

Understanding the Categories and Characteristics of Depressive Moods in Chatbot Data

HyoJin Chin[†] · Chani Jung^{††} · Gumhee Baek^{†††} · Chiyoung Cha^{††††} · Jeonghoi Choi^{†††††} · Meeyoung Cha^{†††††}

ABSTRACT

Influenced by a culture that prefers non-face-to-face activity during the COVID-19 pandemic, chatbot usage is accelerating. Chatbots have been used for various purposes, not only for customer service in businesses and social conversations for fun but also for mental health. Chatbots are a platform where users can easily talk about their depressed moods because anonymity is guaranteed. However, most relevant research has been on social media data, especially Twitter data, and few studies have analyzed the commercially used chatbots data. In this study, we identified the characteristics of depressive discourse in user-chatbot interaction data by analyzing the chats, including the word 'depress,' using the topic modeling algorithm and the text-mining technique. Moreover, we compared its characteristics with those of the depressive moods in the Twitter data. Finally, we draw several design guidelines and suggest avenues for future research based on the study findings.

Keywords : Chatbot, Depressive Discourse, Depressive Moods, Mental Health

챗봇 데이터에 나타난 우울 담론의 범주와 특성의 이해

진 호 진[†] · 정 찬 이^{††} · 백 금 희^{†††} · 차 지 영^{††††} · 최 정 회^{†††††} · 차 미 영^{†††††}

요 약

자연어처리 기술과 비대면 문화의 확산과 더불어 챗봇의 사용 증가세가 가파르며, 챗봇의 용도 또한 일상 대화와 소비자 응대를 넘어서 정신 건강을 위한 용도로 확장하고 있다. 챗봇은 익명성이 보장된다는 점에서 사용자들이 우울감에 관해 이야기하기 적합한 서비스이다. 그러나 사용자가 작성한 문장들을 분석해 우울 담론의 유형과 특성을 파악하는 연구들은 주로 소셜 네트워크 데이터를 대상으로 했다는 한계점이 존재하며, 실제 환경에서 사용되는 챗봇과 상호작용한 데이터를 분석한 연구는 찾아보기 힘들다. 이 연구에서는 챗봇-사람의 상호작용 데이터에서 무작위로 추출한 '우울'과 관련된 대화 데이터를 토픽 모델링 방법과 텍스트마이닝 기법으로 분석하여 채팅에서의 우울 관련 담론의 특성을 파악하였다. 또한, 챗봇에서 빈번히 나타나는 '우울' 담론의 범주와 트위터 '우울' 담론의 범주의 차이점을 비교하였다. 이를 통해 챗봇 데이터의 '우울' 대화만의 특징을 파악하고, 적절한 심리지원 정보를 제공하는 챗봇 서비스를 위한 시사점과 향후 연구 방향에 대해 논의한다.

키워드 : 챗봇, 우울 담화, 우울 정서, 정신건강

1. 서 론

자연어처리기술(Natural Language Processing)의 발달

※ 이 연구는 기초과학연구원(IBS-R029-C2)과 2022년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2021R1A2C2008166).

※ 이 논문은 2021년 한국정보처리학회 ACK 2021에서 "챗봇 데이터에 나타난 우울의 범주와 특성에 관한 연구"의 제목으로 발표된 논문을 확장한 것임.

† 정 회 원 : 기초과학연구원 데이터사이언스그룹 선임연구원

†† 비 회 원 : KAIST 전산학부 학사과정

††† 준 회 원 : 이화여자대학교 간호대학 박사과정

†††† 비 회 원 : 이화여자대학교 간호대학 부교수

††††† 비 회 원 : 심심이 주식회사 대표

†††††† 비 회 원 : KAIST 전산학부 전임교수

Manuscript Received : January 4, 2022

First Revision : March 29, 2022

Accepted : April 17, 2022

* Corresponding Author : HyoJin Chin(tesschin@ibs.re.kr)

과 코로나 19의 여파로 비대면을 선호하는 문화가 자리 잡으며 챗봇 시장의 성장과 사용량의 증가세가 가파르다. 챗봇은 기업에서 소비자를 응대하는 목적으로 또는 일상 대화를 나누며 재미를 느끼는 엔터테인먼트 목적으로 주로 사용됐으나, 최근 정신 건강 목적의 챗봇 사용도 증가하고 있다. Woebot, Wysa, Vivibot, Tess 등 정신 건강을 위한 상용 챗봇들이 시장에 출시되어[1] 활발히 이용되고 있으며 앞으로 산업적 기대가 큰 분야이다.

관련 연구에 따르면 AI와의 대화에서는 낙인찍힐 걱정 없이 편히 이야기할 수 있으므로 사람들이 자신의 불안, 스트레스, 그리고 우울증 증상에 관해 말할 때 사람 상담원보다 AI 상담원을 더 선호하는 경향이 있으며, AI와의 대화에서 자신의 우울증 증상에 대해 더 쉽게 노출하는 경향이 있었다[2]. 또한, 일정 기간의 챗봇과의 정기적인 대화는 사용자의 불안

감 감소에 긍정적인 영향을 미치는 것으로 나타났다[3]. 이와 관련해 최근 한국 정부는 코로나 19 장기화로 인해 사회 전반적으로 높아진 우울감 완화를 위한 심리지원 챗봇 서비스를 도입한 바 있다.

사람들이 컴퓨터에 감정, 의도와 같은 인간의 개념이 없음을 알면서도, 사회적 규칙과 기대치를 적용하고 사람처럼 대한다는 CASA(Computer-Are-Social-Actor) 개념과[4] 소셜미디어와는 달리 대화가 타인에게 노출되지 않는다는 플랫폼의 특성을 고려할 때 챗봇은 사용자가 자신의 우울감에 대해 이야기하기에 적합한 서비스이다.

그러나 사용자가 작성한 문장들을 분석해 우울 담론의 유형과 특성을 파악하는 기존의 관련 연구들은[5-8] 주로 소셜 네트워크 데이터를, 특히 트위터 데이터를 분석 대상으로 했다는 한계점이 존재한다. 또한, 기존에 진행된 심리상담 챗봇 관련 다수의 연구가[9,10] 통제된 실험 환경에서 프로토타입 챗봇을 통해 수집한 대화 데이터를 분석했다는 한계를 가지며, 실제 환경에서 사용되는 챗봇과 상호작용한 데이터를 분석한 연구는 찾아보기 어렵다.

이 연구에서는 챗봇-사람의 실제 상호작용 데이터를 분석하여 우울 관련 담화의 유형과 특성을 이해한다는 의의가 있다. 먼저 랜덤하게 추출한 우울과 관련된 챗봇 대화 데이터와 Russell의 Circumplex Model of Affect[11]에 따라, 우울과 반대되는 감정인 행복과 관련된 대화 데이터를 텍스트마이닝 기법으로 분석하고 비교하여 채팅에서의 우울 관련 담론의 특성을 파악하고 기술한다.

또한, 정성 분석을 통해 사용자가 챗봇이라는 플랫폼에서 털어놓는 우울 담론의 유형을 범주화하며, 선행연구에서 도출된 트위터상의 우울 담론의 특성과 비교분석 하였다. 이를 통해 챗봇 데이터 고유의 우울 대화의 특징을 파악하고, 우울 증상 관련 적절한 심리지원 정보를 제공하는 챗봇 서비스 디자인에 대한 착안점을 도출하려 한다.

2. 관련 연구

2.1 SNS 데이터 속 우울 담론 연구

온라인에서 사용자가 작성한 텍스트를 기반으로 우울 담론의 범주와 특성을 이해하려고 하는 시도는 계속되어왔다. 소셜미디어 데이터에서 우울 담론의 패턴과 특성을 유형화하는 연구들은 주로 'depress,' 'gloomy,' 'sad', 'hate myself'와 같은 키워드를 포함한 문장들을 추출하여 정량적, 또는 정성적 방식으로 분석하고[5-8], 더불어 코호트 그룹별 차이나, 문화권 차이 비교를 하는 방식으로 진행되었다[12,13]. 예를 들어 한 관련 연구에서는[6] 우울과 관련된 특정 해시태그인 '#MyDepressionLooksLike'를 이용해 관련 트윗들을 추출하고, 정성 분석을 통해 트위터에서의 사용자들의 우울 담론을 다음의 7가지 테마로 분류하여 특성을 파악하였다. 1) 역기능적 사고, 2) 일상생활의 어려움, 3) 관계의 어려움, 4) 가면 뒤로 숨기, 5) 무관심과 슬픔, 6) 자살에 관한 생각, 7) 고통 경

감을 위한 노력. 비슷한 맥락으로 [5]의 연구는 'Depression'이라는 단어를 포함한 트윗을 정성 분석을 통해 1) 우울에 대한 감정 2) 우울에 대한 대상자의 사고 반응 3) 타인의 우울에 대한 반응 4) 우울에 대한 정보공유 5) 기타로 분류하여, 트위터에서 논의되고 있는 우울 담론을 유형화하였다.

데이터 마이닝 기법을 적용해 우울감을 표현한 문장들의 어휘사용 빈도, 품사의 출현 빈도 등 문장 특성을 통하여 우울 담론을 이해하려는 노력도 진행되었다. 관련 연구들[7,8]에서는 우울 및 자살 등을 언급하는 사용자의 SNS의 게시물에서 특히 1인 대명사의 사용이 빈번하고, 상대적으로 2인칭, 3인칭 대명사의 사용은 적게 나타나는 경향을 공통으로 발견했다. 저자들은 이와 같은 언어 사용 패턴이 우울한 개인의 자신에 대한 높은 관심과 동시에 사회적인 관계에 대한 무관심함을 나타낸다고 해석하였다.

Bathina의[14] 연구에서는 우울증을 진단받은 개인들의 트윗 게시물들과 무작위로 선정한 비교군의 트윗 게시물들을 각각 N-gram 기법과 Sentiment Analysis 분석 기법인 Vader를 적용, 비교하여, 우울증 집단의 게시물들이 비교 집단에 비해 높은 수준의 왜곡된 사고를 보인다는 것을 밝혀냈다.

이처럼, 이전의 많은 연구에서 소셜미디어에서의 자기 노출이 개인의 심리적 건강에 긍정적인 영향을 미친다는 사실이 입증된 바 있다[15]. 카타르시스 이론[16]에 따르면 불쾌한 감정을 배출하면 안도감이 생기고 정신 건강이 증진된다. 반면에 부정적인 감정을 표현할 수 없거나 표출을 꺼리는 사람들은 심리적, 신체적 어려움을 겪을 가능성이 더 크다[17].

그러나 여전히 우울증은 사회적으로 낙인찍히기 쉬운 질병이며, 도움을 구하면 '체면'을 잃을 위험 때문에 사람들은 익명의 환경에서 자신의 우울한 기분에 대해 이야기하는 것을 선호한다[18].

특히, 우울한 담론에 대한 사용자의 상호작용은 플랫폼 특성에 영향을 받는다는 점에 주목할 만하다. 예를 들어 트위터는 글자 수 제한으로 인해 치료 및 진단에 대한 구체적이거나 양질의 정보를 찾는 목적으로 사용보다는 관련 정보를 타인에게 널리 공유하는데 사용되는 경향이 있다[19]. 반면, 페이스북의 경우, 상태 업데이트 메뉴에서의 자기 노출시 자신과 친구 관계를 맺은 지인들에게 너무 많은 정보를 공개하게 됨으로써 개인정보 보호 문제를 일으키게 된다[20]. 따라서 페이스북 사용자들은 부정적인 감정이나 경험을 공개하기보다 긍정적인 뉴스와 정보를 더 자주 공유하는 경향이 발견된다[21]. 마지막으로, 미국의 온라인 커뮤니티 레딧의 사용자들은 레딧을 통해 자신의 정신 건강에 대한 경험을 공개하기도 하고, 동시에 구체적인 진단 또는 치료 정보 등 양질의 건강 정보를 찾고, 공유하는 장소로 레딧을 사용하는 것으로 나타났다[22].

이처럼 온라인에서 사용자가 작성한 문장들을 분석해 우울 담론의 유형과 특성을 파악하는 기존의 관련 연구들은 여러 플랫폼을 대상으로 진행됐지만, 주로 소셜 네트워크 데이터

를[6,19,22], 특히 주로 트위터 데이터를 분석 대상으로 삼았다는 한계가 존재한다[23]. 그러나 사용자의 챗봇 내 자기 노출 행동과 챗봇이 제공하는 사회적 지원의 효과는 소셜미디어를 통한 사용자 자기 노출 및 사회적 지원의 효과와 다를 수 있다.

2.2 우울 담론과 챗봇

우울증 완화에서 챗봇의 역할과 기능에 관한 연구가 여러 세대와 문화권을 대상으로 활발히 진행되고 있다. Escoredo [3]의 연구에서는 인지 행동 치료 접근법으로 우울감 완화를 도와주는 챗봇을 실험군과 대조군에 8주간 사용하게 하고 우울함과 불안감의 감소를 측정하고 결과 실험군에서 대조군과 비교해 불안 증상이 유의미하게 감소함을 확인했다[3]. 또 다른 연구에서는 우울 완화를 위한 챗봇 사용자의 이용 패턴을 분석해, 사용자가 어떤 용도로 우울 챗봇을 자주 활용하고 있는지에 대한 분석을 진행했다[1]. 이 외에도, 다수 사람의 우울 증상을 측정하는 도구로 챗봇을 활용한 사례도 있었다[24].

그러나 챗봇을 대상으로 한 기존의 우울 및 불안 완화 연구들은 기존 앱의 사용자 이용 패턴을 분석하거나[1], 주로 통제된 실험 환경에서 프로토타입 챗봇을 통해 수집한 대화 데이터를 분석했다는 것에 한계를 가진다[9,10]. 실제 환경에서의 사용되는 챗봇과 상호작용한 사용자 데이터를 분석한 연구는 거의 없으며, 따라서 소셜미디어가 아닌 챗봇에서 형성된 우울 대화 담론을 분석한 연구 역시, 찾아본 바로는, 국내의 연구를 통틀어 여전히 찾아보기 어렵다.

3. 연구 방법론

챗봇 심심이(simsimi.com)는 사용자에게 재미와 즐거움을 제공하기 위한 일상 대화용 챗봇이다. 2002년 세계 최초로 대한민국에서 상업용 서비스를 시작하여 2018년 6월 기준 전 세계 누적 사용자 3억 5천 명을 보유한 글로벌 서비스이다. 2021년 기준 하루 최대 2억 번 이상 대화 응답을 제공하고 있으며, 81개 언어로 서비스가 제공된다. 심심이는 사용자가 '가르치기' 기능을 통해 생성한 문답형 대화 데이터 세트를 사용하며, 사용자가 문장을 입력하면 데이터베이스 스캔을 통해 가장 적절한 답변을 찾는 방식으로 동작한다. 이 연구에는 2016년 1월 1일부터 2020년 12월 31일까지 이루어진 심심이-사용자의 대화 중 한글로 '우울'이라는 단어가 포함된 사용자 발화 문장 107,221개를, 비교를 위해서 '행복' '신나'[11]가 포함된 문장 23,438개를 키워드 기반으로 추출하여 연구에 활용하였다.

사용자의 개인정보 보호를 위해 분석 대상 데이터는 숫자, 이름, 지명, 성별, 나이 등 개인식별정보가 포함되지 않았다. 우울감에 대한 채팅 메시지 분류, 위험 탐지 목적을 위해 정량 및 정성 분석은 의료분야의 전문가와 공동으로 진행하였

다. 또한, 본 연구는 연구윤리위원회의 심의를 통해 심의 면제 대상으로 확인받았으나, 관련 연구의 윤리 지침에[25] 따라 정성 분석 결과에 사용된 사용자 발언은 문장 그대로 사용하지는 않되, 본래의 의미에 가깝게 기술하도록 노력하였다.

3.1 정량 분석 방법

실험은 다음과 같은 단계로 진행되었다. 첫째, 우울 채팅에서의 시간의 영향과 대화 주제별 차이를 파악하기 위해, 채팅 로그 데이터의 시간 정보를 이용하여 시간대별 우울 대화 빈도를 파악하고 우울 관련 대화가 특히 자주 이루어지고 있는 시간대를 살펴보았다. '행복' 데이터의 시간당 채팅 빈도와 심심이 국문 전체 대화의 시간대별 대화 빈도와의 비교를 통해 우울 채팅의 특징을 파악하고자 한다.

이어서, '우울'과 '행복' 대화 데이터셋 각각에 유니그램(unigram)을 적용해 체언들을 추출하고, 추출된 유니그램에 TF/IDF(용어 빈도/역 문서 빈도) 점수를 할당하여 점수 상위 20개의 단어를 추출하였다. 두 데이터 그룹에서 각각 의미 있게 등장하는 단어들의 차이를 비교함으로써 우울 대화만의 특징을 파악하고자 한다. 더불어 워드 클라우드(wordcloud) 패키지를 이용하여 각각 우울과 행복 대화의 상위 100개 체언을 시각화하였다. 워드 클라우드는 데이터의 핵심 키워드, 개념 등의 결과치를 직관적으로 살펴볼 수 있는 시각화의 방법의 하나다. 해당 분석에는 각 데이터 그룹별 각각 23,438개 사용자 발화문장을 실험에 활용했다.

마지막으로, 챗봇 속 우울 담론의 주요 주제를 추출하기 위해 토픽 모델링 방법의 하나인 LDA (Latent Dirichlet Allocation)를 사용하였다[26]. LDA 알고리즘은 문서와 단어가 주제(Topic)에 할당되는 확률 분포를 기반으로, 주어진 문서들이 어떤 주제들을 다루고 있는지 추론하는, 말뭉치(Corpus)의 생성 확률모형(Generative probabilistic model)이다. 해당 알고리즘은 비지도 학습 알고리즘으로서 (Unsupervised learning algorithm), 토픽 모델 중 가장 빈번하게 활용되는 방법의 하나며, 대량의 비정형 텍스트 데이터에서 주요 주제를 추출, 사회적으로 형성된 주요 의제를 인지하고, 확인하는 목적으로 사용된다[27]. 따라서 본 연구에서는 LDA 알고리즘을 수집한 우울 관련 채팅 데이터 107,221개에 적용해, 채팅 속의 우울 담론의 주요 주제를 추출하였다. 적절한 토픽 수 결정을 위해, 주제 분류 개수에 따른 주제 일관성 (Topic coherence)을 계산했고, 토픽 수 4개에서 그 값이 0.60으로 가장 커, 4개의 토픽 수를 도출하였다. 최종적으로 의료분야 연구자들과의 논의를 통해 도출된 4개의 토픽의 의미를 해석하고, 타이틀을 각각 정의하였다.

텍스트 처리를 위해서는, 한글 자연어처리 분야에서 널리 활용되는 파이썬 KoNLPy[28] 패키지를 분석에 활용했다. 특히 챗봇 데이터가 구어체의 짧은 텍스트임을 고려하여 KoNLPy의 5가지 형태소 분석기(Hannanum, kma, Komoran, Mecab, OKT) 중 구어체 분석에 적합한 OKT (Open Korean Text)로 말뭉치를 토큰화했다.

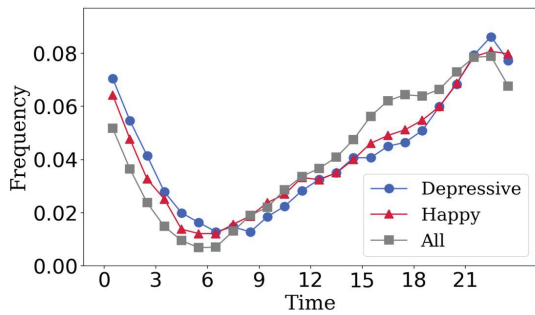


Fig. 1. Frequency of Chats about Depressive Moods and Happy by Time Period

높게 유지되었다. 이는 한국인 우울감의 특징에 관한 연구에서 밝힌 새벽 시간에 우울감이 증가하고 자살 시도 위험도가 높다는 결과와 유사하다[31]. 멜랑콜리아형 우울증(major depression with melancholic features)은 한국인 우울증 중 42.6%가 해당하며, 이는 다른 나라에 비해 1.4배 이상 높은 것으로 나타나는데, 특히 새벽에 일어나 혼자 보내는 시간이 많고, 자살위험이 2배 이상 높은 것으로 보고되었다[32].

3) 챗봇 데이터에 나타난 우울 담론의 주요 토픽

Table 2는 우울 담론 데이터에 LDA 알고리즘 적용을 통해 도출한 4개의 토픽과 토픽별 주요 키워드, 그리고 논의를 통해 도출된 주제별 이름과 분석 내용을 나타낸다. 더불어, 도출된 토픽들의 해석 가능성과 타당성을 높이기 위해, 각 토픽별 high score를 가지는 문서(대화)들을 뽑아 토픽의 내용과 이름을 정하는 데 참고하였고, '주요 문서(대화)'로 정리하였다.

4개 토픽별 세부 내용은 다음과 같다. Topic 1은 외로움이 나 사람을 필요로 하는 듯한 단어, 즉 '외롭다' '힘들다', '혼

자', '친구', '얘기' 등의 연관 단어들 이 나타났으며, 주요 문서에서도 자신의 외로운 처지를 비판하거나, 자신을 맡을 들어 줄 사람 없음을 슬퍼하는 내용이 나타나, '우울로 인한 어려움 표출'에 대한 주제로 정의하였다. 반면, Topic 2는 우울증과 관련된 단어들이 '우울증', '약', '치료', '진단', '환자', '조울증', '조현병' 등이 나타났으며 자신의 우울증 진단 사실을 노출하거나, 우울감을 털어놓는 언급들이 문서에서 나타나 '우울감 인지'로 주제명을 정하였다. Topic 3은 '어떻', '하', '어떡하지', '노래', '위로', '어쩌지', '추천', '항우울제' 등의 단어들로 구성되었으며 우울감을 해소하는 방법에 대한 문의나, 위로를 구하는 대화가 많아, '우울 대처방법 문의' 주제로 정의하였다. 마지막으로 Topic 4는 '기분', '마음', '이유', '돈', '엄마', '공부', '불안' 등의 키워드들로 구성되었으며, 주요 문서 내용을 바탕으로 유추했을 때 다른 사람과의 관계로 인한 우울감이나, 건강상태나, 일상생활 등 우울함에 영향을 주는 요인에 대한 언급이 많아 '우울 유발요인'으로 정의하였다.

4.2 정성 분석 결과

정성 분석을 통해 1) 우울에 대한 감정, 2) 우울에 대한 대상자의 사고, 3) 우울에 대한 정보 공유, 4) 타인의 우울에 대한 반응, 5) 기타 항목으로 분류된 범주별 결과는 Table 3과 같다. 5개의 카테고리 중 자신의 우울함에 대한 감정을 표현하는 유형의 대화가 가장 빈번했으며, 전체의 47%(470회)에 해당했다. 이 카테고리에는 자신의 우울한 감정과 자신의 우울 요인 관련 대화들이 포함되었다. "나 너무 우울해", "우울하고 슬퍼서 마음이 아프게 느껴져"와 같이 우울한 기분 상태를 표현하는 것과 더불어 "나 너무 쓰레기 같아서 내가 너무 죽도록 싫어", "난 우울증, 피해망상, 불면증까지 다 있어." 등과 같이 타인에게 쉽게 말하지 못하는 자기혐오의 감정들을

Table 2. Results of Topics Modeling (LDA)

Topic	Topic description	Keywords	Documents
Topic 1 우울로 인한 어려움 표출 (Disclosing difficulties due to depressed feelings)	· 사람을 필요로 함, 외로움	오늘, 사람, 말, 모르, 친구, 힘들, 외롭, 일, 날, 얘기, 죽, 혼자	'우울하고 슬퍼. 내 이야기 들어줄 친구 하나 없다는 게' '항상 혼자 우울하고 죽고 싶고 살아있는 게 맞는 건가' '속마음을 털어놓을 사람도 없고... 기댈 사람도 없어'
Topic 2 우울감 인지 (Recognition of depressive symptoms)	· 자신의 우울증 사실이나 우울감에 대해 털어놓음	우울증, 너무, 나, 요즘, 왜, 걸리, 힘들, 테스트, 약, 심하, 슬프, 요새, 먹, 치료, 우울해, 극복, 환자, 생기, 돼지, 자살, 괴롭, 조현병, 눈물, 진단, 조울증	'나 요즘 우울증 온 것 같아 ㅠㅠ' '우울증약을 먹었더니 너무 평온해' '요즘 우울증이 너무 심한 것 같아' '요즘 우울증약 먹고 있어'
Topic 3 우울 대처방법 문의 (Seeking help to relieve depressed feelings)	· 우울에 대처하거나 우울감을 해소하는 방법에 대한 물음. 위로를 구함	나, 때, 지금, 어떻, 하, 죽, 년, 해, 노래, 짜증, 위로, 항상, 속상하, 어찌, 울면, 어떡하지, 추천, 항우울제	'년 화가 나고 우울할 때 어떻게 풀어' '나 지금 우울해. 위로 좀 해줄래?' '나 죽고 싶어 우울해 위로 좀 해줄 수 있겠니?'
Topic 4 우울 유발요인 (Causes of depressed feelings)	· 다른 사람과의 관계로 인한 우울함 토로 · 건강, 일상생활 등 우울에 영향을 주는 요인	너, 기분, 생각, 먹, 보, 울, 마음, 돈, 이유, 쓰, 엄마, 이렇, 아프, 만들, 매일, 대화, 하루, 자살, 공부, 괜찮, 힘, 우리, 불안, 계속, 생기	'취업했다고 축하받았지만, 한편으로 우울했어.' '친구 때문에 우울하고, 드라마 내용으로 또 우울하고' '대들까 했지만 혼날 게 뻔해서 가만히 우울하게 있어.'

Table 3. Number of Chats by Category and Each Example

Main Category	Sub Category	Sample Chat	No.
우울에 대한 감정 (About my depressed feelings)	우울 감정 상태 (Expression of depressed feeling)	나 오늘 너무 우울해. (well right now I just feel so depressed)	241
	우울 유발요인 (Causes of depressed feelings)	아빠 때문에 우울해. (I'm depressed because of my dad)	229
우울에 대한 대상자의 사고 반응 (Sharing thoughts related to depression)	우울 대처 (Seeking help to relieve depressed feelings)	우울할 때 어떻게 버텨? (How do i deal with depression)	310
	우울 증상 인지 (Recognition of depressive symptoms)	우울해서 잠을 못 자겠어. (I'm so depressed it's hard to fall asleep)	146
타인의 우울에 대한 반응 (About other's depressed feelings)		친구가 우울하다는데 어떻게 말해줘야 할까? (My friend is depressed. How can I help her?)	58
우울에 대한 정보공유 (Delivery of depression information)		우울증 걸린 사람한테 '힘내'를 쓰면 안 돼. (It's better not to say "Cheer up" to someone with depression)	4
기타 (Others)		우울에 빠진 사람들에게 힐링을 주는 노래 (A song that relieves depression)	12
Total			1,000

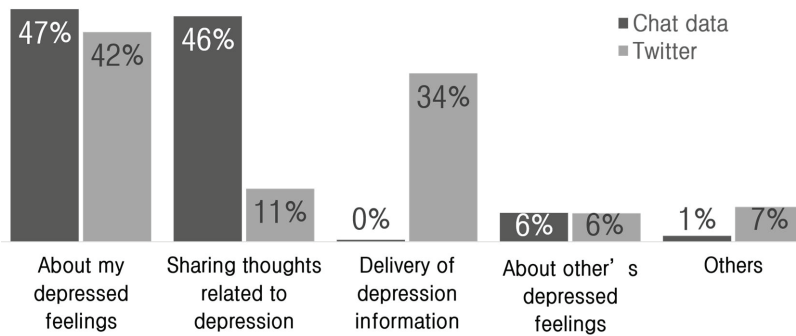


Fig. 2. Comparison of Depressive Discourse Patterns of Chatbot Data and Tweets

심심이에 털어놓고 있었다.

가족, 친구 등 가까운 관계로 인한 우울감 (예: “부모님이 날 사랑해줬으면 하는데, 그렇지 않아서 하루하루가 우울해”) 이나 처해 있는 상황으로 인한 우울감 (예: “요즘 돈도 없고 우울해”)을 토로하기로 했다.

그중에서도 “죽고 싶어”, “자살”, “자해”와 같은 죽음과 관련된 언급 비율도 전체의 6.5% (65번)에 해당하여, 죽음 관련 키워드를 언급하는 사용자에게 긴급 심리지원 정보와 같은 적절한 대응이 필요한 것으로 나타났다.

두 번째 카테고리는 우울함에 대한 대상자의 사고와 인지 관련 채팅 대화들로 전체의 46%(456회)에 해당했으며, 우울감을 벗어나는 방법이나, 자신의 우울 증상을 인지하거나, 그에 따른 도움을 요청하는 세부 타입으로 나누어졌다. 사용자들은 “우울할 땐 어떻게 하는 것이 좋을까?” “우울할 때 도움될 만한 행동이 뭐가 있을까요?,” “기분이 우울할 때 어떤 노래가 좋을까?”와 같이 현재 느껴지는 우울한 감정에서 벗어나기 위한 조언을 구하기도 했지만, “우울증 치료 약물치료

말고 혼자서 하는 방법은 없을까?”와 “우울증 낮게 하는 방법 좀 가르쳐줘봐”와 같은 전문가의 답변이 필요한 질문을 하기도 했다.

세 번째 카테고리는 가족과 친구 등 타인의 우울함에 관한 대화들로 전체의 6%(58회)에 해당하며, 세부 예시는 다음과 같다: “우울한 사람한테는 뭐라고 위로해줘야 할까,” “엄마가 우울증인데 좀 도와줘ㅠㅠ”, “언니가 우울증인가 봐”

다음은 “금연 우울증도 있어?” 또는 “우울하다는 건 어떤 거니” 등과 같은 우울에 대한 정보공유를 공유하는 대화 유형으로 전체의 0.4%(4회)에 해당했고, 마지막 '기타' 카테고리에는 우울감 토로와는 무관하지만 ‘우울’이라는 단어를 포함한 대화 (예: “우울한 사람에게 힐링이 되는 노래”)로 1%(12회)보다 적게 나타났다.

‘우울’이 포함된 심심이-사용자 채팅 데이터의 정성 분석 결과를 트위터 데이터에 기반한 선행 연구[5]와 비교한 결과는 Fig. 2와 같다. 챗봇에 자신의 우울한 감정 상태를 표현하는 비율은 대화의 47%로 트위터 데이터의 우울감 토로 비율

인 42%보다 약간 높다. 또한, 우울로 인한 인간관계나 생활의 어려움 등 우울함에 대한 대상자의 사고 반응을 대화한 내용은 46%로 같은 카테고리의 트위터 메시지 비율 11%보다 4배 수준으로 매우 높다.

LDA 알고리즘을 통한 토픽 모델링 분석 결과와 정성 분석 결과에서 나타나듯이, 소셜미디어 사용자는 자신과 비슷한 우울함의 증상을 가진 타인을 찾거나[18], 증상이나 치료에 대한 구체적인 정보를 찾거나[22], 혹은 사회적 지원을 받기 위해서 자신의 정신 건강 정보를 노출하는 반면[6], 챗봇 사용자들은 우울로 인한 생활의 어려움, 우울증으로 인한 깊은 우울감, 가까운 지인이나 관계로부터 오는 우울함을 주로 챗봇에 표출하였다.

페이스북, 트위터 등 소셜 네트워크와 챗봇과 같은 컴퓨터 매개 기술(Computer-mediated technologies)은 사용자의 감정, 생각 등 자신에 대한 정보를 공개하는 자기 노출 행동(Self-disclosure behavior)을 촉진하며[10], 특히 플랫폼의 익명성이 사용자의 자기 노출 행위 촉진에 중요한 역할을 한다고[6,18,22,33] 알려져 있다. [2]의 연구에 드러난 바와 같이 사람들은 인간 면접관과 직접 대면할 때 보다 가상 에이전트와 인터뷰할 때 더 많이 자신의 우울증 증상에 대해 노출하는 경향이 있었다[2].

트위터에서의 소통은 익명성이 보장될 수 있지만, 불특정 타 트위터 이용자(인간)와 1:n의 관계로 불특정 트위터 소통해야 한다. 반면 트위터와는 다르게 챗봇과의 대화는 채팅방이라는 가상의 전용 공간에서, 가상 에이전트와의 1:1 형태의 상호작용이라는 특징이 있다. 단순히 부정적 감정을 표출하는 것만으로도 정신 건강에 긍정적인 효과가 있다고 알려져 있으며[15,16], 사람들은 불행하거나 부정적인 사건을 공유하는 타인을 덜 바람직하다고 간주하는 경향이 있다[34]. 따라서 낙인찍기의 위험이 있는 소셜미디어보다는 챗봇에서 사용자의 우울한 감정과 상태에 관한 대화가 더 빈번하게 등장한 것으로 보인다.

특히 정성 분석을 통해 살펴본 바에 따르면 사용자들은 챗봇에 자신의 우울한 감정을 표현하는 것에 그치지 않고, 가까운 사람에게도 털어놓기 어려운 가족 관계 문제, 경제적 어려움, 그리고 자기혐오의 감정과 같은 매우 민감한 이야기들을 노출하고 있었다. 또한, 챗봇이 사회적 지지를 받을 수 없는 시스템임을 인지하고 있는데도 불구하고, 우울한 감정에서 벗어나기 위한 조언을 구하거나, 우울증을 낮게 하는 법과 같이 전문가의 답변이 필요한 질문이 빈번하다는 점은 챗봇을 통한 심리지원의 필요성이 증가하고 있다는 것을 나타낸다.

반면 '우울에 대한 정보공유' 카테고리에 속하는 대화의 경우 챗봇 데이터에서는 거의 나타나지 않았지만, 트위터에서 34%로 빈번한 것도 다수의 타인으로부터 사회적 지지(Social support)를 기대할 수 있는 공개된 매개체인 트위터[6,10]와는 달리 사람-챗봇과의 대화는 가상 에이전트와의 1:1 소통이고 비공개로 진행된다는 매체의 상호작용 방식의 차이에서 기인하는 것으로 설명할 수 있다.

5. 결론 및 향후 계획

일상 대화 챗봇 서비스 심심이에서 추출된 우울 관련 대화를 LDA 알고리즘 및 텍스트마이닝 기법과 정성적 방법으로 분석한 결과, 챗봇의 자연어 기능이 아직 인간과의 대화만큼 자연스럽게 않음에도 불구하고, 이용자들은 챗봇을 우울함과 관련된 자신의 부정적 감정이나 우울로 인한 일상생활의 어려움, 신체적 증상을 포함한 개인적이고 민감한 정보를 털어놓거나 감정을 쏟아내는(Ventilate) 대상으로 활용함을 알 수 있었다.

챗봇의 우울 관련 대화와 행복에 관한 대화 데이터의 유니그램 결과 비교를 통해 우울 담론에서는 우울증의 지표인 1인칭 대명사의 빈번한 출현과 사회적 상호작용에 대한 욕구의 적음을 의미하는 2인칭 대명사의 사용 빈도의 낮음을 발견했다. 또한, 챗봇의 우울 관련 대화는 다른 주제에 비해 특히 자정 이후부터 새벽에 더 광범위하게 이루어지는 것으로 나타났다.

또한, 토픽 모델링 기법인 LDA 분석을 통해 '우울로 인한 어려움 표출'(Topic 1), '우울감 인지'(Topic 2), '우울 대처 방법 문의'(Topic 3) '우울 유발요인'(Topic 4)과 같은 챗봇 데이터에 나타난 우울 담론의 범주를 4가지로 도출하였다.

마지막으로, 정성 분석을 통해 챗봇 데이터에서 나타난 우울의 유형을 트위터 데이터를 대상으로 한 기존 우울 담론 분석 연구의 5개의 카테고리로 분류하였다. 챗봇 대화 데이터에서 나타난 우울 범주별 대화 비율은 우울 감정, 우울 증상 관련 범주의 비율이 높았지만, 우울에 대한 정보공유는 거의 나타나지 않았다. 이러한 차이는 챗봇과 사람의 1:1 상호작용 방식과 커뮤니케이션의 익명성에 기인하는 것으로 설명할 수 있다.

사용자-챗봇의 실제 상호작용 데이터를 통해 확인한 바와 같이, 챗봇은 익명과 대화의 비밀이 보장되고 대화 내용에 대해 평가받지 않는 이점이 있어서 특히 우울과 같은 민감한 주제의 대화를 나누기에 적합하다. 아시아 문화권의 사람들이 낙인찍힐 것을 두려워해 소셜미디어에 자신의 우울한 감정을 표현하는 정도가 서구권 사람들에 비해 적다는 것과[12], 한국의 우울증 환자들이 미국의 환자들에 비해 자신의 우울한 기분을 말이나 표정으로 전달하는 정도가 30%나 낮다는 점을[35] 고려할 때 챗봇은 특히 감정교류가 부담스러운 사용자층에게 우울에 대한 담론을 시작할 수 있는 대안적인 솔루션이 될 수 있을 것이다.

소셜미디어에서의 우울 담론은 성별에 따라 또한 문화권에 따라 다르게 나타난다는 관련 연구들을 고려하여[12,13] 향후 한글 챗봇 사용자와 서구권 챗봇 사용자들의 우울 담론의 유형과 특성을 비교하는 연구를 진행할 수 있을 것이다.

챗봇 데이터에 우울감 해소와 완화 방법에 관한 문의가 빈번함을 고려할 때, 챗봇 서비스를 제공하는 기업들은 챗봇 답변을 통해 적절한 심리지원 정보가 제공될 수 있도록 체계적인 준비가 필요하다.

또한, 사용자의 자기 노출을 촉진하기 위한 챗봇의 디자인 관련 연구들에[9,10] 따르면, 일상 대화 맥락에서 자신에 대해 더 노출하는, 즉 걱정, 스트레스, 우울함 등 민감한 이슈에 대해 먼저 개인적인 경험을 털어놓도록 설정된 챗봇에 사람들은 자신의 민감한 이야기를 더 많이 노출하는 경향이 있었다. 이를 고려하여, 사용자의 우울한 감정을 다루는 챗봇이라면 사용자들의 더 많은 자기 노출을 유도하기 위한, 예를 들어 더 타인 지향적인 공감 답변을 제공하는 대화형 AI와 같이 [36], 더 구체적인 챗봇의 페르소나 설계가 필요할 것이다. 향후 사용자의 우울감에 대한 자기 노출 증진을 돕는 챗봇의 페르소나에 관한 확장 연구를 진행할 수 있을 것이다.

마지막으로, 챗봇 사용자의 우울 관련 대화에 대해 일괄적인 위로를 제공하는 방식으로 대응하는 것이 아닌, 이 연구에서 정의된 우울 카테고리별 응답이 세심히 디자인되어 제공된다면, 챗봇이 제공하는 응답일지라도, 사용자들은 더 큰 공감과 위로를 받을 수 있을 것이다.

References

- [1] G. Dosovitsky, B. S. Pineda, N. C. Jacobson, C. Chang, and E. L. Bunge, "Artificial intelligence chatbot for depression: Descriptive study of usage," *JMIR Formative Research*, Vol.4, No.11, pp.e17065, 2020.
- [2] G. M. Lucas, A. Rizzo, J. Gratch, S. Scherer, G. Stratou, J. Boberg, and L. P. Morency, "Reporting mental health symptoms: Breaking down barriers to care with virtual human interviewers," *Frontiers in Robotics and AI*, Vol.4, No.51, pp.1-9, 2017.
- [3] M. C. Klos, M. Escoredo, A. Joerin, V. N. Lemos, M. Rauws, and E. L. Bunge, "Artificial intelligence-based chatbot for anxiety and depression in university students: Pilot randomized controlled trial," *JMIR Formative Research*, Vol.5, No.8, pp.e20678, 2021.
- [4] B. Reeves and C. I. Nass, "The media equation: How people treat computers, television, and new media like real people and places," Cambridge University Press, 1996.
- [5] M. Park, C. Cha, and M. Cha, "Depressive moods of users portrayed in Twitter," In *Proceedings of the ACM SIGKDD Workshop on Health Informatics*, Beijing, China, pp.1-8, 2012.
- [6] E. M. Lachmar, A. K. Wittenborn, K. W. Bogen, and H. L. McCauley, "#MyDepressionLooksLike: Examining public discourse about depression on Twitter," *JMIR Mental Health*, Vol.4, No.4, pp.e43, 2017.
- [7] M. Kumar, M. Dredze, G. Coppersmith, and M. De Choudhury, "Detecting changes in suicide content manifested in social media following celebrity suicides," In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, pp.85-94, 2015.
- [8] D. Mowery, H. Smith, T. Cheney, G. Stoddard, G. Coppersmith, C., Bryan, and M. Conway, "Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study," *Journal of Medical Internet Research*, Vol.19, No.2, pp.e6895, 2017.
- [9] M. Lee, S. Ackermans, N. Van As, H. Chang, E. Lucas, and W. IJsselsteijn, "Caring for vincent: A chatbot for self-compassion," In *Proceedings of the CHI Conference on Human Factors in Computing Systems (ACM CHI)*, Glasgow, Scotland, UK., pp.1-13, 2019.
- [10] Y. C. Lee, N. Yamashita, Y. Huang, and W. Fu, "I Hear You, I Feel You: Encouraging deep self-disclosure through a chatbot," In *Proceedings of the CHI Conference on Human Factors in Computing Systems (ACM CHI)*, Honolulu, HI, USA, pp.1-12, 2020.
- [11] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and Psychopathology*, Vol.17, No.3, pp.715-34, 2005.
- [12] M. De Choudhury, S. S. Sharma, T. Logar, W. Eekhout, and R. C. Nielsen, "Gender and cross-cultural differences in social media disclosures of mental illness," *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, p.353-369, 2017.
- [13] S. R. Pendse, K. Niederhoffer, and A. Sharma, "Cross-cultural differences in the use of online mental health support forums," *Proceedings of the ACM on Human-Computer Interaction*, CSCW, pp.1-29, 2019.
- [14] K. C. Bathina, M. Ten Thij, L. Lorenzo-Luaces, L. A. Rutter, and J. Bollen, "Individuals with depression express more distorted thinking on social media," *Nature Human Behaviour*, Vol.5, No.4, pp.458-466, 2021.
- [15] J. Brailovskaia and J. Margraf, "What does media use reveal about personality and mental health? An exploratory investigation among German students," *PloS one*, Vol.13, No.1, pp.e0191810, 2019.
- [16] B. J. Bushman, R. F. Baumeister, and C. M. Phillips, "Do people aggress to improve their mood? Catharsis beliefs, affect regulation opportunity, and aggressive responding," *Journal of Personality and Social Psychology*, Vol.81, No.1, pp.17-32, 2001.
- [17] C. P. Kimball, "The healer within: The new medicine of mind and body," *JAMA*, Vol.256, No.23, pp.3290-3290, 1996.
- [18] N. Andalibi, P. Ozturk, and A. Forte, "Sensitive self-disclosures, responses, and social support on instagram: The case of# depression," *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pp.1485-1500, 2017.

- [19] M. Paul and M. Dredze, "You are what you tweet: Analyzing twitter for public health," *Proceedings of the International AAAI Conference on Web and Social Media*, Vol.5, No.1, 2011.
- [20] M. W. Newman, D. Lauterbach, S. A., Munson, P. Resnick, and M. E. Morris, "It's not that I don't have problems, I'm just not putting them on Facebook: Challenges and opportunities in using online social networks for health," *In Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work*, pp.341-350, 2011.
- [21] A. Kramer and C. Chung, "Dimensions of self-expression in Facebook status updates," *Proceedings of the International AAAI Conference on Web and Social Media*, Vol.5, No.1, pp.169-176, 2011.
- [22] M. De Choudhury, and S. De, "Mental health discourse on reddit: Self-disclosure, social support, and anonymity," *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.
- [23] R. Skaik and D. Inkpen, "Using social media for mental health surveillance: A review," *ACM Computing Surveys (CSUR)*, Vol.53, No.6, pp.1-31, 2020.
- [24] G. Dosovitsky, E. Kim, and E. L. Bunge, "Psychometric properties of a chatbot version of the PHQ-9 with adults and older adults," *Frontiers in Digital Health*, Vol.3, pp.41, 2021.
- [25] M. Conway, "Ethical issues in using Twitter for public health surveillance and research: Developing a taxonomy of ethical concepts from the research literature," *Journal of Medical Internet Research*, Vol.16, No.12, pp.e290, 2014.
- [26] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, Vol.3, No.1, pp.993-1022, 2003.
- [27] Y. R. Suh, K. P. Koh, and J. Lee, "An analysis of the change in media's reports and attitudes about face masks during the COVID-19 pandemic in South Korea: A study using Big Data latent dirichlet allocation (LDA) topic modelling," *Journal of the Korea Institute of Information and Communication Engineering*, Vol.25, No.5, pp.731-740, 2021.
- [28] E. L. Park and S. Cho, "KoNLPy: Korean natural language processing in Python," *Annual Conference on Human and Language Technology*, Human and Language Technology, pp.133-136, 2014.
- [29] A. Shrestha and F. Spezzano, "Detecting depressed users in online forums," *In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, Vancouver, Canada, pp.945-951, 2019.
- [30] M. De Choudhury, S. Counts, E. J. Horvitz, and A. Hoff, "Characterizing and predicting postpartum depression from shared facebook data," *In Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing (ACM CSCW)*, Baltimore, Maryland, USA, pp.626-638, 2014.
- [31] H. J. Jeon et al., "Melancholic features and hostility are associated with suicidality risk in Asian patients with major depressive disorder," *Journal of Affective Disorders*, Vol.148, No.2-3, pp.368-374, 2013.
- [32] S. Park et al., "The association of suicide risk with negative life events and social support according to gender in Asian patients with major depressive disorder," *Psychiatry Research*, Vol.228, No.3, pp.277-282, 2015.
- [33] X. Ma, J. Hancock, and M. Naama, "Anonymity, intimacy and self-disclosure in social media," *In Proceedings of the CHI Conference on Human Factors in Computing Systems (ACM CHI)*, San Jose, CA, USA, pp.3857-3869, 2016.
- [34] R. Zhang, "The stress-buffering effect of self-disclosure on Facebook: An examination of stressful life events, social support, and mental health among college students," *Computers in Human Behavior*, Vol.75, pp.527-537, 2017.
- [35] H. J. Jeon et al., "Differences in depressive symptoms between Korean and American outpatients with major depressive disorder," *International Clinical Psychopharmacology*, Vol.29, No.3, pp.150-156, 2014.
- [36] H. Chin and M. Y. Yi, "Voices that care differently: Understanding the effectiveness of a conversational agent with an alternative empathy orientation and emotional expressivity in mitigating verbal abuse," *International Journal of Human-Computer Interaction*, Vol.38, No.12, pp.1153-1167, 2022.



진 효 진

<https://orcid.org/0000-0003-4773-9518>

e-mail : tesschin@ibs.re.kr

2014년 KAIST 정보미디어 MBA

2020년 KAIST 산업및시스템공학과(박사)

2020년~ 현재 기초과학연구원

데이터사이언스그룹 선임연구원

관심분야 : Human-AI Interaction, Data Science, Natural Language Processing



정 찬 이

<https://orcid.org/0000-0002-3442-2434>
e-mail : 1016chani@gmail.com
2017년~현 재 KAIST 전산학부 학사과정
관심분야 : Natural Language
Processing, Computational
Social Science



최 정 회

<https://orcid.org/0000-0002-8540-5191>
e-mail : sijay@simsimi.com
2007년 서울대학교 산업디자인학과(학사)
2002년~현 재 심심이 주식회사 대표
관심분야 : Chatbot, Natural Language
Processing, AI Ethics



백 금 희

<https://orcid.org/0000-0003-1999-0158>
e-mail : dnjsxka486@naver.com
2022년 이화여자대학교 간호학과(석사)
2022년~현 재 이화여자대학교 간호대학
박사과정
관심분야 : Mobile Health, Mental Care,
Symptom Management of
Post-Covid-19 Syndrome



차 미 영

<https://orcid.org/0000-0003-4085-9648>
e-mail : meeyoungcha@kaist.ac.kr
2008년 KAIST 전산학과(박사)
2008년~2010년 독일 막스플랑크연구소
박사후연구원
2010년~현 재 KAIST 전산학부 전임교수
2015년~2016년 미국 페이스북 초빙교수
2019년~현 재 기초과학연구원 Chief Investigator
관심분야 : Data Science, Computational Modeling, AI



차 지 영

<https://orcid.org/0000-0003-0115-1348>
e-mail : chiyoung@ewha.ac.kr
2010년 University of Washington(박사)
2011년 이화여자대학교 간호대학
박사후연구원
2012년~2017년 이화여자대학교
간호대학 조교수

2018년~현 재 이화여자대학교 간호대학 부교수
2019년 University of Washington 초빙교수
관심분야 : Mobile Health, Symptom Clusters, Depression