

상담 및 심리치료에서 인공지능 기술의 활용: 국외사례를 중심으로*

김도연 조민기 신희천[†]

아주대학교 심리학과 / 학생

아주대학교 심리학과 / 교수

본 연구는 상담 및 심리치료 분야에서 컴퓨터와 정보통신 기술을 활용해왔던 기존의 시도들을 살펴보고, 여섯 가지 최근 국외사례를 중심으로 인공지능 기술을 활용한 상담의 현황을 소개하면서, 앞으로의 발전 가능성을 기술적 이슈와 전망을 통해 살펴보았다. 기존의 웹이나 모바일 기반 온라인 심리치료들이 정보통신(IT) 기술의 발전을 바탕으로 서비스를 제공하여 접근성과 효율성을 높임으로써 치료적 개입의 효과를 높이는데 초점이 있었다면, 최근에는 챗봇이나 대화형 에이전트와 같이 내담자와 치료적 상호작용을 할 수 있는 형태로 발전하면서 그 효과를 검증하는 연구들이 활발히 이루어지고 있다. 최근 국외에서 개발된 인공지능 기반 심리치료의 대표적인 사례들(웹기반 심리치료 MOST, 모바일 기반 TESS, 대화형 에이전트 WOEBOT, 대화형 인공지능 내담자 CLIENTBOT, 가상현실 기반 임상적 면담자 ELLIE)을 통해 인공지능 상담의 현재를 살펴보았다. 끝으로 인공지능 상담의 미래 전망과 관련된 기술적 이슈로 자연어 처리, 대화형 에이전트가 공감적인 표현을 할 수 있게 하는 메커니즘, 그리고 자기학습과정을 통해 인간의 도움 없이도 지속적인 발전을 시도하는 인공지능의 개발 등을 살펴보고, 본 연구의 의의 및 제한점을 언급하였다.

주요어 : 인공지능, 온라인 심리치료, 자연어 처리, 대화형 에이전트

* 본 연구는 '2018학년도 아주대학교 일반연구비 지원'에 의하여 연구되었음.

[†] 교신저자 : 신희천, 아주대학교 심리학과, 경기도 수원시 영통구 월드컵로 206 아주대학교 율곡관 506호

Tel : 031-219-2747, E-mail : hcshin@ajou.ac.kr

인공지능(AI: Artificial Intelligence)은 ‘인간의 학습 능력과 추론 능력, 지각 능력, 자연언어의 이해 능력 등과 같은 지능을 컴퓨터 프로그램으로 실현한 기술’이라고 정의할 수 있다. 즉, 컴퓨터와 같은 기계가 인간의 지능으로 할 수 있는 능력들을 모방하는 것을 말한다. 인공지능(Artificial Intelligence)이란 용어는 1956년 미국 다트머스 대학에서 열린 컨퍼런스에서 컴퓨터 과학자 John McCarthy가 처음 공식적으로 사용하였다(McCarthy, 1956; McCarthy, Minsky, Rochester, & Shannon, 2006에서 재인용). McCarthy는 인공지능을 ‘인텔리전트한 기계를 만드는 과학과 공학’이라고 정의했으며, 기계가 지식을 가지고 스스로 학습하고 행동하는 시대가 올 것이라 주장했다. 그로부터 약 60년 후인 2016년 3월 언론과 대중의 큰 관심 속에 Google Deepmind사의 인공지능 ‘알파고’가 이세돌 9단과의 바둑 경기 승부를 펼치게 되었다. 이 대국은 인공지능의 승리로 끝이 났는데, 기계가 인간을 넘을 수 없는 것으로 여겨졌던 분야 중 하나인 바둑에서 인공지능 기술의 일종인 ‘딥 러닝(Deep Learning)’을 통해 인간을 넘어섰다는 점에서 큰 의미가 있는 것으로 받아들여진다(박태진, 2017; Silver et al., 2016). 이 사건을 통해 한국뿐만 아니라 전 세계적으로 인공지능 및 관련 기술에 대한 관심이 매우 커졌다. 이러한 관심은 심리학 분야에서도 마찬가지로, 다양한 측면에서 심리학과 인공지능 기술을 접목하고 활용하고자 하는 움직임이 나타나고 있다. 대표적인 예로 미국심리학회(APA)는 2019년 10월 ‘TECHNOLOGY, MIND & SOCIETY’ 컨퍼런스를 개최하여 인공지능과 같은 신기술의 혁신적 설계와 인공지능 기술의 활용에 있어 심리학의 역할에 대해 논의한 바 있다. 국내 심리

학계에서도 이런 흐름은 나타나고 있는데, 지난 2018년에 개최된 제72차 한국심리학회 연차학술대회에서는 ‘심리학, 인간과 기계의 마음을 말하다’라는 주제로 미래사회의 기술변화에 대해 심리학이 적극적으로 대응해야 할 필요가 있다는 점, 그러한 맥락에서 심리학의 다양한 분야에서 인공지능 기술을 활용할 필요성과 가능성에 대한 논의가 이루어진 바 있다.

사실 상담 및 심리치료 분야에서 보다 효율적이고 효과적인 치료적 개입을 위해 새로운 기술을 활용하려는 시도는 처음이 아니며, 인공지능 이전에도 컴퓨터 및 정보통신(IT) 등과 같은 기술들을 활용하고자 하는 노력들이 지속적으로 있어왔다. 상담 및 심리치료 분야에서 컴퓨터를 활용하고자 했던 최초의 시도는 1966년에 개발된 ELIZA 프로그램이었다(Weizenbaum & Freeman, 1976; Luxton, 2014에서 재인용). ELIZA는 인본주의 심리학자인 Carl Rogers의 공감적인 의사소통을 모방하도록 설계되었으며, 사용자가 키보드를 통해 입력한 문장을 분석하여 일정한 규칙에 따라 대답을 만들어내어 출력하는 형태로 작동했다. 당시에는 인간의 질문을 이해하고 대화를 이끌어 나갈 수 있을만한 프로그램을 개발할 기술이 부족했기 때문에 사전에 약속된 규칙에 따른 정형화된 반응을 하면서 대화를 모방하는 수준에 그쳤다. 이후 1970년대 초반에는 스탠포드 대학의 정신과 의사인 Colby(1977)가 편집증적인 조현병 환자를 시뮬레이션하여 타인과 대화할 수 있는 PARRY라는 프로그램을 개발한 바 있다(Franchi & Guzeldere, 1995). 이 프로그램은 튜링 테스트를 통과한 최초의 컴퓨터 프로그램이었는데, 숙련된 정신의학자조차 PARRY가 컴퓨터 프로그램인지 조현병을 가진

진짜 인간인지 구분하지 못한 것으로 알려져 있다(Teuscher & Hofstadter, 2006).

이후 정보통신 분야의 기술 발전이 두드러짐에 따라 인터넷에 대한 접근성과 활용성이 높아졌고 사람들은 다양한 분야에서 인터넷을 활용하기 시작했다. 이러한 움직임은 상담 및 심리치료 분야에서도 나타나 이 시기부터 인터넷을 기반으로 하는 온라인 심리치료가 본격적으로 개발되고 활용되기 시작했으며, 국내에서도 비슷한 시기에 사이버 상담이라는 명칭으로 프로그램들이 개발되고 그 효과가 검증되기 시작했다(이재영, 정영운, 2001, 심혜숙, 이현진, 2003). 온라인 심리치료는 치료적 개입이 인터넷을 통해 제공되는 형태의 심리치료를 의미한다(Childress, 2000). 초창기 컴퓨터의 발달과 함께 의학적 정보를 제공하기 위한 목적으로 웹사이트를 활용한 것을 시초로 보고 있으며(Kolata, 2000), 이후에는 단순한 정보 제공을 넘어 환자의 신체적, 정신적 증상을 줄이기 위한 목적으로 활용되고 있다(Ritterband et al., 2003).

온라인 심리치료 중 현재 가장 보편적으로 활용되고 있는 방식은 웹사이트 기반 심리치료이며, 세부적으로는 다양한 형태가 존재하지만 대표적으로는 웹사이트를 플랫폼으로 하여 심리장애와 관련된 유용한 정보들을 제공하고, 상담자와의 채팅을 통해 상담을 진행하며, 회기 사이에 주어진 과제를 수행하는 형태로 운용된다. 이런 웹사이트 기반 방식은 사람들이 인터넷을 통해서 얻은 정보들을 자신의 건강관리 행동에 반영한다는 사실을 기반으로 한다. 웹사이트 기반 심리치료는 다양한 장점을 가지고 있는데 무엇보다 웹사이트라는 하나의 플랫폼을 통해 동시에 수많은 사람들에게 동일한 정보를 적은 비용으로 제공

할 수 있다. 또한 웹사이트 자체는 24시간 운용될 수 있기 때문에 내담자의 관점에서 보면 시간적 제약이 없고 물리적인 거리 또한 의미가 없다(이동훈, 김주연, 김진주, 2015). 이는 비용 대비 효율성과 치료에 대한 접근성을 높이는 것뿐만 아니라 사회적 편견 혹은 낙인으로 인해 가질 수 있는 심리치료에 대한 접근장벽(Corrigan & Rüsch, 2002)을 낮추는 것과 같은 장점으로 이어진다. 실제로 미국 재향군인회에서는 퇴역군인들의 외상 후 스트레스 장애(Post Traumatic Disorder; PTSD)를 치료할 목적으로 웹사이트를 통해 운용되는 자기관리형 인지행동치료 프로그램을 개발하여 운용하였으며, 그 결과 PTSD뿐만 아니라 우울 및 불안 증상이 유의미하게 개선되는 것으로 보고한 바 있다(Litz, Engel, Bryant, & Papa, 2007). 그 뿐만 아니라 이명으로 인한 심리적 고통의 완화(Andersson, 2002), 섭식장애(Gulec et al., 2011, Gollings & Paxton, 2006), 불안장애(Spence, Holmes, March, & Lipp, 2006; Cohen & Kerr, 1999), 신체 이미지(Celio et al., 2000), PTSD와 병적인 슬픔(Lange, van de Ven, Schrieken, & Emmelkamp, 2001)에 대한 웹사이트 기반 심리치료가 진행되어 그 효과가 검증되고 있는 중이다.

온라인 심리치료의 또 다른 형태는 화상상담이다. 화상상담은 치료자와 내담자가 물리적으로 떨어져 있는 상태에서 광대역 통신망이나 무선 통신망을 통해 실시간으로 전달되는 영상과 음성을 기반으로 상담을 진행하는 방식으로, 화상상담이 가능하기 위해서는 광대역 통신망 연결 회선, 컴퓨터, 웹캠, 스마트폰 등이 필요하다. 화상상담의 경우 웹사이트를 이용하는 방식에 비해 전통적인 대면 상담과 유사한 측면이 있다. 대면 상담과 같이 내

담자의 상태에 대한 청각적, 시각적 정보를 얻을 수 있다는 장점이 있지만, 웹캠을 기반으로 하기 때문에 동공의 확대나 축소, 미묘한 움직임, 분위기의 변화와 같은 미묘한 정보를 전달하는 데는 한계가 있는 것으로 알려져 있다(Yuen, Goetter, Herbert, & Forman, 2012). 또한 인터넷 회선이 느리면 영상과 음성 전송이 늦어지거나 아예 끊어져 상담에 방해가 될 뿐만 아니라 아예 불가능한 상황이 있을 수 있다. 화상상담을 통한 심리적 개입은 외상 후 스트레스 장애(Frueh et al., 2007; Germain, Marchand, Bouchard, Guay, & Drouin, 2010; Tuerk, Yoder, Ruggiero, Gros, & Acierno, 2010), 사회불안(Yuen et al., 2010), 공황장애(Bouchard et al., 2004), 강박장애(Himle et al., 2006), 청소년과 아동의 우울증(Pesämaa et al., 2004) 등 다양한 심리장애에 적용되고 있으며 대면 치료만큼이나 효과가 있음이 검증되고 있는 중이다.

최근에는 IT 및 무선 인터넷 기술의 발전과 이를 활용하는 스마트폰의 보급으로 스마트폰을 활용한 디지털 심리치료에 대한 관심 또한 높아지고 있다. 스마트폰을 기반으로 하는 심리적 개입의 경우 현재는 챗봇(Chatbot)과 같은 텍스트 메시징 서비스를 통해 상호작용하면서 사전에 입력된 구조화된 개입 절차를 제공하는 방식이 많이 개발 및 연구되고 있다(Woebot: Fitzpatrick, Darcy, & Vierhile, 2017; Shimbobot: Ly, Ly, & Andersson, 2017). 사실 이런 애플리케이션 형태의 심리치료 프로그램들은 인터넷 연결만 가능하다면 개인용 컴퓨터를 통해서도 제공될 수 있다. 하지만 스마트폰을 활용하는 방식의 경우 휴대용 기기의 장점인 높은 휴대성을 바탕으로 실제 생활 장면 속에서 적절한 시점에 치료적 개입을 진행할 수

있고, 즉각적인 피드백을 통해 개입의 효과를 높일 수도 있다. 또한 스마트폰뿐만 아니라 스마트워치와 같은 웨어러블 기기(wearable device)를 통해 실생활로부터 내담자에 관한 다양한 정보를 수집하고 상담자와 공유함으로써 치료적 개입을 위한 근거 자료로 활용할 수 있고(Riley et al., 2011), 수시로 변하는 내담자의 상태를 모니터링하고 관리하는 데도 도움을 받을 수 있다. 실제로 Dagöö 등(2014)은 사회불안장애를 가진 내담자들을 대상으로 스마트폰 기반 인지행동치료를 진행하여 그 증상이 유의미하게 감소하는 것을 확인하였으며, 대학생들에게 인스턴트 메신저 애플리케이션 형태로 제작된 텍스트 기반 대화형 에이전트를 통해 인지행동치료를 제공한 결과 우울 증상이 감소한 것으로 나타난 연구결과 또한 발표된 바 있다(Fitzpatrick et al., 2017).

정보통신 기술과 더불어 컴퓨터 그래픽 기술의 발전은 가상현실(Virtual Reality; VR)을 활용한 심리치료의 개발 및 보급을 촉진하고 있다. 가상현실을 임상적 장면에서 활용하는 주된 방식은 환자의 치료에 이러한 기술을 사용하는 것으로, Rothbaum, Hodges, Ready, Graap, & Alacon(2001)에 의해 베트남전 참전 군인들의 PTSD 치료에 처음으로 소개된 바 있다. VR을 이용한 심리적 개입은 가상의 상황을 매우 생생하게 경험해주고, 불안을 유발하는 상황의 강도를 자유롭게 조정할 수 있다는 측면에서 PTSD, 사회불안, 특정 공포증과 같은 불안 관련 장애에 대한 노출 치료에 효과적인 것으로 알려져 있다. 노출 치료에서는 스트레스 자극에 대한 노출을 통해 기억 및 정서를 활성화시키고 이를 재구성하는 과정을 거치기 때문에 성공적인 치료를 위해서는 기억의 생생한 재생이 필요하다(Foa & Kozak, 1986). 이

때 가상현실 기술을 활용하면 심상에 의존하는 기존 기법에 비해 공포 혹은 불안 반응을 유발하는 대상에 대한 생생하고 반복적인 노출을 실시할 수 있고 이는 노출 치료의 효과성과 효율성을 높이는 결과로 이어질 수 있다. 임상적 가상현실을 이용한 치료는 PTSD(Rothbaum et al., 1999, 2001; Difede & Hoffman, 2002; Josman et al., 2006; Beck, Palyo, Winer, Schwagler, & Ang, 2007), 비행공포증(Rothbaum, Hodge, Smith, Lee, & Price, 2000), 고기능 자폐군 환자들의 사회적 기술 훈련(Parsons, Leonard, & Mitchell, 2006, Parsons & Cobb, 2014, Parsons, 2016). 사회공포증(Newman, Szkodny, Llera, & Przeworski, 2011) 등에 효과적인 것으로 밝혀져 있다. 특히 Yuen 등(2013)이 성인을 대상으로 실시한 사회불안장애 가상현실 노출 치료 프로그램에서는 온라인 가상현실 속에서 내담자들이 아바타를 만들고 다른 참가자들과 상호작용함으로써 불안을 유발하는 자극을 재경험하는 것에 그치지 않고 실제 상호작용 상황에서 불안을 경험하고 상담자와 함께 이를 다루는 경험을 제공하는 데까지 발전하였다.

본 논문은 위에서 살펴본 사례들 이후에 진행되어져 온 기술의 적용사례들, 특히 인공지능 기술을 상담 및 심리치료 분야에 적용한 최신 사례들을 소개하고, 기술적 측면에서 인공지능 상담의 발전 가능성에 관해 논의하는 것을 목적으로 한다. 앞서 알파고의 예에서 언급한 바와 같이 인공지능 관련 기술들은 최근 비약적으로 발전하여 사람들의 삶 속에 본격적으로 적용되고 있다. 이는 복지 관련 분야에서도 예외는 아니어서, 의료 분야에서는 구글이나 IBM과 같은 거대 IT기업을 중심으로 이미 인공지능을 활용한 진단시스템을 개발

및 도입하고 있는 실정이다. 정신건강 증진과 밀접한 관련을 맺고 있는 상담 및 심리치료 분야의 경우 내담자를 대상으로 하는 개입에 있어 인공지능 관련 기술의 도입이 시작되고 있는 단계이긴 하지만, 관련 서비스를 개발하고 운영하는 과정에서 이러한 기술들을 도입함으로써 전통적인 개입 방식이 가질 수 있는 단점들을 보완하기 위한 노력이 이루어지고 있다. Luxton(2014)은 정신건강 분야에서 인공지능 및 관련 기술을 활용하는 것이 가질 수 있는 이점에 대해 다음과 같이 소개한다. 먼저 온라인 심리치료와 같이 기존의 IT기술을 활용한 개입이 가지는 높은 접근성이라는 장점에 더해, 내담자가 경험하는 개입과정이 더욱 상호작용적이고 관계를 구축한다는 느낌이 들게 한다. 이는 온라인 심리치료가 가지는 약점 중의 하나인 높은 중도탈락률(Aguilera, 2015)을 보완할 수 있는 해결책이 될 수 있다. 또한 정신건강 서비스의 잠재적 이용자들이 우려하는 낙인을 걱정하지 않아도 된다. 특히 한국 문화권에서는 상담을 받는다는 사실에 대한 수치심이 상담장면을 찾는 데 큰 장애물이 된다는 점(이동훈 등, 2015)에서 장점을 가지게 된다. 실제로 특정 주제를 가진 내담자들의 경우 인간 전문가보다 가상의 치료자를 만날 때 덜 불안해하고 이를 더 선호할 가능성을 지지하는 증거들이 발견되고 있다(Gratch, Wang, Gerten, Fast, & Duffy, 2007; Kandalaf, Didehbani, Krawczyk, Allen, & Chapman, 2013; Lucas, Gratch, Kiing, & Morency, 2014). 이는 자신의 이야기를 듣는 상대방이 인간이 아닌 기계이기 때문에 자신을 판단하지 않고 기다려 줄 것이라는 믿음을 가지기 때문일 수 있으며(Bickmore et al., 2010), 그런 환경에서 덜 방어하고 더 개방적인 경향이 있음을 발견한 연구

들이 존재한다(Hart, Gratch, & Marsella, 2013). 그뿐만 아니라, 인공지능 관련 기술을 활용할 경우 인간 전문가의 개입을 더욱 효과적으로 촉진하고 이를 보완할 수 있는 다양한 방법들을 활용할 수 있다. 예를 들어, 인간 상담자와의 회기 사이에서 피드백을 통해 행동 수정을 돕거나, 응급상황에서 적절한 조치를 제공하고, 내담자를 객관적으로 평가하는데 도움이 될 수 있는 데이터들을 수집 및 분석할 수 있다. 특히 대규모 데이터 속에서 규칙을 찾아내는데 특화되어 있는 인공지능 관련 기술들을 활용할 경우 심리평가 및 진단 분야에서 인간 전문가가 가질 수밖에 없는 인지 및 판단능력의 한계를 뛰어넘어 훨씬 짧은 시간에 더 정확한 판단을 내릴 수 있도록 돕고, 인지적 오류 및 실수의 가능성을 획기적으로 낮출 수 있을 것이다.

앞으로 인공지능 관련 기술이 급속도로 발전하여 인간 삶의 다양한 영역에 영향을 미칠 것이고, 이러한 흐름을 막을 수는 없음을 자명해 보인다. 정신건강 분야 역시 예외는 아닐 것이며 이 분야의 전문가들이 개입을 하던 그렇지 않던 기술은 정신건강 개입에 큰 영향을 미칠 것이다. 때문에 정신건강 분야의 전문가들이 이러한 흐름을 주도하지 못한다면 인간이 아닌 기술이 중심이 되는 흐름이 이어질 수 있다. 즉, 인간의 행복과 복지를 위해서 분명한 목적을 가지고 기술을 활용하는 것이 아니라, 기술적으로 가능하기 때문에 이를 정신건강 개입에 적용하고자 하는 움직임이 나타날 수 있는 것이다. 이런 움직임에 대한 우려는 이미 시작되어, 2019년 3월에는 국제전기전자공학회(IEEE) 표준화 기구는 인공지능의 설계 및 개발, 활용 과정에서 인권, 복지, 투명성, 책임성의 중요성을 강조하는 자율·지

능 시스템 원칙인 ‘윤리적으로 정렬된 설계’를 발표하였고, 11월에는 유럽의회와 유럽위원회의 지원을 받는 ‘사람을 위한 인공지능(AI4people)’ 포럼에서 기술과 윤리, 시장, 사회규범 간의 균형을 강조하는 보고서를 발표한 바 있다. 그뿐만 아니라 경제협력개발기구(OECD) 및 주요 20개국(G20) 무역 및 디지털 경제 분야 장관 회의에서도 인공지능의 개발 및 활용에서 사람 중심 원칙이 핵심임을 강조하고 있다. 이러한 움직임들이 공통적으로 강조하는 바와 같이 인공지능 기술의 개발 및 활용이 사람 중심 원칙에 따라 이루어지기 위해서는 인공지능 기술의 전문가들이 이러한 원칙을 잘 이해해야 하는 것뿐만 아니라 기술이 실제로 적용되는 분야의 전문가들이 이 기술을 충분히 이해하고, 그 위험성과 가능성을 고려하여 활용 방향을 결정하는 것이 매우 중요할 것이다. 따라서 이러한 움직임을 주도하기 위한 노력의 일환으로 인공지능 기술이 상담 및 심리치료 분야에서 어떤 식으로 활용되고 있으며, 앞으로 어떻게 활용될 수 있는지를 살펴보는 것은 충분한 의미가 있을 것으로 생각된다. 이를 위해 본 연구에서는 인공지능 이전의 관련 기술들이 상담 및 심리치료 분야에 적용되어 온 역사를 검토하고, 인공지능 및 관련 기술을 적용한 치료적 개입의 개발이 어디까지 이루어지고 있는지 구체적인 사례를 중심으로 살펴보고자 한다.

방 법

본 연구를 진행하기 위해 인공지능 및 관련 분야의 기술을 상담 및 심리치료 분야에 적용한 국내외 연구물들을 조사했다. 이를 위해

먼저 최근 3년(2017년~2019년)을 기준으로 학술 데이터베이스에 등재된 논문 및 관련 자료들 중 인공지능 기술이 상담 및 심리치료에 실제 적용된 사례들을 살펴보았다. 자료 검색에 사용한 학술 데이터베이스는 PubMed, Scopus, JMIR(Journal of Medical Internet Research), Frontiers in Psychology, PsycARTICLES, Google Scholar였으며, 검색에 사용한 주제어(keyword)는 “artificial intelligence”, “digital psychotherapy”, “health technology”, “online counseling”, “chatbot”, “conversational agent”였다. 최초 검색 결과 290건의 연구물들이 검색되었고, 논문의 주제, 초록, 논의를 검토하여 42개의 논문 및 자료를 예비 검토 대상으로 선정하였다. 이후 자료들을 재검토하여 다섯 가지의 주요 적용 사례를 선정하였는데, 주요 사례를 선정한 기준은 다음과 같았다. 정신건강, 상담 및 심리치료의 맥락에서 이루어진 연구 중 심리적 개입을 목적으로 인공지능 관련 기술을 활용하여 개발해낸 결과물의 개발과정 및 적용방식, 효과성에 대한 구체적인 정보가 제시된 연구들을 포함시켰으며, 인공지능 기술을 적용한 방식이 동일하여 기술적 영역이 중복되는 연구, 가상상담, 웹사이트 기반 심리치료, 모바일 애플리케이션, 가상현실 등과 같이 단순히 특정 IT기술만을 적용한 연구, 심리교육적 정보의 제공이나 정형화된 대화형 도구와 같이 인간과 기계의 상호작용이 아닌 단방향적 개입을 활용한 연구는 분석에서 제외하였다. 국내 연구의 경우 DBPIA, e-Article, KISS, RISS, 국회도서관 학술 데이터베이스에서 “챗봇”, “상담 챗봇”, “인공지능 상담”, “대화형 에이전트”를 주제로 검색하였으나 그 사례가 많지 않았고, 검색된 연구물들조차도 인공지능 기술의 적용이 상담 및 심리치료 분야의 치료적 개입 맥락에

서 이루어지지 않았거나, 개발배경 및 과정, 실제 치료적 맥락에서의 적용 방식, 치료적 개입의 효과성이 구체적으로 제시되지 않았거나 혹은 국외 사례들과 비교하여 미흡한 경우가 많았다. 이에 국내 연구물들은 본 연구의 연구범위에 포함시키지 않았다. 주요 사례들에 대한 분석은 각 사례의 개발배경, 작동원리(mechanism), 효과성, 그리고 장단점을 포함한 상담 및 심리치료 분야에서의 함의를 중심으로 이루어졌으며, 분석결과는 독자들의 가독성 및 이해가능성을 높이기 위해 재구조화하여 제시하였다.

결 과

ELIZA로부터 시작되어 온라인 심리치료로 이어진 흐름은 현재 쌍방향적이고 자연스러운 대화를 수행할 수 있는 가상현실 대화형 에이전트의 개발을 시도하는 데까지 이르렀다. 가상현실 시뮬레이션, 기계 학습, 자연어 처리 기술, 추론 능력을 갖춘 지식 기반 인공지능 기술의 발전이 이를 가능하게 하고 있다. 여기서는 발전된 인공지능 및 관련 기술들을 심리치료에 활용하고자 시도하고 있는 다섯 가지 최신 사례들(웹 기반 심리치료 MOST, 모바일 기반 TESS, 대화형 에이전트 WOEBOT, 대화형 인공지능 내담자 CLIENTBOT, 가상현실 기반 임상적 대화형 에이전트 ELLIE)을 각각의 개발배경, 작동원리(mechanism), 효과성, 장점과 한계를 포함한 상담 및 심리치료 분야에서의 함의를 중심으로 소개하고자 한다.

웹 기반 심리치료 MOST

MOST 프로젝트(Álvarez-Jiménez & Gleeson, 2012; Lederman, Wadley, Gleeson, Bendall, & Álvarez-Jiménez, 2014; D'Alfonso et al., 2017)는 웹사이트 상에서 페이스북과 유사한 형태의 대화형 소셜 미디어 플랫폼의 형태로 서비스되는 온라인 사회치료(Social Therapy) 프로그램이다. 웹사이트를 기반으로 하는 인터넷 심리치료는 정보통신 기술을 활용한 상담 및 심리치료 개입의 전통적인 형태 중 하나로, 여전히 다양한 방식으로 개발되고 있으며, 그 효과성 또한 지속적으로 검증되고 있는 중이다. 인터넷 기반 심리치료는 낙인의 감소(Houston, Cooper, & Ford, 2002), 자기 공개의 증가(Weisband & Kiesler, 1996), 고립의 감소(Dennis, 2003), 그리고 지지적 관계의 발달(O'Keeffe & Clarke-Pearson, 2011)에 이르기까지 다양한 장점을 가진 것으로 평가되고 있다. 이런 맥락 속에서 MOST 프로젝트의 주된 목적은 정신질환을 가지고 있는 청년들을 지원하는 것으로, 정신질환과 관련된 사회적 낙인으로 인해 고립을 경험하고 관계를 형성·유지하는 것이 어려운 청년들(Morgan et al., 2012)을 돕기 위한 목적으로 개발되었다.

MOST는 소셜 미디어 플랫폼을 활용함으로써 사용자들의 관심을 유도하고 참여를 극대화하고자 한다. 이 개입방식은 긍정심리치료의 강점 기반 모델(Seligman, Rashid, & Parks, 2006)을 이론적 바탕으로 하며, 총 4가지 개입 영역으로 구성되어 있다. 첫째, 사용자가 페이스북 스타일의 뉴스피드를 게시하고 코멘트를 달면서 자신의 경험을 공유하고 지원을 주고받으면서 관점을 검증해볼 수 있는 소셜 네트워킹 및 카페(social networking & cafe) 영역. 둘

째, 제공되는 심리적 기술들을 연습하고 개발하도록 설계된 상호작용 치료 모듈인 Take a step. 셋째, 증거 기반 사회 문제 해결 프레임워크(evidence-based social problem solving framework; McFarlane, 2004)에 따라 사용자의 문제나 어려움을 꺼내놓고 증재자 집단 속에서 토론하면서 도움을 얻을 수 있는 공간인 Talk it out 영역. 넷째, 마음챙김, 자기 연민, 개인적 강점 활용 등의 제안을 현실 상황에 적용할 수 있도록 도와주는 행동실험이나 행동화 프로세스를 제공하는 Do it! 등 총 4가지 영역으로 구성된다. 본 시스템은 여기에 그치지 않고 인공지능 기술의 적용을 통해 서비스 제공의 효율성과 효과성을 높이는 방향으로 개선되고 있다는 점에서 더 큰 가능성을 가진다. 먼저, 시스템 내에서 사용자들의 호소문제와 배경 정보 등을 검토하고 개별화된 적절한 치료 콘텐츠를 제안하는 역할을 하는 증재팀(7명의 임상심리사와 임상사회복지사로 구성)을 대체하기 위한 목적으로 사용자의 게시물에 대한 언어분석을 실시하고, 게시물로부터 필요한 정보를 추출하여, 적절하다고 판단되는 치료적 제안을 선택하여 제시하는 자동화된 프로세스를 개발하고 있다. 이러한 프로세스의 개발은 내담자의 언어적/심리 측정적 특성이 치료 내용과 결합될 수 있다는 연구결과에 근거하며(Tausczik & Pennebaker, 2010), 이를 통해 개입의 즉시성을 높일 수 있을 것으로 보고 있다. 또 다른 개선점은 시스템 내 인공지능을 기반으로 하는 챗봇(Chatbot)의 도입이다. 챗봇은 문자 또는 음성 기반 인터페이스를 통해 인간의 대화를 모방하는 컴퓨터 프로그램이다(Abdul-Kader & Woods, 2015). 챗봇은 시스템 내에서 사용자가 대화를 통해 정보를 검색할 수 있는 보조 인터페이스의 형태로 제

공될 계획이지만, 단순한 보조 인터페이스로만 기능하는 것이 아니라 사용자와의 대화를 통해 정보를 수집하고, 이를 바탕으로 적절한 개입방안을 검색하여 인간 치료사들에게 제안함으로써 인간 치료사와 사용자 간의 온라인 채팅 세션의 효과성을 높이는 방식으로 활용될 예정이다.

MOST는 기본적으로 대화형 소셜 미디어 플랫폼을 통해 제공되는 서비스이기 때문에 사용자들이 시스템을 사용하고 심리적 도움을 얻는 과정에서 상호작용 경험이 자주 일어나게 된다. 거기에 더해 인공지능 중재 시스템을 통한 개별화된 개입의 즉각적인 제공과 챗봇 시스템과의 상호작용 경험은 일방향성으로 인해 높은 중도탈락률이 문제가 되는 기존의 인터넷 심리치료의 한계를 극복하는데 효과적인 방법이 될 수 있다. 즉, 사용자들은 낙인으로부터 비교적 자유로운 온라인상에서 서로 소통하며 긍정적 사회화, 지지적 관계의 형성, 자존감 향상, 소통의 촉진, 집단 구성원으로서의 소속감 경험과 같은 심리적인 이득을 얻을 수 있다(Collin, Rahilly, Richardson, & Third, 2011; O’Keeffe & Clarke-Pearson, 2011).

현재 MOST는 4가지 파일럿 연구를 포함하여 총 6가지 연구를 통해 그 효과성이 검증되고 있으며, 이러한 방식의 개입이 심리치료의 긍정적 효과를 장기적 개선으로 확장시킬 수 있다는 점을 보여준 바 있다(D’Alfonso et al., 2017). 전통적인 면대면 심리치료를 통해 정신증이나 우울증 증상의 완화를 경험한 경우라도 증상이 재발하는 비율이 청년층의 경우 최대 80%에 이르는 등(Alvarez-Jimenez et al., 2012) 재발을 방지하는 것이 상담 및 심리치료의 중요한 목표 중 하나인 상황에서, MOST와 같은 웹사이트 기반 인터넷 심리치료가 증

상의 장기적인 개선과 재발 방지로 이어질 수 있다는 점은 상담 및 심리치료의 적용에 많은 시사점을 가진다.

모바일 기반 정신건강 개입 TESS

TESS(Fulmer, Joerin, Gentile, Lakerink, & Rauws, 2018)는 정신건강 지원 및 심리교육을 목적으로 X2AI Inc에 의해서 개발된 심리학적인 인공지능 챗봇(Psychological AI chatbot)으로, 웹사이트뿐만 아니라 주로 모바일 기기를 통해 애플리케이션 형태로 서비스되며 사용자들의 정서표현과 자신의 심리적 상태에 대한 걱정들을 기반으로 개인화되고 전문적인 대화를 제공한다. 이 프로그램은 심리치료를 보조하기 위한 수단으로 개발되었으며 치료사의 역할을 대체하기 위한 목적은 아니라는 점을 분명히 하고 있다.

최근에는 웹사이트를 기반으로 하는 인터넷 심리치료에서 더 나아가 스마트폰과 같은 모바일 기기를 활용한 정신건강 개입이 활발하게 연구되고 있다. Poushter(2017)에 따르면 2017년 미국인의 77%가 스마트폰을 소유하고 있으며, 2020년까지 전 세계에서 60억 명이 사용하게 될 것으로 보았다. 특히 우리나라의 스마트폰 보급률은 95%에 육박해(Taylor & Silver, 2019) 인구의 대부분이 스마트폰을 사용하는 것으로 볼 수 있다. 즉, 모바일 기기를 활용할 경우 웹사이트를 기반으로 하는 심리치료를 뛰어넘는 접근성이라는 이점을 가지므로써 정신건강 개입의 가능성을 상당히 확장시켜 줄 수 있음을 알 수 있다. 그뿐만 아니라 스마트폰을 활용한 새로운 측정 방식들은 이전에는 상상하기 힘들었던 수준의 개인화되고 정량화된 신체적 건강, 사회적·정서

적 경험에 대한 데이터를 제공할 수도 있다 (Torous & Roberts, 2017). 인간 전문가는, 내담자의 동의하에, 모바일기기를 통해 수집된 자료를 활용하여 내담자를 평가하고 보다 효과적인 개입 전략을 수립할 수 있을 것이다.

TESS가 제공하는 치료적 개입은 모두 정신건강 전문가들에 의해 개발되고 검토된 사항들로, 기본적으로 통합적 접근을 지향한다. 구체적으로는 우울 및 불안 증상을 감소시키는 사실이 반복적으로 확인된 CBT(Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012)를 비롯하여 마음챙김 기반 치료, 수용 전념 치료, 정서 중심 치료, 해결 중심 치료, 그리고 동기 강화 면담 등 임상적으로 그 효과가 검증된 접근법들에 기반한 다양한 개입이 사용자들에게 제공된다. 기술적인 측면에서는 인공지능을 통해 사용자의 입력 중 정서반응이나 특정 주제에 대한 언급 등 특정한 요소를 감지하고, 이에 맞추어 치료적으로 적절한 반응을 제공하도록 설계되어 있다. 제공되는 모든 반응과 지시문은 정신건강 전문가들에 의해서 검토된 내용으로 이루어져 있으며, 사용자가 정서적 반응 혹은 걱정을 표현하는 경우 이에 적절한 공감적 반응을 하고, 이완전략과 같이 특정 정서적 상태를 해소하는데 도움이 되는 기법을 제공한다(Miner et al., 2016). 그뿐만 아니라 TESS는 특정한 치료적 관점이나 사용자 집단의 인구통계학적 특성을 반영하여 제공되는 내용을 개별화 및 맞춤화할 수 있으며, 자신이 제시한 내용에 대한 사용자의 피드백을 수집하고 추가적인 대안을 제공할 수도 있다.

74명의 대학생 참가자들을 대상으로 실시한 연구에서 eBook 전자링크를 통해 정신건강에 대한 정보만을 제공받은 통제집단과 비교했을 때, 2-4주간 TESS를 사용한 처치집단은 범불안

장애 척도(GAD-7)와 우울증 선별도구(PHQ-9)를 통해 측정된 불안과 우울 증상에서 유의미한 감소를 경험하는 것으로 나타났으며, 기술적 한계로 인해 부적절한 상호작용 과정을 경험하거나 적합하지 않은 내용을 제공받은 문제점이 나타나긴 했으나, 사용자들은 전반적으로 높은 수준의 참여도와 만족도를 보고했다(Fulmer et al., 2018).

대화형 에이전트 WOEBOT

Woebot(Fitzpatrick et al., 2017)은 인지행동치료(CBT)를 제공하기 위한 목적으로 구축된 텍스트 기반 대화형 에이전트(Conversational agent)로, 상업적 목적으로 개발된 서비스이다. 이 대화형 에이전트는 간단하고 일상적인 대화 및 기분 추적 기능을 활용하여 사용자에게 CBT의 기본원리와 관련된 치료적 개입을 제공한다. 대화형 에이전트는 자연언어를 사용하는 인간 사용자들에게 사용 맥락에 맞게 적절하게 반응하는 소프트웨어 애플리케이션으로, 잘 알려진 예로는 애플의 시리나 아마존의 알렉사 등이 있다. 이러한 형태의 에이전트는 시각적 인터페이스나 별도의 입력장치를 사용해야 할 필요성을 줄이고, 인간과 기계 사이에 원활하고 연속적인 상호작용을 가능하게 한다(Morris, Kouddous, Kshirsagar, & Schueller, 2018). 초기의 대화형 에이전트들은 if-then 형태의 규칙 목록을 가진 규칙 기반 프로그래밍에 의존했기 때문에 규칙 목록에 포함되지 않은 입력에 대한 반응이나 내장되지 않은 규칙이 적용되는 새로운 환경에서의 상호작용은 불가능했다. 하지만 최근의 대화형 에이전트는 기계학습(Machine Learning)과 보다 발전된 음성인식기술 및 자연어 처리(Natural

Language Processing; NLP)를 기반으로 보다 환경적응적이고 인간과 유사한 상호작용을 제공하는 것을 목표로 한다. 이러한 대화형 에이전트들은 인간이 입력한 대화 규칙에 의존하지 않으며, 많은 사례의 말뭉치를 학습하면서 스스로 그럴듯한 말을 만들어낼 수 있는 방식으로 개발되고 있다(Tanana, Soma, Srikumar, Atkins, & Imel, 2019). Woebot 또한 그 결과물의 일종이라고 할 수 있다.

사용자와 Woebot의 대화가 진행되는 기본 메커니즘은 임상적 의사결정(clinical decision making)과 사회적 담론 역동(the dynamics of social discourse)을 모델링하여 만들어졌으며, 심리교육의 내용은 CBT의 자기 도움(self-help) 내용을 사용했다. 통상적으로 에이전트가 사용자의 현재 맥락이나 기분을 물어보는 방식으로 시작되며, 사용자와의 상호작용은 화면에 제시된 응답을 선택하는 것을 기본으로 하여, 대화의 특정 노드(nod)에서는 자연어 입력을 처리할 수 있는 기능을 갖춘 의사결정 트리 방식을 사용한다. Woebot의 특징 중 하나는 붓이 제공하는 반응 속에 공감적 반응, 반영하기, 테일러링, 목표 설정하기, 책임감 촉진시키기, 동기부여, 관여 높이기와 같은 심리치료의 과정 지향적 특성을 반영했다는 점이다.

다양한 연구들을 통해 인터넷이나 모바일 기반 정신건강 관리 서비스의 효과가 검증되었음에도 불구하고 실제 사용단계에서는 사용자들의 중도탈락이 문제가 되고 있는데(Donkin et al., 2013), Woebot과 같은 발전된 형태의 대화형 에이전트는 이러한 문제에 대한 효과적인 해결책이 될 수 있을 것으로 보인다. 중도탈락 현상이 나타나는 큰 이유 중 하나는 기존의 인터넷 심리치료 서비스나 챗봇 기반 정신건강 개입 서비스가 인간 전문가와 진행하

는 전통적인 면대면 심리치료와 달리 상호작용을 하면서 상대와 관계를 구축하는 감각을 제공하지 못하기 때문인데(Fitzpatrick et al., 2017), Woebot과 같은 보다 발전된 형태의 대화형 에이전트는 사용자들에게 실제 상호작용에 가까운 사용자 경험을 줄 수 있기 때문이다. 물론 현재 단계에서 완벽하게 인간의 상호작용을 모사할 수는 없지만, 관련 분야의 기술 발전 속도는 우리의 생각보다 훨씬 빠르다. 실제로 Bickmore, Gruber, & Picard(2005)는 건강 관련 대화형 에이전트가 사용자와 치료적 관계를 구축할 수 있다는 증거를 제시한 바 있다.

18세에서 28세 사이의 대학생 70명을 대상으로 무선 통제 시행 절차를 통해 Woebot의 사용효과를 검증한 결과, 우울과 불안에 관한 전문적인 정보만을 제공받는 통제집단에 비해 2주간(최대 20회기) Woebot을 사용한 처치 집단에서 우울증 선별도구(PHQ-9)를 통해 측정된 우울 증상이 유의미하게 감소하는 것으로 나타났다. 특히 사용자 경험에 대한 질적 분석에서는 Woebot을 통해 제공된 내용도 중요하지만, 공감의 전달과 같은 치료의 과정 요인이 효과를 발휘했음을 시사하는 내용들이 드러났다(Fitzpatrick et al., 2017). 이는 심리치료 및 정신건강 개입을 위한 대화형 에이전트를 개발하고 발전시켜 나가는 과정에서 대화의 내용뿐만 아니라 대화의 방식을 충분히 고려해야 할 필요가 있다는 점을 시사한다.

대화형 에이전트 CLIENTBOT

대화형 에이전트를 심리치료 및 정신건강 개입에 활용하는 또 다른 방향은 내담자 역할을 하는 챗봇인 Clientbot의 개발이다. 특정 행

등의 효과적인 학습을 위해서는 즉각적인 피드백이 중요하다는 오래된 학습원리(Kluger & DeNisi, 1996)에도 불구하고, 상담 및 심리치료 슈퍼비전 맥락에서의 피드백은 물리적, 시간적 제약으로 인해 즉각적이지 못하고 모호한 경향이 있다(Beutler & Harwood, 2004). Clientbot은 이런 어려움을 극복하기 위해 개발된 것으로, 실제 상호작용을 하는 과정 속에서 수련 상담자들에게 즉각적이면서 성과에 기반한 피드백을 제공할 수 있도록 개발된 신경 대화형 에이전트이다.

Clientbot은 동기 강화 면접의 원리를 바탕으로 열린 질문(open-question)과 반영(reflection) 기술을 훈련시키는데 초점이 맞춰져 있다. 즉, 이 시스템은 사용자가 열린 질문과 반영을 사용하도록 촉진하며, 사용자의 반응에 대해 기계학습 기반 피드백을 제공한다. 이를 통해 사용자는 자신의 반응이 열린 질문 혹은 반영인지 즉각적으로 피드백 받게 됨으로써 해당 기술을 훈련할 수 있는 기회를 가진다(Tanana et al., 2019). 사용자와의 대화를 위한 Clientbot의 신경 대화 모델에는 장단기 기억 셀(Long short-term memory; LSTM cell)이라는 신경망을 기반으로 하는 두 개의 서로 다른 LSTM 순환 신경망이 사용된다. 하나는 사람의 잡담과 같은 자연스러운 반응을 만들어내지만 긴 반응을 만들어 내는 데는 서투른 Sequence to Sequence 모델이며, 나머지 하나는 길이가 길고 비결정적인 반응을 만들어낼 수 있는 단순한 LSTM 모델이다. 또한 사용자의 반응을 기초 상담 기술로 분류하고 이에 대한 피드백을 제공하기 위해 기계 학습 모델이 포함되며, 이 모델은 동기 강화 면접을 하기 위해 수련 상담자들을 훈련하는데 사용된 2,354개의 축어록 속에 등장하는 514,118개의 말뭉치를 통해

학습이 이루어졌다.

총 151명의 상담자 훈련 경험이 없는 비상담자들 대상으로 무선 할당 처치를 통해 Clientbot을 통한 훈련 성과와 사용자 경험을 탐색한 결과, 인간 모의 내담자와 대화하며 자신의 반응에 대한 피드백을 받지 못한 처치 집단에 비해, 20분간 내담자 봇을 사용하면서 열린 질문과 반영을 연습하도록 하는 과정에서 Clientbot으로부터 반응에 대한 피드백을 받은 처치 집단이 더 많은 반영을 사용한 것으로 나타났다. 이러한 결과는 정신건강 전문가를 양성하기 위한 효과적인 훈련 시스템을 개발하는데 인공 신경망과 기계 학습을 기반으로 하는 발전된 형태의 대화형 에이전트가 어떠한 역할을 할 수 있는지를 제시한다는 점에서 의미가 있다.

가상현실 기반 대화형 에이전트 ELLIE

가상현실(Virtual Reality, VR)은 인간-컴퓨터 인터페이스의 일종으로, 사용자가 컴퓨터가 만들어낸 가상의 환경을 경험하고 상호작용하게 한다(Rizzo, Buckwalter, & Neumann, 1997). 현재 가장 널리 알려진 VR 기반 대화형 에이전트는 Ellie로(Katharine, 2019), PTSD를 가진 군인을 검진하는 것을 보조할 목적으로 미국 남캘리포니아 대학(USC)의 창의기술연구소에서 개발된 SimSensei 키오스크 프로그램(DeVault et al., 2014) 속에 등장하는 면담자이다. SimSensei 키오스크는 미국의 Defense Advanced Research Projects Agency(DARPA)에서 진행하고 있는 ‘심리적 신호에 대한 탐지 및 컴퓨터 분석 프로젝트(Detection and Computational Analysis of Psychological Signals; DCAPS Project)’와 관련하여 개발한 가상 인간 면담시스템(Virtual

human interviewer system)이다. 이 프로그램은 임상적인 의사결정을 보조하는 도구로 개발되었으며, 모니터 화면 속에 3D로 구현된 Ellie라는 가상의 상담자와 내담자 간의 상호작용을 통해 언어적, 비언어적 정보를 수집하고 이를 바탕으로 우울, 불안, PTSD와 같은 내담자의 심리적 스트레스의 지표들을 평가하는 시스템을 구축하는 것을 목표로 한다(Devault et al., 2014). 개념적으로 봤을 때 이 시스템은 VR기술을 활용하여 시각적으로 제공되는 진보된 형태의 대화형 에이전트와 인간의 심리적 상태에 대한 임상적 의사결정 지원 시스템의 통합으로 볼 수 있다.

사용자 경험 측면에서 이 시스템의 목표는 Ellie와 대화하는 내담자가 편안함을 느끼고 정보를 공유하는데 불편함을 경험하지 않을 것, 시스템은 내담자의 비언어적인 반응과 행동에 민감할 것, 수집한 비언어적인 행동들을 수량화하여 특정한 심리학적 상태를 평가하는데 적용할 수 있을 것, 적절한 비언어적 행동(예, 제스처, 표정, 시선, 포즈 등)을 구현할 수 있을 것 등이다. 이를 위해 기존의 경험적 연구들과 이론들, 임상적 인터뷰 전문가들의 자문을 바탕으로 심리적 스트레스를 경험하고 있는 사람들의 언어적, 비언어적 특성을 구분하여 데이터베이스화시켰다. SimSensei는 시청각 지각 시스템(Visual-Audio perception system)인 MultiSense 프레임워크를 통해 실시간으로 사용자의 언어적, 비언어적 행동을 수집하고 이를 데이터베이스와 비교함으로써 임상적 의사결정에 도움을 줄 수 있는 정보를 찾아낸다. 또한 음성인식과 자연어 처리 시스템, 그리고 대화 관리 모듈을 통해 사용자의 발화 내용을 분석한다. 이상의 정보들을 통합하여 사용자의 정서적 상태를 추론하고, 이를 바탕으로

임상적 인터뷰를 진행하기 위한 질문 및 반응이 이루어진다. 이 과정에서 Cerebella라는 행동 생성 시스템을 통해 적절한 비언어적 행동을 구현하여 음성반응과 함께 시각적으로 제공한다.

최근 많이 연구되고 있는 텍스트 기반 대화형 에이전트의 경우 사용자와 상호작용하기 위해서 주로 메신저와 같은 형태의 채널을 사용하는 반면, Ellie를 구현하는 가상현실 기술을 활용한 대화형 에이전트의 경우 가상의 면담자를 시각적으로 보고 있는 상태에서 음성을 통해 상호작용하기 때문에 사용자가 더 자연스럽게 경험하고 몰입할 수 있는 면담 상황을 유도할 수 있다. 그뿐만 아니라 표정, 움직임, 포즈, 목소리의 높낮이, 빠르기, 시선 등 텍스트 기반 상호작용에서는 수집할 수 없는 언어적·비언어적 행동들을 수집하고 분석하여 면담에 활용할 수도 있다.

Simsensei와 인터뷰를 실시한 사용자들을 대상으로 한 연구결과를 살펴보면, 해당 시스템은 가까워지고자 하는 의도, 타인에게 추천 의향, 전반적인 만족도에서 우수한 점수를 보이는 것으로 나타났다. 또한 Ellie와의 대면 인터뷰 상황을 편안하게 느끼는 것으로 보고했다. 물론 현재로서는 인간과 비교했을 때 라포의 수준도 낮고, 비언어적인 의사소통에서도 능숙하지 못해 사람들 간의 자연스러운 상호작용을 완벽히 구현할 수 있는 수준은 아니지만 인간의 개입 없이 자율적으로 인터뷰를 비교적 원활하게 진행할 수 있는 인공지능 시스템이 개발될 가능성이 있는 것으로 평가된다(Devault et al., 2014).

논 의

본 연구는 인공지능 기술을 활용한 치료적 개입의 현재를 알아보기 위해 최근 3년 간 학술 데이터베이스에 등록된 상담 및 심리치료 분야 연구논문들 중 인공지능 기술을 적용한 치료적 개입을 개발하고, 이를 활용한 결과를 발표한 논문들을 검토하였다. 검토 과정에서 연구결과물의 실효성이나 효과성 검증 결과가 보고되었고 기술적 영역이 중복되지 않은 최신의 연구들을 대상으로, 인공지능 상담 연구 경험이 있는 상담전공 교수의 자문을 통해 다섯 가지 주요 사례들이 선정되었다. 결과에서는 주요 사례들의 개발배경, 작동방식, 개입의 효과성, 장점과 한계를 포함한 상담 분야에서의 의의가 제시되었으며, 이를 통해 인공지능 상담의 현재 기술 수준이 사람들의 행동 및 정신건강을 효과적으로 관리하는데 도움이 될 뿐만 아니라, 앞으로 치료적 개입의 효율성과 효과성을 높이는 방향으로 발전할 가능성이 높다는 점을 확인할 수 있었다.

Luxton(2014)은 인공지능 기술이 인간 전문가의 직업을 위협하거나 대체하기 위한 것이 아니며, 오히려 정신건강 분야의 전문가들이 치료적 개입에 활용될 수 있는 인공지능의 개발에 직접 참여하거나, 치료적 개입 과정에서 인공지능 관련 기술을 활용함으로써 경제적 이익을 얻을 뿐만 아니라 개입의 효과성을 높여 궁극적으로 환자와 내담자들의 정신건강 복지에 도움이 될 수 있음을 강조한다. 구체적으로는 인공지능 기술을 활용한 치료적 개입의 개발에 바탕이 되는 이론적 기반과 근거 기반 치료에 관한 전문적인 지식을 상담 및 심리치료 분야의 전문가들이 제공함으로써 시스템 개발자를 도울 수 있을 것이다.

인공지능 상담의 기술적 발전은 치료적 역명성을 높여 낙인 등의 이유로 상담 및 심리 치료에 저항하는 치료적 사각지대에 놓인 내담자들의 복지에 기여할 것이다. 홍진표 등(2016)의 정신질환실태 역학조사 연구에 의하면, 한국의 성인 4명 중 1명이 평생 1번 이상의 정신질환을 겪지만, 그 중 22.2%만이 전문가에게 정신건강 문제를 의논하거나 치료받은 경험이 있는 것으로 나타났다. 한국문화에서는 자신의 정신건강의 문제를 드러내는 것을 꺼리며, 체계적이고 전문적인 도움을 받는 것에 대해서도 거부감이 있는 것으로 보인다(박성희, 이동렬, 2008). 인공지능 기술을 접목한 치료적 개입을 통해 자신의 심리적 어려움을 외부에 드러내지 않고도 전문적인 도움을 받을 수 있는 환경이 구축된다면 정신건강의 회복을 위해 개입이 필요하지만 낙인 등의 이유로 행동에 나서지 못하고 있는 잠재적 내담자들의 심리적 문제 해결에 상당한 도움이 될 것이다.

그러나 앞서 소개한 사례들과 각 사례에 적용된 기술들을 살펴보면 현재의 기술 수준으로 구축한 인공지능은 아직은 인간 사용자와의 자연스러운 상호작용이 어려우며, 따라서 인간 상담자의 역할을 완전히 대체할 수도 없다는 것을 알 수 있다. 하지만 인공지능을 통해 구동되는 챗봇이나 대화형 에이전트의 예에서 볼 수 있듯, 사람들의 정신건강을 증진하기 위해 인공지능 관련 기술을 활용하려는 노력은 이미 시작되었고, 실제로 성과를 나타내고 있으며, 다양한 가능성을 보여주고 있는 분명하다. 이에 인공지능 상담의 발전과 밀접한 관련이 있는 최근의 첨단 기술적인 이슈인 자연어 처리(NLP), 인공지능의 정서 인식 및 공감 표현, 인공지능의 자기학습(self-feeding)

의 측면에서 현재 진행되고 있는 인공지능 상담의 현재와 향후 전망을 논의하고자 한다.

자연어 처리

상담 및 심리치료 분야에서 인공지능 기술을 활용하기 위해 반드시 필요한 기술 중의 하나가 바로 기계와 인간 사이의 의사소통을 가능하게 하는 기술이다. 기계와 인간의 의사소통을 위해 다양한 방식이 활용될 수 있지만, 그 중에서도 특히 인간의 자연어(natural language)를 인식하고 이해하여 그에 적합한 반응을 출력해내는 것이 중요하며 이를 연구하는 분야가 바로 자연어 처리이다. 자연어 처리는 기계학습의 일종으로, 인간의 자연언어를 해석하고 처리하는 기계의 능력을 의미하며 컴퓨터 과학과 언어학이 통합된 인공지능 연구의 하위 분야이다. 앞서 설명한 ELIZA나 PARRY가 이런 자연어 처리 시스템의 초기 형태로 볼 수 있으며, 발전된 형태의 자연어 처리 시스템의 대표적인 예는 앞서 언급한 IBM의 Watson이다. Watson은 진화된 자연어 처리 시스템(Advanced Natural Language Processing System), 의미 분석(semantic analysis), 정보 검색, 통계적 분석, 자동화된 추론 과정, 기계 학습 등의 기술이 집약된 DeepQA라는 기술 프레임워크를 기반으로 제시된 문제를 인식 및 이해하고, 가능한 모든 대답을 산출한 뒤, 찾아낸 자료로부터 얻어낸 증거를 바탕으로 각 대답을 점수화하여 각각이 정답일 확률을 계산해내는 방식으로 작동한다. 한국에서도 자연어 처리 시스템의 개발이 이루어지고 있다. 한국 전자통신연구원(ETRI)의 총괄 하에 진행되고 있는 EXOBRAIN 프로젝트는 2023년까지 인간의 지적 노동을 보조할 수 있는 자연어 처리

분야의 인공지능 기술 개발을 목표로 하고 있으며, 이를 위한 세부 과제로 자연어 분석 기술, 지식 학습 및 추적 기술, 자연어 질의응답 기술을 개발하고 있다. 이 시스템은 지난 2016년 11월에, Watson과 유사하게, 장학퀴즈 프로그램에서 4명의 인간 우승자들을 상대로 가장 높은 점수를 획득하여 우승하였으며, 2017년 10월에는 인공지능 기술 및 데이터의 활용 보급을 위해 언어분석, 어휘 관계 분석, 질문 분석, 음성인식이 가능한 공개 API(Open Application Programming Interface)를 제공하고 있다(<http://exobrain.kr>). 자연어 처리 기술은 사용자와의 자연스러운 상호작용을 가능하게 한다는 점에서 인공지능 기술을 활용한 정신건강 관리에 상당히 중요하고 유용한 기술이며, 향후 인공지능 상담의 ‘완전한’ 실현 가능성을 예측하는 가늠자 역할을 할 것으로 보인다. 예를 들어, 기계 학습 기술과 결합된 자연어 처리 기술은 가상 인간(virtual human agent)과 인간 사용자가 문자 혹은 음성으로 자연스럽게 의사소통할 수 있는 수단을 제공하고, 음성 혹은 의미 분석을 통해 사용자의 심리적·정서적 상태에 대한 정보를 제공하며, 상담 회기의 내용을 분석하여 치료에 도움이 되는 패턴을 식별하는 활용될 수도 있다(Luxton, 2014). 이미 이와 관련된 연구 또한 활발히 진행되고 있다(Imel, Steyvers, & Atkins, 2015).

정서 인식 및 공감 표현

공감은 심리치료의 결과를 예측할 수 있게 하는 치료적 동맹의 강한 예측 변인이며(Nienhuis et al., 2018), 내담자의 변화를 이끌어내는 주요한 기제(Elliott, Bohart, Watson, & Greenberg, 2011)이다. 따라서 상담 및 심리치

료를 제공하기 위한 인공지능을 설계할 때 인공지능이 사용자의 정서적 경험을 진심으로 이해하는 것처럼 보이게 하는 개인화된 반응을 생성하고 전달하는 것은 매우 중요한 문제가 된다. 하지만 현재까지 개발된 인공지능을 활용하여 치료적 개입을 하는 대화형 에이전트들은 공감적인 반응을 제공하지 못하거나 매우 제한적인 수준에서만 가능한 것으로 보인다. 앞서 살펴보았던 Woebot 역시 최신의 기술력을 반영한 대화형 에이전트이며 사용자의 정서적 고통에 대한 표현을 식별하고 친절히 반응하긴 하지만, 통상적이고 보편적인 의미의 위로인 경우가 많다(예, “당신이 외로움을 느낀다니 유감입니다. 때때로 나는 우리 모두가 외로움을 느낀다고 생각해요”). 특히 상담자가 내담자의 고유한 상황을 이해하고 이에 맞게 반응하는 세련된 방식의 공감을 표현하는 것은 현재로서는 여전히 불가능한 것으로 보인다(Morris et al., 2018).

이러한 문제를 해결하기 위해 진행된 Morris 등(2018)의 연구는 대화형 에이전트의 공감 표현에 대한 가장 진보된 결과를 보여주는 연구 결과 중 하나이다. 이들은 기계 학습의 일종인 비지도학습을 활용하여 실제와 비슷한 공감을 즉각적이고 설득력 있게 전달할 수 있는 대화형 에이전트의 초기 단계를 구축하고 사용자의 경험을 분석하였다. 이를 위해 정신건강 관리를 목적으로 웹사이트 기반으로 서비스되고 있는 상업적 서비스 플랫폼인 KoKo를 통해 수집한 정서적 상호작용 말뭉치(72,786개의 게시글과 339,983개의 반응)로부터 정서적 지지 진술을 추출하는 말뭉치 기반 접근법을 사용하여 개별 사용자에게 맞추어진 공감적 표현을 만들어냈다. 기본적으로는 사용자의 정서가 표현된 입력이 주어지면 기존의 말뭉

치에서 주어진 입력과 유사한 글을 찾고 그 글에 뒤따라온 반응을 마치 해당 사용자만을 위해 애플리케이션이 만들어 낸 고유한 표현처럼 전달하는 방식이다.

연구자들은 개발된 에이전트로부터 제시된 공감적 반응들에 대한 사용자의 평가와 인식을 조사했다. 3점 리커트 척도(bad, ok, good)로 평가했을 때 ‘good’ 평가를 받은 반응의 비율은, 에이전트가 만들어낸 반응에 비해 인간이 만들어낸 반응에서 유의미하게 높긴 했지만, 에이전트가 만들어낸 반응의 79.2%가 ‘ok’ 또는 ‘good’으로 평가된 것으로 나타났다. 이러한 결과 자체도 흥미롭지만 주목해야 할 점은 에이전트의 학습에 사용된 말뭉치의 수량이 많아질수록 에이전트의 반응은 더 정교하고 인간의 반응과 유사해질 수 있으며, 에이전트와 인간이 각각 만들어낸 반응에 대한 평가자들의 평가 결과를 에이전트가 학습하여 스스로 더 나은 반응을 탐색하는 알고리즘을 만들어낼 수 있다는 점이다. 물론 공감은 단순한 표현방식의 문제가 아니며 에이전트가 만들어낸 반응이 공감의 본질에 접근했다고 할 수도 없다. 하지만 이러한 접근 방식은 대화형 에이전트가 사람들 간에 이루어지는 사회적 지지를 데이터화하고 이로부터 공감적 표현을 시뮬레이션해낼 수 있음을 보여준다는 점에서 인공지능 상담의 미래에 시사하는 바가 크다.

인공지능의 자기학습, Self-Feeding bot

앞서 살펴본 바와 같이 최근에 개발되고 있는 대화형 에이전트들은 사용자와의 상호작용 능력을 향상시키는 기계 학습을 활용한다. 특히 상담 축어록과 같이 이미 완료된 말뭉치를 통해서 학습하는 일종의 지도학습이 많이 활

용되고 있다. 하지만 이런 지도학습을 위해 상담 및 심리치료 맥락에서 이루어진 말뭉치들을 대량으로 수집하는 것이 쉽지 않고, 가능하더라도 큰 비용이 들기 때문에 이에 대한 대안으로 일반적인 맥락에서 이루어진 대화를 스크랩한 말뭉치를 학습시키는 경우가 있다. 하지만 이 경우에도 말뭉치 속의 대화들이 이루어졌던 맥락이 정신건강 개입을 위한 에이전트가 실제로 사용자와 대화를 하게 될 상황이나 환경과는 차이가 있는 경향이 있다(Ross, Zaldivar, Irani, & Tomlinson, 2009).

이러한 어려움을 극복하기 위해서 제안된 대화형 에이전트가 Self-feeding 챗봇이다(Hancock, Bordes, Mazare, & Weston, 2019). 이러한 형태의 대화형 에이전트는 주어진 기존의 자료를 통해서도 학습하지만, 자신이 참여하고 있는 현재의 대화 속에서 학습에 필요한 예제를 수집한다. 이 메커니즘 하에서 에이전트는 사용자와 대화하면서 상대방이 보이는 반응을 통해 에이전트가 직전에 보인 반응에 대한 상대방의 만족도를 예측한다. 좋은 만족도가 예상되면 대화로부터 학습할 자료를 추출하고, 나쁜 만족도가 예상되면 어떤 반응이 적절할지를 사용자에게 피드백을 요청함으로써 자신의 학습에 사용할 추가적인 데이터를 확보해나간다. 이러한 방식을 통해 대화형 에이전트는 풍부하고, 맥락 혹은 과제 특정적이며, 역동적이고, 경제적인 방식으로 자신의 능력을 발전시켜 나갈 수 있다. 이는 에이전트가 인간이 대화하는 것을 관찰하여 학습하는 것에 그치지 않고, 마치 아동이 대화능력을 습득하는 것과 마찬가지로, 실제 대화에 참여하여 그 속에서 주어지는 상대방의 피드백에 기초해 자신의 대화 능력을 능동적으로 조정해나가는 것을 의미한다(Bassiri, 2011; Werts,

Wolery, Holcombe, & Gast, 1995). 물론 현재는 시작단계에 있지만, 치료적 개입을 목적으로 하는 인공지능에 이러한 형태의 자기학습 과정을 적용할 수 있다면 내담자와의 상호작용의 질을 향상시키는데 상당한 진보가 있을 것이며, 이는 인공지능 상담의 발전에 중요한 기반이 될 수 있을 것이다.

이상의 기술적 이슈들을 살펴보면 인공지능 상담은 아직 기술적 측면에서 넘어야 할 장애물과 한계를 지니고 있는 동시에, 그러한 장애물과 한계를 넘기 위한 개념들이 제안되고 실제로 구체화시키기 위한 연구 및 시도들이 활발히 이루어지고 있다는 것을 알 수 있다.

본 연구는 상담 및 심리치료 분야에서 인공지능 기술을 활용한 상담의 최근의 발전 과정을 살펴보고, 구체적인 사례를 소개하면서 이와 관련된 기술적 이슈를 살펴보았다는 점에서 의의가 있다. 지금까지 국내 연구 중에서는 ‘온라인’ 심리치료에 대한 전반적인 리뷰(이동훈 등, 2015)가 유일하며, 최근 들어 특정 인공지능 기술을 상담 및 심리치료 장면에 적용한 연구들, 즉 웹사이트 기반 심리치료(안정광, 권정혜, 윤혜영, 2014), 챗봇(이지원, 양현정, 김지근, 2019), 클라우드소싱 기반 심리치료 애플리케이션(한소영, 김향숙, 김주섭, 2017) 등과 같은, 특정 기술을 상담 및 심리치료 장면에 적용하고 그 효과를 검증한 개별적인 논문들은 있었지만, 인공지능 기술을 활용한 상담 전반의 최근 현황과 기술적 이슈를 다룬 연구가 없었다.

그럼에도 불구하고 본 논문은 다음과 같은 한계점을 가지며 이를 넘어서기 위한 후속 연구가 필요하다. 첫째, 인공지능 기술이 활용된 상담 및 심리치료 개입의 주요 사례들을 선정하는 과정에서 국내 연구결과들이 배제되었다

는 점이다. 국내 학술 데이터베이스를 검색한 결과를 살펴보면 현재 국내에는 인공지능 기술을 적용하여 개발한 치료적 개입 프로그램 혹은 이와 관련된 연구가 극히 드문 것으로 보인다. 이러한 현상에는 여러 가지 이유가 있겠지만 한국어의 자연어 처리 기술 수준과도 관련이 있는 것으로 보인다. 상대적으로 앞선 연구결과들을 보여주고 있는 영어의 경우 비교적 자연스러운 자연어 처리가 가능한 기술이 개발되어 있고 이미 여러 분야에서 활용되고 있는 반면, 한국어의 경우 자연어 처리 자체가 타 언어에 비해 어렵고, 개발된 기술조차도 쉽게 접근할 수 없는 상황이다. 다행히 현재 한국전자통신연구원(ETRI)의 EXOBRAIN 프로젝트를 통해 한국어 자연어 처리 기술 개발이 진행되고 있으며, 개발이 완료되어 일반에 공개된다면 앞으로 이를 활용한 치료적 개입의 개발이 활성화될 것으로 생각된다. 이는 자연어 처리에서의 제한점에도 불구하고 국내에서도 인공지능 기술을 상담 및 심리치료 장면에서 적용하고자 하는 시도들이 시작되고 있다는 점에서도 확인할 수 있는데, 예를 들어 이지원 등(2019)은 상담 챗봇 구현을 위한 시나리오를 개발하고 그 유용성을 확인하면서 챗봇만의 고유한 장점을 소개한 바 있으며, 강은빈(2019)은 접수 면접 상황에서 효과적인 상담 챗봇 설계 전략 방안을 제시한 연구를 발표하기도 했다. 앞으로 이러한 연구결과들이 축적되면서 이론적 기반을 형성하고, 기술적 발전이 뒷받침된다면 국내에서 또한 인공지능 상담에 대한 본격적인 연구들이 이루어질 수 있을 것이다.

둘째, 본 연구를 통해 소개한 사례들의 효과성 검증을 위해 진행된 연구들은, 일부를 제외하면, 대부분 단기 연구이며 기존의 개입

효과성 검증 연구들과 같이 임상적인 수준에서 엄격한 검증이 이루어지고 있지는 않았다. 이러한 현상은 본 연구에서 소개되지 않은 다른 연구들에서도 마찬가지였으며, 이는 인공지능 관련 기술의 적용이 아직 시작 단계에 있어 적용에 따른 효과를 철저히 검증하기 보다는 기술 자체를 소개하고 이를 활용하는 방법에 대한 소개가 필요하기 때문으로 보인다. 하지만 심리적 개입을 통한 정신건강 상의 변화는 단기간에 이루어지는 것이 아니며 대개 수 주 또는 수개월 이상의 지속적인 행동 변화를 필요로 한다. 이러한 측면에서 단기 연구에 의해 검증된 효과성이 실제 내담자들에게 임상적인 수준에서의 효과를 나타낼지에 대한 의문이 존재할 수 있으며, 이러한 의문은 상담 및 심리치료 분야에서 인공지능 기술의 활용 및 관련 연구들을 촉진하는데 장애물이 될 수 있다. 따라서 장기적인 관점에서 임상적으로 엄격하게 통제된 연구들을 통해 그 효과성을 검증하는 연구들이 필요할 것으로 보인다.

셋째, 본 연구에서는 인공지능을 활용한 치료적 개입 연구들을 중심으로 소개했으나 심리적 진단 및 평가 장면에서 인공지능 기술을 활용한 사례들에 대한 소개도 필요할 것이다. 대표적인 예로 임상적 의사결정 지원 시스템(Clinical Decision Support Systems; CDSSs)이 있다(Bennett, Doub, & Selove, 2012). 기존의 CDSSs의 경우 규칙 기반 시스템으로 설계되어 현장에서의 실용성이 떨어지거나, 판단의 준거가 되는 증거 기반 규칙을 주기적으로 업데이트하는 과정에서 시스템의 최신성이 떨어지는 단점이 있었다. 하지만 기계학습, 인공 신경망과 같은 인공지능 기술을 활용한 CDSSs는 자료 기반 확률 추론이 가능하기 때문에(Bellazzi

& Zupan, 2008) 복잡하고 방대한 데이터를 분석하여 사실들(검사결과, 인구통계학적 특성, 증상, 처치 등) 간의 관계를 추론해내고 이를 바탕으로 인간 전문가가 임상적 의사결정에 활용할 수 있는 정보, 더 나아가 권고사항을 제공할 수 있다. 이러한 시스템은 독립적으로 개입을 제공하거나 상담을 진행할 수는 없지만 인간 상담자가 인식할 수 없는 정보를 획득하거나, 인간 상담자로서는 처리할 수 없는 수준의 대량의 정보를 신속하게 처리하여 사실들 간의 관계를 발견하여 제공하는 등 다양한 방식으로 상담자를 지원할 수 있다는 장점이 있기 때문에 앞으로 활용의 가능성이 많으며 따라서 후속 연구에서는 이에 대한 소개가 필요할 것이다.

인공지능 기술에 대한 사회적 관심과 투자는 막대하며 그에 따라 발전 속도 또한 매우 빠르다. 한국전자통신연구원은 최근 ‘2020년 인공지능(AI) 7대 트렌드’라는 정책 보고서를 발표하며 인공지능이 연구자로서 인간이 생각하는 방식을 바꿈으로써 산업 자체를 혁신하고 있을 뿐만 아니라, 인간의 고유한 영역으로 여겨졌던 창작분야에서도 기능할 수 있다고 언급하였다. 이는 인공지능이 기존의 예상보다 훨씬 광범위하고 깊은 수준에서 우리 삶의 변화를 일으킨다는 것을 의미한다. 그리고 이러한 추세는 상담 및 심리치료 분야에서도 마찬가지로의 변화로 나타날 것이다. 인공지능의 사용은 인류의 복지 증진에 초점을 맞추어 이루어져야 한다는 ‘사람 중심 인공지능’의 맥락에서, 상담 및 심리치료 분야에서의 인공지능 기술의 활용은 무엇보다도 윤리적인 이슈들이 중요한 문제로 떠오를 것으로 예상된다. 따라서 후속 연구에서는 인공지능 상담의 윤리적 문제들이 이론적 현실적 측면에서 자세

히 고찰될 필요가 있다.

참고문헌

- 강은빈 (2019). 챗봇의 의인화 특성이 사용자 의 자기노출과 에이전트와의 관계형성에 미치는 영향. 연세대학교 대학원 석사학위논문.
- 박성희, 이동렬 (2008). 한국 문화와 상담. 서울: 학지사.
- 박태진 (2017, 5, 25). 알파고 바둑 실력의 비밀, ‘딥 러닝(Deep Learning)’. UNIST News Center. https://news.unist.ac.kr/kor/column_202/에서 검색.
- 심혜숙, 이현진 (2003). 사이버 상담에서 상담자의 경력에 따른 자기 대화 및 가설형성 차이 분석. 한국심리학회지: 상담 및 심리치료, 15(1), 1-16.
- 안정광, 권정혜, 윤혜영 (2014). 인터넷 기반 사회불안장애 인지행동 자가 치료 프로그램 개발 및 효과검증. *Korean Journal of Clinical Psychology*, 33(4), 695-721.
- 이동훈, 김주연, 김진주 (2015). 온라인 심리치료의 가능성과 한계에 대한 탐색적 연구. 한국심리학회지: 상담 및 심리치료, 27(3), 543-582.
- 이지원, 양현정, 김지근 (2019). 상담 챗봇 구현을 위한 시나리오 개발 및 유용성 검증. 한국콘텐츠학회논문지, 19(4), 12-29.
- 이재영, 정영운 (2001). 사이버 집단상담 효과에 관한 연구: 공업계 고교생을 위한 진로 탐색 프로그램을 중심으로. 한국심리학회지: 상담 및 심리치료, 13(2), 143-160.
- 한소영, 김향숙, 김주섭 (2017). 클라우드소싱

- 기반 심리치료앱 UX 디자인. 한국 HCI 학회 학술대회, 158-161.
- 홍진표, 이동우, 함봉진, 이소희, 성수정, 윤탁, 하태현, 손상준, 손정우, 유제춘, 김정란, 박종익, 김성환, 조성진, 정영철, 김문두, 장성만, 김병수, 안준호, 김봉조, 윤진상, 신일선 (2016). 2016년도 정신질환 실태조사. 보건복지부 · 삼성서울병원.
- Abdul-Kader, S. A., & Woods, J. C. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, 6(7), 72-80.
- Aguilera, A. (2015). Digital technology and mental health interventions: Opportunities and challenges. *Arbor: Ciencia, Pensamiento y Cultura*, 191(771), 210.
- Álvarez-Jiménez, M., & Gleeson, J. F. (2012). Connecting the dots: twenty-first century technologies to tackle twenty-first century challenges in early intervention. *Australian & New Zealand Journal of Psychiatry*, 46(12), 1194-1196.
- Álvarez-Jiménez, M., Gleeson, J. F., Bendall, S., Lederman, R., Wadley, G., Killackey, E., & McGorry, P. D. (2012). Internet-based interventions for psychosis: A sneak-peek into the future. *Psychiatric Clinics*, 35(3), 735-747.
- Andersson, G. (2002). Psychological aspects of tinnitus and the application of cognitive-behavioral therapy. *Clinical Psychology Review*, 22(7), 977-990.
- Bassiri, M. A. (2011). Interactional feedback and the impact of attitude and motivation on noticing 12 form. *English Language and Literature Studies*, 1(2), 61.
- Beck, J. G., Palyo, S. A., Winer, E. H., Schwagler, B. E., & Ang, E. J. (2007). Virtual reality exposure therapy for PTSD symptoms after a road accident: An uncontrolled case series. *Behavior Therapy*, 38(1), 39-48.
- Bellazzi, R., & Zupan, B. (2008). Predictive data mining in clinical medicine: current issues and guidelines. *International Journal of Medical Informatics*, 77(2), 81-97.
- Bennett, C. C., Doub, T. W., & Selove, R. (2012). EHRs connect research and practice: Where predictive modeling, artificial intelligence, and clinical decision support intersect. *Health Policy and Technology*, 1(2), 105-114.
- Beutler, L. E., & Harwood, T. M. (2004). Virtual reality in psychotherapy training. *Journal of Clinical Psychology*, 60(3), 317-330.
- Bickmore, T., Gruber, A., & Picard, R. (2005). Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Education and Counseling*, 59(1), 21-30.
- Bickmore, T. W., Pfeifer, L. M., Byron, D., Forsythe, S., Henault, L. E., Jack, B. W., ... & Paasche-Orlow, M. K. (2010). Usability of conversational agents by patients with inadequate health literacy: Evidence from two clinical trials. *Journal of health communication*, 15(S2), 197-210.
- Bouchard, S., Paquin, B., Payeur, R., Allard, M., Rivard, V., Fournier, T., & Lapierre, J. (2004). Delivering cognitive-behavior therapy

- for panic disorder with agoraphobia in videoconference. *Telemedicine Journal and E-health*, 10(1), 13-25.
- Celio, A. A., Winzelberg, A. J., Wilfley, D. E., Eppstein-Herald, D., Springer, E. A., Dev, P., & Taylor, C. B. (2000). Reducing risk factors for eating disorders: Comparison of an Internet-and a classroom-delivered psychoeducational program. *Journal of Consulting and Clinical Psychology*, 68(4), 650-657.
- Childress, C. A. (2000). Ethical issues in providing online psychotherapeutic interventions. *Journal of Medical Internet Research*, 2(1), e5.
- Cohen, G. E., & Kerr, B. A. (1999). Computer-mediated counseling: An empirical study of a new mental health treatment. *Computers in Human Services*, 15(4), 13-26.
- Colby, K. M. (1977). Appraisal of four psychological theories of paranoid phenomena. *Journal of Abnormal Psychology*, 86(1), 54-59.
- Collin, P., Rahilly, K., Richardson, I., & Third, A. (2011). Literature review: The benefits of social networking services. *YAW-CRC's Partner Organisations*, 1-29.
- Corrigan, P. W., & Rüsch, N. (2002). Mental illness stereotypes and clinical care: Do people avoid treatment because of stigma?. *Psychiatric Rehabilitation Skills*, 6(3), 312-334.
- Dagöö, J., Asplund, R. P., Bsenko, H. A., Hjerling, S., Holmberg, A., Westh, S., ... & Andersson, G. (2014). Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: A randomized controlled trial. *Journal of Anxiety Disorders*, 28(4), 410-417.
- D'Alfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., Gleeson, J., & Alvarez-Jimenez, M. (2017). Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in Psychology*, 8, 796.
- Dennis, C. L. (2003). Peer support within a health care context: A concept analysis. *International Journal of Nursing Studies*, 40(3), 321-332.
- DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., ... & Lucas, G. (2014, May). SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (pp. 1061-1068). International Foundation for Autonomous Agents and Multiagent Systems.
- Difede, J., & Hoffman, H. G. (2002). Virtual reality exposure therapy for World Trade Center post-traumatic stress disorder: A case report. *Cyberpsychology & Behavior*, 5(6), 529-535.
- Donkin, L., Hickie, I. B., Christensen, H., Naismith, S. L., Neal, B., Cockayne, N. L., & Glozier, N. (2013). Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: Randomized controlled trial. *Journal of Medical Internet Research*, 15(10), e231.
- Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy*, 48(1), 43.
- Fitzpatrick, K. K., Darcy, A., & Vierhile, M.

- (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, 4(2), e19.
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99(1), 20-35.
- Franchi, S., & Guzeldere, G. (1995). Constructions of the mind: Artificial intelligence and the humanities. *Stanford Humanities Review*, 4(2).
- Frueh, B. C., Monnier, J., Grubaugh, A. L., Elhai, J. D., Yim, E., & Knapp, R. (2007). Therapist adherence and competence with manualized cognitive-behavioral therapy for PTSD delivered via videoconferencing technology. *Behavior Modification*, 31(6), 856-866.
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., & Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Mental Health*, 5(4), e64.
- Germain, V., Marchand, A., Bouchard, S., Guay, S., & Drouin, M. S. (2010). Assessment of the therapeutic alliance in face-to-face or videoconference treatment for posttraumatic stress disorder. *CyberPsychology, Behavior, and Social Networking*, 13(1), 29-35.
- Gollings, E. K., & Paxton, S. J. (2006). Comparison of internet and face-to-face delivery of a group body image and disordered eating intervention for women: a pilot study. *Eating Disorders*, 14(1), 1-15.
- Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007, September). Creating rapport with virtual agents. In *International workshop on intelligent virtual agents* (pp. 125-138). Springer, Berlin, Heidelberg.
- Gulec, H., Moessner, M., Mezei, A., Kohls, E., Túry, F., & Bauer, S. (2011). Internet-based maintenance treatment for patients with eating disorders. *Professional Psychology: Research and Practice*, 42(6), 479-486.
- Hart, J., Gratch, J., & Marsella, S. (2013). How virtual reality training can win friends and influence people. *Human Factors in Defence. Ashgate*, 235-249.
- Hancock, B., Bordes, A., Mazare, P. E., & Weston, J. (2019). Learning from dialogue after deployment: Feed yourself, chatbot!. *arXiv preprint arXiv:1901.05415*. Retrieved from <https://arxiv.org/pdf/1901.05415v4.pdf>
- Himle, J. A., Fischer, D. J., Muroff, J. R., Van Etten, M. L., Lokers, L. M., Abelson, J. L., & Hanna, G. L. (2006). Videoconferencing-based cognitive-behavioral therapy for obsessive-compulsive disorder. *Behaviour Research and Therapy*, 44(12), 1821-1829.
- Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, 36(5), 427-440.
- Houston, T. K., Cooper, L. A., & Ford, D. E. (2002). Internet support groups for depression: A 1-year prospective cohort study. *American Journal of Psychiatry*, 159(12), 2062-2068.
- Imel, Z. E., Steyvers, M., & Atkins, D. C. (2015). Computational psychotherapy research: Scaling

- up the evaluation of patient-provider interactions. *Psychotherapy*, 52(1), 19-30.
- Josman, N., Somer, E., Reisberg, A., Weiss, P. L., Garcia-Palacios, A., & Hoffman, H. (2006). BusWorld: Designing a virtual environment for post-traumatic stress disorder in Israel: A protocol. *CyberPsychology & Behavior*, 9(2), 241-244.
- Kandalaf, M. R., Didehbandi, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders*, 43(1), 34-44.
- Katharine C. (2019, January 9), "How computer-assisted therapy helps patients and practitioners(pt. 1)", Retrieved from <https://www.apa.org/members/content/computer-assisted-therapy>.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Kolata, G. (2000). Web research transforms visit to the doctor. *New York Times*, 6.
- Lange, A., van de Ven, J. P., Schrieken, B., & Emmelkamp, P. M. (2001). Interapy. Treatment of posttraumatic stress through the internet: A controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*, 32(2), 73-90.
- Lederman, R., Wadley, G., Gleeson, J., Bendall, S., & Álvarez-Jiménez, M. (2014). Moderated online social therapy: Designing and evaluating technology for mental health. *ACM Transactions on Computer-Human Interaction*, 21(1), 1-26.
- Litz, B. T., Engel, C. C., Bryant, R. A., & Papa, A. (2007). A randomized, controlled proof-of-concept trial of an Internet-based, therapist-assisted self-management treatment for posttraumatic stress disorder. *American Journal of Psychiatry*, 164(11), 1676-1684.
- Lucas, G. M., Gratch, J., King, A., & Morency, L. P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94-100.
- Luxton, D. D. (2014). Recommendation for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence in Medicine*, 62(1), 1-10.
- Ly, K. H., Ly, A. M., & Andersson, G. (2017). A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interventions*, 10, 39-46.
- McFarlane, W. R. (2004). *Multifamily groups in the treatment of severe psychiatric disorders*. New York: Guilford Press.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, 27(4), 12-12.
- Miner, A. S., Milstein, A., Schueller, S., Hegde, R., Mangurian, C., & Linos, E. (2016). Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health.

- JAMA Internal Medicine*, 176(5), 619-625.
- Morgan, V. A., Waterreus, A., Jablensky, A., Mackinnon, A., McGrath, J. J., Carr, V., ... & Galletly, C. (2012). People living with psychotic illness in 2010: The second Australian national survey of psychosis. *Australian & New Zealand Journal of Psychiatry*, 46(8), 735-752.
- Morris, R. R., Kouddous, K., Kshirsagar, R., & Schueller, S. M. (2018). Towards an artificially empathic conversational agent for mental health applications: System design and user perceptions. *Journal of Medical Internet Research*, 20(6), e10148.
- Newman, M. G., Szkodny, L. E., Llera, S. J., & Przeworski, A. (2011). A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: Is human contact necessary for therapeutic efficacy?. *Clinical Psychology Review*, 31(1), 89-103.
- Nienhuis, J. B., Owen, J., Valentine, J. C., Winkeljohn Black, S., Halford, T. C., Parazak, S. E., ... & Hilsenroth, M. (2018). Therapeutic alliance, empathy, and genuineness in individual adult psychotherapy: A meta-analytic review. *Psychotherapy Research*, 28(4), 593-605.
- O'Keeffe, G. S., & Clarke-Pearson, K. (2011). The impact of social media on children, adolescents, and families. *Pediatrics*, 127(4), 800-804.
- Parsons, S. (2016). Authenticity in Virtual Reality for assessment and intervention in autism: A conceptual review. *Educational Research Review*, 19, 138-157.
- Parsons, S., & Cobb, S. (2014). Reflections on the role of the 'users': Challenges in a multi-disciplinary context of learner-centred design for children on the autism spectrum. *International Journal of Research & Method in Education*, 37(4), 421-441.
- Parsons, S., Leonard, A., & Mitchell, P. (2006). Virtual environments for social skills training: Comments from two adolescents with autistic spectrum disorder. *Computers & Education*, 47(2), 186-206.
- Pesämaa, L., Ebeling, H., Kuusimäki, M. L., Winblad, I., Isohanni, M., & Moilanen, I. (2004). Videoconferencing in child and adolescent telepsychiatry: A systematic review of the literature. *Journal of Telemedicine and Telecare*, 10(4), 187-192.
- Poushter, J. (2017). *Smartphones are common in advanced economies, but digital divides remain*. Washington, DC: Pew Research Center.
- Riley, W. T., Rivera, D. E., Atienza, A. A., Nilsen, W., Allison, S. M., & Mermelstein, R. (2011). Health behavior models in the age of mobile interventions: Are our theories up to the task?. *Translational Behavioral Medicine*, 1(1), 53-71.
- Ritterband, L. M., Gonder-Frederick, L. A., Cox, D. J., Clifton, A. D., West, R. W., & Borowitz, S. M. (2003). Internet interventions: In review, in use, and into the future. *Professional Psychology: Research and Practice*, 34(5), 527-534.
- Rizzo, A., Buckwalter, J. G., & Neumann, U. (1997). Virtual Reality and Cognitive

- Rehabilitation: A Brief review of the future. *Journal of Head Trauma Rehabilitation*, 12(6), 1-15.
- Ross, J., Zaldivar, A., Irani, L., & Tomlinson, B. (2009). Who are the turkers? worker demographics in amazon mechanical turk. Department of Informatics, University of California, Irvine, USA, Tech. Rep.
- Rothbaum, B. O., Hodges, L., Alarcon, R., Ready, D., Shahar, F., Graap, K., ... & Baltzell, D. (1999). Virtual reality exposure therapy for PTSD Vietnam veterans: A case study. *Journal of Traumatic Stress*, 12(2), 263-271.
- Rothbaum, B. O., Hodges, L. F., Ready, D., Graap, K., & Alarcon, R. D. (2001). Virtual reality exposure therapy for Vietnam veterans with posttraumatic stress disorder. *The Journal of Clinical Psychiatry*, 62(8), 617-622.
- Rothbaum, B. O., Hodges, L., Smith, S., Lee, J. H., & Price, L. (2000). A controlled study of virtual reality exposure therapy for the fear of flying. *Journal of Consulting and Clinical Psychology*, 68(6), 1020-1026.
- Seligman, M. E., Rashid, T., & Parks, A. C. (2006). Positive psychotherapy. *American Psychologist*, 61(8), 774-788.
- Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484-493.
- Spence, S. H., Holmes, J. M., March, S., & Lipp, O. V. (2006). The feasibility and outcome of clinic plus internet delivery of cognitive-behavior therapy for childhood anxiety. *Journal of Consulting and Clinical Psychology*, 74(3), 614-621.
- Tanana, M. J., Soma, C. S., Srikumar, V., Atkins, D. C., & Imel, Z. E. (2019). Development and evaluation of clientbot: Patient-like conversational agent to train basic counseling skills. *Journal of Medical Internet Research*, 21(7), e12529.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.
- Taylor, K., & Silver, L. (2019). Smartphone ownership is growing rapidly around the world, but not always equally. *Pew Research Center*.
- Teuscher, C., & Hofstadter, D. R. (2006). *Alan Turing: Life and legacy of a great thinker*. New York, NY: Springer.
- Torous, J., & Roberts, L. W. (2017). Needed innovation in digital health and smartphone applications for mental health: Transparency and trust. *JAMA Psychiatry*, 74(5), 437-438.
- Tuerk, P. W., Yoder, M., Ruggiero, K. J., Gros, D. F., & Acierno, R. (2010). A pilot study of prolonged exposure therapy for posttraumatic stress disorder delivered via telehealth technology. *Journal of Traumatic Stress*, 23(1), 116-123.
- Weisband, S., & Kiesler, S. (1996, April). Self disclosure on computer forms: Meta-analysis and implications. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3-10). ACM.
- Werts, M. G., Wolery, M., Holcombe, A., &

- Gast, D. L. (1995). Instructive feedback: Review of parameters and effects. *Journal of Behavioral Education, 5*(1), 55-75.
- Yuen, E. K., Goetter, E. M., Herbert, J. D., & Forman, E. M. (2012). Challenges and opportunities in internet-mediated telemental health. *Professional Psychology: Research and Practice, 43*(1), 1-8.
- Yuen, E. K., Herbert, J. D., Forman, E. M., Goetter, E. M., Juarascio, A. S., Rabin, S. J., & Bouchard, S. (2010, August). *Using Skype videoconferencing and Second Life virtual environments to deliver acceptance-based behavior therapy for social anxiety disorder*. In 44th Annual Convention of the Association for Behavioral and Cognitive Therapies, San Francisco, CA.
- Yuen, E. K., Herbert, J. D., Forman, E. M., Goetter, E. M., Comer, R., & Bradley, J. C. (2013). Treatment of social anxiety disorder using online virtual environments in second life. *Behavior Therapy, 44*(1), 51-61.

원 고 접 수 일 : 2019. 12. 05

수정원고접수일 : 2020. 02. 25

게 재 결 정 일 : 2020. 03. 13

The Application of Artificial Intelligence Technology in Counseling and Psychotherapy: Recent Foreign Cases

Doyoun Kim

Ajou University, Department of Psychology
/ Student

MinKi Cho

Ajou University, Department of Psychology
/ Professor

Heecheon Shin

This study examined existing attempts to utilize Information Technology (IT) in the field of counseling and psychotherapy, and current trends in recent overseas cases of the application of Artificial Intelligence (AI) technology in the field. While traditional online psychotherapies have focused on increasing accessibility and efficiency using IT technologies, recent studies have focused on the utilization of therapeutic interactions using chatbot and interactive agent, which can interact with clients independently, and the validation of the effects of such interventions. Accordingly, the current status of AI-based counseling and psychotherapy was reviewed through representative cases developed overseas. Technical issues related to the future prospects of AI counseling were reviewed, including natural language processing, mechanisms that enable interactive agents to express empathy, and the development of artificial intelligence that continues to develop without human assistance through self-learning processes. Finally, the significance and limitations of this study are presented.

Key words : Artificial Intelligence, Online-psychotherapy, Natural Language Processing, Conversational agent