. 다만, Sprint는 각 Intent별 또는 업무단위별로 설정하여 Daily Scrum Meeting을 진행하였으며, 각 Sprint 기간은 프로젝트 일정을 고려하여 최대 2주, 반복개발 주기(Iteration)는 최대 5일을 넘지 않았다. 이것은 반복을 통해 최대한 고객요건을 맞추기 위함이다.

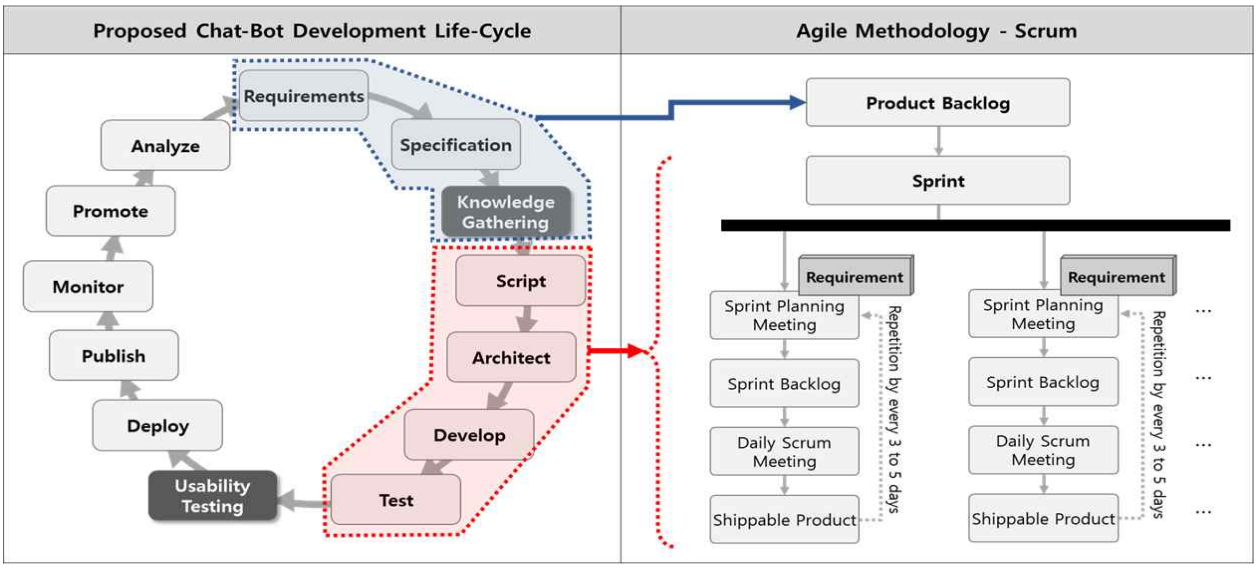


Fig. 4. Application of Agile Methodology in the Proposed 챗봇 Development Lifecycle

5. 개선결과

Gillian Cameron et al.,의 논문에서 제시한 챗봇 개발방법론과 비교하여 본 논문에서 제시하는 챗봇 개발방법론의 특징은, 일반적인 답변을 하는 챗봇 시스템의 관점을 넘어, UX(사용자 경험, User Experience)를 반영한 정보획득 과정의 구조화 관점으로 발전했다는 것이다. 그러한 의미에서 기존 개발방법론은 구현에 초점을 둔 방법론인 반면, 본 논문에서 제시한 방법론은 실제 활용 관점에서 사용자와 챗봇 간 상호작용에 중점을 두었다고 볼 수 있으며, 'Q&A 세트 확보', '스트럭처 정의', '대화모델 설계' 및 'C&T 테스트 단계'에서 해당 활동을 구체화하였다. 이러한 주장의 근거는 대화모델의 정확도를 확보하는 기간, 즉 고객이 챗봇에 대해 일정수준의 만족도를 확보하기 위한 개발기간을 분석함으로써 확인이 가능하다. Table 2.는 Gillian Cameron et al.,의 개발방법론을 적용했던 첫 번째 프로젝트와, 본 논문에서 제시한 개발방법론으로 보완된 이후의 수행 기간 및 대화모델의 정확도를 비교한 것이다. 이전 개발방법론에서는 전체 13주, 테스트 단계는 6주가 소요되었으나, 본 논문에서 제시한 개선된 개발방법론에서는 전체 11주로 약 2주 단축하였고, 특히 테스트 단계는 약 4주로 33.3% 단축하였다. 이것은 재학습 횟수가 기존에는 평균 8회였으나 5회로 단축(100개 질문 기준)하면서 기간이 단축된 것으로 확인하였다. 정확도가 기존 개발방법론과 차이가 크지 않은 것은 프로젝트 시작 시점에 고객과 대화모델 정확도를 80%로 협의하였기 때문이며, 재학습 횟수가 5회에 해당 정확도를 획득하였다는 것을 의미한다.

개선된 개발방법론의 각 단계별 상세활동은 하기와 같으며, 기존 개발방법론 상 반복개발 주기(Iteration)의 주요 대상이였던 'Script 단계'부터 'Test 단계'까지의 기존 4개 단계를 8개 단계의 세분화된 프로세스로 개선하였으며 해당 내용은 Fig.5와 같다.

5.1 페르소나 설정단계

챗봇이 사용자에게 어떠한 모습으로 보일 것인지를 설정하는 단계를 의미한다. 일반적으로 기업은 자신들이 원하는 기업의 이미지를 만들기 위해 많은 예산을 투입하고 있으며, 페르소나는 이러한 기업의 이미지를 챗봇에 적용하는 것을 의미한다. 페르소나를 챗봇에 설정하는 방법은, 먼저 나이, 성별, 성격 등 가상의 인물을 설정하고, 이 가상의 인물이 자주 사용하는 단어, 말투, 끝 맺음말 등을 정의한다. 이러한 것들은 사용자의 질문에 대한 답변에 가장 중점적으로 쓰이며 환영인사, 끝 맺음말, 그리고 챗봇의 이모티콘 설정 등에 사용되게 된다. 페르소나 세팅에서는 해당 업무를 처리하는 고객의 KPI를 활용하는 것도 의미가 있으며, 이것은 KPI가 해당 조직에서 제공하고자 하는 서비스의 핵심내용이기 때문이다. 또한 해당 업무를 직접 지켜보고 관찰하여 어떠한 말투를 쓰는지, 어떠한 단어를 쓰는지 그리고 그 속도는 어떠한지, 업무 숙련도는 어떠한지 등에 대한 체크리스트를 만들어 추후 답변을 만들거나 대화모델을 구성할 때 활용하면 유용하다.

Table 2. The Result of Methodology Improvement

Division		Proposed Chatbot Methodology *	Improved Chatbot Methodology **	Improvement Rate
Development period				
	Total Phase Period	13 Weeks	11 Weeks	15.4% ↑
	Test Phase Period	6 Weeks	(C&T Phase) 4 Weeks	33.3% ↑
Number of Relearning at C&T Phase		8 Times (Per 100 Questions)	5 Times (Per 100 Questions)	37.5% ↑
Conversation Accuracy Rate		79.7%	81.5%	2.2% ↑

* Fig. 4. Gillian Cameron et al, / ** Fig. 7. This Paper

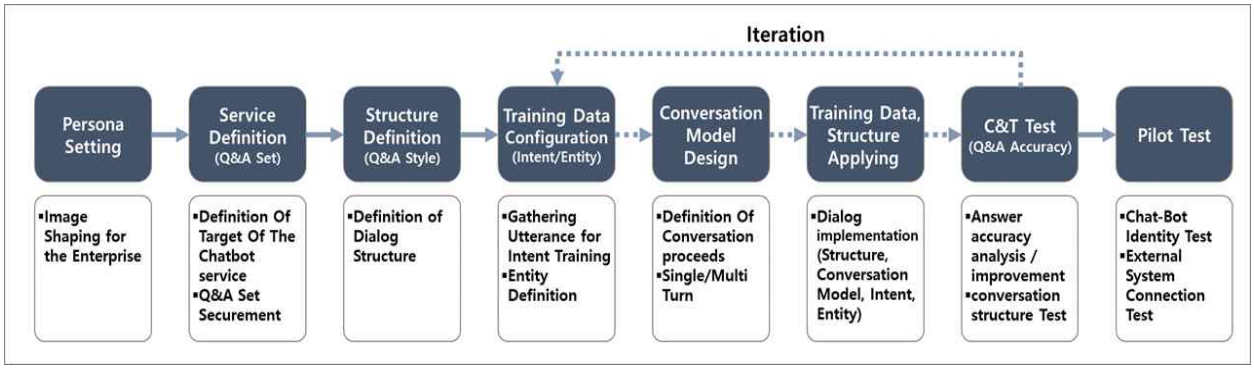


Fig. 5. Tasks Improvement Through Research

이와 관련된 연구에서는, 페르소나의 프로필 이미지와 대화 유형 간 적합성이 높을수록 서비스 신뢰성, 지속적 이용의도 그리고 서비스에 대한 긍정적 이미지가 향상되는 것으로 나타나, 변인 간 강한 상관관계가 있음을 확인되었다[15].

5.2 서비스 정의 및 Q&A 세트 확보

두 번째는 '서비스 정의' 및 'Q&A 세트 확보' 단계이다. 필수적인 과정이지만 프로젝트를 수행하는 과정에서 소홀하기 쉬워 별도의 단계로 구성하였다. 챗봇 서비스는 정의하고 학습하지 않은 영역에 대해서는 답변할 수 없으므로, '서비스 정의' 단계에서는 어느 업무영역 또는 어떤 지식영역에 대해 서비스를 제공할 것인가를 정의를 하고, knowledge Base라고 하는 지식 영역을 확정하는 과정을 의미한다. 'Q&A 세트 확보' 단계에서는 챗봇을 통해 답변하고자 하는 지식이 이미 Q&A로 정의되어 있다면(예를 들어, 기존 운영중인 콜센터의 매뉴얼 등), 해당 세트를 챗봇에 적용하기 위한 정리과정이 필요하다. 주의해야 할 사항은, 이 과정에서 업무시스템과 연결되어 처리되어야 하는 대화 세트도 정의해두는 것이 다음 단계에서의 업무처리에 유용한다.

가장 중요한 것은, 서비스 하고자 하는 챗봇이 사용자에게 어떤 역할을 할 수 있는 챗봇으로 인식시켜 활용을 극대화 하냐는 것이므로, 가장 많은 문의가 들어오는 주제가 챗봇의 영역이 될 수 있고, 또한 답변하기 쉬운 영역이 챗봇의 영역이 될 수 있다. 챗봇 기술은 Classification, 즉 분류를 핵심으로 한 서비스 형태이기 때문에, 챗봇의 성능을 위해서라도 챗봇의 영역을 특정 범위로 구분하고 정의하는 것이 챗봇의 성능과 구현 관점에서 매우 중요하다.

5.3 스트럭쳐 정의 단계

스트럭처란 챗봇에서 사용자가 대화를 시작하고 이어나가는 방식을 의미하며, 질문을 받는 구조와 답변을 하는 구조 모두에 대해 한 번의 문장으로 정보를 제공할 것인지, 아니면 여러 번의 문장으로 정보를 취합하여 제공할 것인지를 정의하는 단계이다. 예를 들어 처음 대화를 시작할 때 사용자의 문장 입력을 바로 받을 것인지, 아니면 많이 물어보는 질문을 먼저 제시해 주거나 추천해줄 것인지, 또는 버튼이나 리스트를 이용하여 대화의 종류를 1차 분류한 후 그 다음 사용자의 문장을 입력 받을 것인지를 결정하는 것으로, 서비스의 특성을 고려하여 챗봇의 전체 스트럭처를 구성한다. 이렇게 스트럭처가 구성되면 질문유형 별로 사용자에게 어떻게 입력을 받을지 그리고

어떤 답변을 할지를 확정한다. 어떤 경우에는 정확히 '예/아니오'라고 입력이 필요할 수도 있고, 선택해야 하는 것을 한정하거나 현재 할 수 있는 서비스나 답변이 한정되어 있을 수도 있다. 또한 입력을 문자 형태 외에 이미지나 동영상 그리고 URL Link로 추가 정보를 안내할 수 있다. 이처럼 챗봇 사용자가 원하는 정보를 잘 전달하기 위해, 문자로 입력하는 형태뿐만 아니라 다양하게 구성함으로써 챗봇 서비스의 활용성을 극대화하는 것을 목적으로 한다.

스트럭처는 챗봇의 성능과 매우 관련이 높고 사용자의 활용 만족도에서도 가장 큰 영향을 미침으로, 스트럭처를 정의하는 단계에서는 실제 과정에서 하나하나의 질문과 대답에 대해서 고객과의 워크샵을 통해 각각의 스트럭처를 정의하는 것이 매우 중요하다.

5.4 학습데이터 구성 단계

앞 단계에서 어떤 챗봇을 만들 것인지, 질문과 대답을 어떤 방식으로 처리할지 정의했다면, 현재 단계에서는 챗봇에게 해당 내용을 어떻게 학습시킬 것인가에 대해 정의하는 단계이다. 여기서 중요한 개념이 Intent와 Entity로, Intent는 인공지능 기반으로 사용자의 발화 의도를 파악하는 것을, Entity는 룰(Rule) 기반으로 입력된 문장 내 정보를 파악하는 것을 말한다. Intent는 일반적으로 지도학습(Supervised Learning) 기반임으로, Intent의 유형은 머신러닝(Machine Learning)의 목표함수(정답)가, 예시 문장은 학습 데이터가 된다.

예시문장은 솔루션마다 다소 상이하나 일반적으로 Intent별 약 30개의 예시 문장을 권장한다. Intent에서 중요한 것은, 10개의 Intent 예시 문장을 만든다면 Intent별 한 명이 각각의 Intent 예시문장을 30개씩 만들기 보다는, 10명이 10개 Intent에 예시문장을 각 3개씩 만드는 것을 권장한다. 왜냐하면 사람마다 같은 질문을 해도 물어보는 방식과 사용하는 단어 그리고 문장의 길이가 다르기 때문에, 이러한 다양한 사용자 발화 형태를 확보하기 위함이며, 이를 통해 어느 한쪽에 치우치지 않고 인식률을 향상시킬 수 있기 때문이다. 두 번째로 Entity는 입력된 문장 내 정보를 파악하는 것으로, 필요한 정보가 어떤 것인지 정의해야 한다. 상품을 주문하는 챗봇 서비스라면 상품명과 수량을, 공연을 예약하는 서비스라면 공연이름, 공연시간 그리고 티켓 수량을 정확하게 획득하는 것이다. Entity에서는 동의어와 특정 패턴, 그리고 이미 등록되어 있는 시스템 Entity를 활용할 수도 있다. 예를 들어, '세탁물'이라는 단어를 어떤 사용자는 '빨래'라고 사용할 수 있고, 또 어떤 사용자는 '빨랫감'이라고 사용할 수 있는 것이므로, 이러한 단어들을 동의어로 등록하고 같은 의미로 정보를 파악할 수 있도록 관리를 해야 한다. 또한 휴대폰 번호 정보를 파악하고자 한다면, 010으로 시작하는 세 자리, 네 자리 그리고 네 자리 숫자의 패턴은 휴대폰 번호로 인식할 수 있도록 특정 패턴 정보를 정의하고 등록 관리하여야 한다. 마지막으로 시스템 Entity는 시간, 주소, 날짜, 환율 등이 있으며 이러한 Entity를 등록하여 활용한다면 보다 효율적으로 정보를 관리할 수 있다. 이러한 Intent 및 Entity는 다음 단계인 대화 모델 설계와 테스트 과정에서 등록, 수정 및 삭제과정을 거치게 된다.

5.5 대화모델 설계 단계

대화 모델의 설계라는 것은, 질문이나 어떠한 업무의 특성을 반영해서 대화 모델을 Single Turn으로 설계할지, Multi Turn으로 설계할지를 결정하는 단계를 말한다. 간단한 예로, "공연시간은 어떻게 되나요?" 라는 질문에 "평일은 오후 7시 그리고 10시, 주말은 오전 11시, 오후 3시, 5시, 7시입니다" 라고 답변할 수도 있지만, 같은 질문에 "평일과 주말 중 언제가 편하신가요?" 라는 질문을 두어 구분한 후 답변할 수 있을 것이다. 즉, 최종 답을 얻기까지 몇 번의 트랜잭션을 일으키고 처리할지에 대해 확정하는 단계이다.

5.6 학습데이터 스트럭처 적용 단계

앞 단계에서 정의한 사항들을 바탕으로 챗봇의 대화모델를 구현하는 것으로, 슬롯(Slot) 또는 디그레션(Digression) 기능을 설정하고, 사용자의 의도를 파악하지 못했을 경우 예외처리를 어떻게 할지 결정한다.

일반적으로 챗봇 솔루션은 효과적으로 대화 모델을 개발하고 파악할 수 있도록 GUI 기반의 개발도구를 제공하고 있다. 스트럭처 정의단계에서 수립한 대화구조를 기준으로 전체 노드를 구성한 후, 각 개별 노드를 설계하며 Intent 및 Entity를 해당 노드에 설정하고 페르소나에 맞추어서 적용한 답변을 개발도구(Tool)에 등록한다. 해당 과정은 챗봇 엔진에 따라 상이하겠지만 적용 데이터의 양이 많고 복잡한 경우에는 개발도구의 UI상에서 적용하기 보다는 API방식을 활용하여 정의된 데이터를 한 번에 적용시키는 것이 일반적이다. 실제로 H사, S사 그리고 M사의 경우도 동일한 방식으로 수행하였다.

5.7 구성 및 학습결과 테스트 단계

구성 및 학습결과 테스트 단계는 줄여서 C&T (Configuration and Training) 테스트 단계라고 하며, 현재까지 설계하고 학습된 결과를 테스트하고 대화모델의 정확도 확인한 후, 반복적으로 테스트 및 대화모델 개선 작업을 통해 정확도를 향상시키는 단계를 의미한다.

테스트를 위한 데이터는 일반적으로 학습 데이터를 구성하는 단계에서 학습 데이터의 1.5배를 수집하고, 그 중 학습 데이터로 쓰지 않은 데이터를 테스트 데이터로 활용해야 한다. 프로젝트에서는 해당 프로젝트에 투입되지 않았던 제 3자에게 테스트 문장을 요청하여 테스트에 활용하는 것이 유용하며, 이러한 문장을 바탕으로 크게 두 가지의 테스트를 수행한다. 첫 번째는 사용자의 발화 의도를 얼마나 잘 인식하는지를 분석하는 것으로, 원하는 Intent로 분류가 되는지, 정확한 Entity가 분석되는지를 확인한다. 이러한 분석은 신뢰점수(Confidence Score)를 활용하며, 얼마만큼의 신뢰도를 가지고 있는지를 나타내는 정량적인 수치, 정확도를 의미한다. 두 번째는 챗봇의 구조에 대한 테스트로, 사용자가 챗봇과 대화를 이어가던 중 대화 끊김, 대화 오류 또는 대화의 무한루프 등을 테스트한다.

이 테스트 단계에서의 가장 핵심적인 목표는, 원하는 성능(대화의 정확도)이 도출될 때까지 Intent 및 Entity의 학습 데이터 구성을 반복적으로 수행하는 것이다. C&T 테스트 단계에서 주의할 사항은 두 가지가 있으며, 첫 번째는 대화의 목표 정확도를 고객과 사전에 확정한 후 테스트를 진행해야 하며, 그렇지 않으면 대화모델의 정확도 향상을 위해 의미없는 무한 반복적인 업무를 할 가능성이 있기 때문이다. 참고로 본 연구의 사례로 진행했던 프로젝트에서는 대화모델의 정확도를 80%로 협의한 후 진행하였다. 두 번째는 챗봇을 사용해보지 않았던 사용자 그리고 성별, 나이, 업무영역 등 다양한 사용자가 테스트에 참여해야 정확도를 향상시키는데 도움이 된다.

5.8 통합 테스트 단계

본 단계에서는 C&T 테스트 범위를 넘어 업무시스템 연계 테스트뿐만 아니라 페르소나, 아이콘 등과 같은 챗봇 아이덴티티의 내적 요소도 함께 고려하며 테스트를 해야 한다. 또한 테스트의 최종 단계로서 다양한 테스터가 수행해야 함은 물론이고, 활용하는 형태가 앱인지 웹인지 그리고 어떤 디바이스를 중점적으로 생각하는지 등을 고려하는 것이 필요하다.

‘4.1 페르소나 설정단계’ 단계부터 ‘4.8 통합 테스트 단계’ 까지의 요약정리는 Fig. 5와 같다.

6. 결론

본 연구는 기존 논문에서 제시한 챗봇 개발방법론을 기준으로, 실제 프로젝트에 애자일 방법론을 활용하여 반영함으로써 이론적 근거와 실제 사용자의 경험을 바탕으로 프로젝트의 프로세스적 기준을 연구하였다.

연구 결과, Gillian Cameron et al.이 제시한 개선된 개발방법론 13개 Task에 대해, Fig 6.과 같이 업무성격

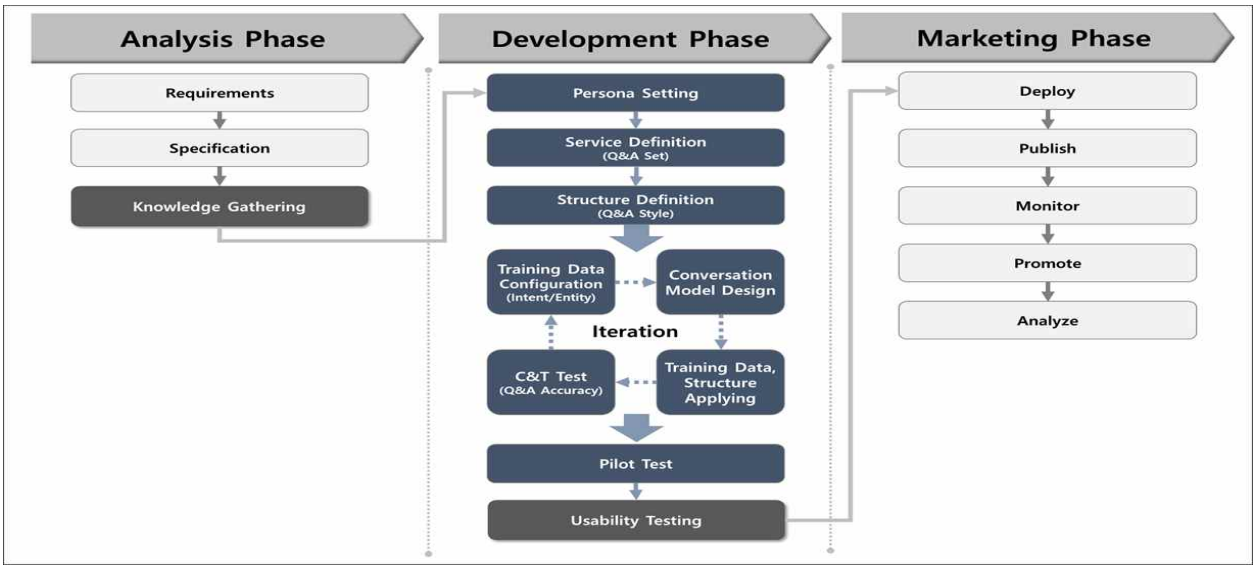


Fig. 6. Improved Chatbot Lifecycle

을 고려하여 3 Phase로 구분하였으며 전체 17개 Task로 세분화하여 업무의 기준을 보다 명확히 하였다.

첫 번째 Phase에서는, 챗봇을 통해 어떤 것을 개발할 것인지, 구체적 요구사항은 무엇인지 그리고 도메인 지식을 바탕으로 전문가 인터뷰는 어떻게 진행할 것인지 등을 분석하고 진행하는 단계로, 기존의 Task를 그대로 따랐다. 두 번째 Phase는 실제 대화형 인터페이스를 개발하고 테스트하는 단계로, 기존 4개 Task(Script – Architect– Develop Test)를 8개 Task(Personal Setting – Service Definition – Structure Definition –Training Data Configuration –Conversation Model Design – Training Data Structure Applying –C&T Test – Pilot Test – Usability Test)로 세분화하여 개선하였다. 이 Phase의 핵심은 챗봇 서비스가 고객사의 기업 이미지에 부합되도록 개발되는지 그리고 챗봇을 통해 개선하고자 하는 현재의 비즈니스 문제에 대해 원하는 결과를 얻을 수 있는지 등을 확인하며 개발하는 것이다. 마지막 세 번째 Phase는 개발이 완료되고 서비스를 시작한 후 모니터링 하는 단계로, 기존 Task를 그대로 준수하였다.

이러한 개선된 개발방법론은 챗봇을 개발하고 운영하는데 있어서, 특정 도메인의 성격과 비즈니스적 특 징을 반영하는 과정에서 고객의 요구사항에 대해 보다 사용자 친화적인 챗봇을 개발하는데 시행착오를 획기적으로 줄일 수 있을 것으로 기대된다.

REFERENCES

[1] Gillian Cameron et al. (2018), *Back to the Future Lessons from Knowledge Engineering Methodologies for chatbot Design and Development,*
DOI : 10.14236/ewic/HCI2018.153

[2] H. J. Baek, S. Y. Kim & S. W. Lee. (2019). *Effects of Interactivity and Usage Mode on User Experience in chatbot Interface.* Journal of the HCI Society of Korea 14(1), 35-43
DOI : 10.17210/jhsk.2019.02.14.1.35

[3] Yan, M., Castro, P., Cheng, P. & Ishakian, V. (2016). *Building a chatbot with Serverless Computing.* In Proceedings of the 1st International Workshop on Mashups of Things and APIs. Trento, Italy. p. 5.

[4] S. G. Kim & J. Y. Yun. (2020). *Multidisciplinary User Experience Research on Task-oriented chatbot Images.* The Korean Society of Science & Art 38(2), 33-43
DOI: 10.17548/ksaf.2020.03.30.33

[5] S. M. Kim. (2019). Comparative Study on the Usability of the Google and Kakao *chatbot* Builders. Master's Dissertation, Konkuk University, Seoul

[6] H. S. Lim. (2018.10.01.) *The app is gone and the AI era has come.* Yonhapnews.
http://www.yonhapnews.co.kr/bulletin/2016/03/31/0200000000AKR20160331005351091.HTML

[7] S. K. Kim, M. C. Shin & J. Y. Kang, (2018), *Introduction to chatbot technology and case analysis,* Korea Institute Of Communication Sciences, 35(2), 21-28.

[8] H. J. Kang & S. I. Kim. (2017). *Evaluation on the Usability of chatbot Intelligent Messenger Mobile Services -Focusing on Google(Allo) and Facebook(M messenger).* Journal of the Korea Convergence Society. 8(9), 271-276.
DOI : 10.15207/JKCS.2017.8.9.271

[9] T. S. Hur, G. Y. Mok, J. H. Kim, S. H. Baek and J. H. Lee (2020). *University Bulletin Chat-bot System based on Natural Language Processing.* The Korean Society of Computer Information Conference 28(1), 2020.1, 169-170.

[10] M. S. Kim, J. K. Cho. et al. (2018). *Design and Implementation of a chatbot System for Smart Sores.* Proceedings of Symposium of the Korean Institute of communications and Information Sciences. 1569-1570

[11] Hunt, E. (Oct. 1. 2018). *Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter.*
https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-*chatbot*-gets-a-crash-course-in-racism-from-twitter.

[12] J. N. Park. (2017). *The role of the chatbot is not 'talker' but 'effectiveness'. Data quality over algorithms.* Dong-A Business Review(DBR).
http://dbr.dong.com

[13] Beerud Sheth. (Jan. 24. 2018). *Chatbots Magazine. The Bot Lifecycle. What to know before you make your chatbot.* https://chatbotsmagazine.com

[14] H. M. Kim, (2013) *A Case Study of Android Applications Development with Agile Methodology.* Doctoral Dissertation, Soongsil University, Seoul

[15] S. K. Lee & J. Y. Yun. (2019). *A Convergence Study on chatbot Persona and User Experience of Financial Service - Focused on Loan Service -.* The Korean Society of Science & Art 37(4), 257-267
DOI : 10.17548/ksaf.2019.09.30.257

현 영 근(Hyun, Young Geun) [정회원]

· 2017년 8월 ~ 현재 : 아주대학교 공과대학 산업공학과 석박사통합과정
· 2018년 1월 ~ 현재 : SK(주) C&C DT 전략 Marketing 그룹
· 2004년 12월 ~ 2017년 12월 : SK(주) C&C 제안전략 Consultant
· 1999년 8월 ~ 2004년 11월 : SI Computer programer
· 2017년 10월 ~ 현재 : 정보통신기술진행센터(IITP) 평가위원
· 2019년 : 한국정보화진흥원(NIA) RPA 자문위원
· 관심분야 : Business Automation, 융합기술연구, AI
· 저서 : 고객의 마음을 움직이는 제안전략(2019)
· 번역서 : Success Factor Modeling (Robert B. Dilts, 2018)
· E-Mail : hyunyg@ajou.ac.kr

임 정 택(Lim, Jung Teak) [정회원]

· 2016년 6월 ~ 현재 : SK(주) C&C AI 솔루션 아키텍트
· 2015년 1월 ~ 2016년 6월 : SK(주) C&C IT 컨설턴트
· 2012년 1월 ~ 2015년 1월 : SK(주) C&C 소프트웨어 엔니지어
· 2008년 3월 ~ 2011년 1월 : 아주대학교 경영대학 E-business 학사
· 관심분야 : AI 서비스 기획 및 아키텍처 설계, 자연어처리
· E-Mail : ryan.lim.intro@gmail.com

한 정 현(Jeong-Hyeon Han) [정회원]

· 2018년 9월 ~ 현재 : 아주대학교 공과대학 산업공학과 박사과정
· 2017년 7월 ~ 현재 : SK 주식회사 C&C Digital 총괄 Vitality 그룹
· 2007년 1월 ~ 2017년 6월 : SK 주식회사 C&C Business Management
· 1997년 7월 ~ 2006년 12월 : SK 주식회사 C&C Software Engineering
· 관심분야 : 융합기술연구, 디지털 헬스케어, 빅데이터
· E-Mail : hann@ajou.ac.kr

채 우 리(U-ri Chae) [정회원]

· 2017년 3월 ~ 현재 : 아주대학교 공과대학 산업공학과 석박사통합과정
· 2015년 3월 ~ 2017년 2월 : 아주대학교 공과대학 산업공학과 석사과정
· 관심분야 : 융합기술연구, 직업병·직업성질환, 데이터분석
· E-Mail : chaeuri@ajou.ac.kr

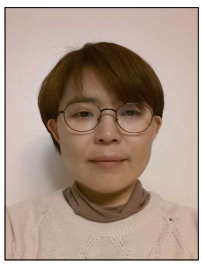
이 기 현(Lee, Gi Hyun) [정회원]

· 2018년 2월 : 남서울대학교 산업공학과(이학사)
· 2018년 3월 ~ 2020년 2월 : 아주대학교 산업공학과 석사
· 2020년 3월 ~ 현재 : 아주대학교 산업공학과 박사과정
· 관심분야 : 융합기술연구, 데이터분석, 신재생에너지
· E-Mail : black9255@ajou.ac.kr

고 진 덕(Ko, Jin Deuk) [정회원]

· 2019년 3월 ~ 현재 : 아주대학교 산업공학과 석박사통합과정
· 2010년 9월 ~ 2015년 6월 : University of Waterloo (B. Eng)
· 관심분야 : 융합기술연구, 데이터분석, 자연어처리, AI
· E-Mail : jdko@ajou.ac.kr

조 영 희(Cho, Young Hee) [정회원]



· 2009년 8월 : 단국대학교 전자계산학과 이학박사
· 2014년 10월 ~ 현재 : 아주대학교 LINC+사업단 전산담당
· 2009년 9월 ~ 2011년 6월 : 모젼스랩
· 관심분야 : 빅데이터, 패턴인식, 기계학습
· E-Mail : zeroch@ajou.ac.kr

이 주 연(Lee, Joo Yeoun) [종신회원]

· 2002년 2월 : 인하대학교 대학원경영학박사
· 2014년 9월 ~ 현재 : 아주대학교 공과대학 산업공학과 교수
· 2015년 2월 ~ 2018년 1월 : 산업통상자원부 산업융합촉진 국가옴부즈만(차관급)
· 2016년 7월 ~ 2019년 6월 : 한국빅데이터서비스학회 학회장
· 2007년 7월 ~ 2011년 6월 : 한국산업정보학회 회장
· 2011년 12월 ~ 2014년 3월 : POSCO ICT 그린사업부문장(전무)
· 2005년 2월 ~ 2011년 11월 : SK C&C 전략마케팅본부장(상무)
· 1999년 12월 ~ 2005년 1월 : Oracle 전략솔루션실장(Director)
· 관심분야 : 산업융합기술, 비즈니스인텔리전스, 서비타이제이션, 스마트그리드
· E-Mail : jooyeoun325@ajou.ac.kr