

Generative AI, AI 휴먼 서비스

□ 한종호 / 딥브레인AI

요약

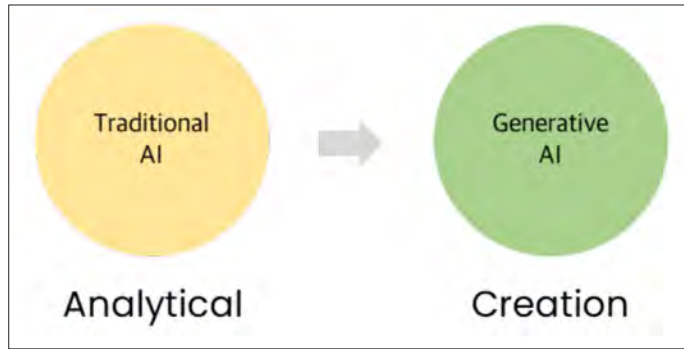
AI 분야는 기존 분석적(Analytical) AI에서 점차 자가학습을 통한 새로운 디지털 이미지, 영상, 음성, 텍스트, 코드 등을 만드는 Generative AI로 너무 빠르게 진화하고 Generative AI 영역을 두고 세계 각 기업들이 비즈니스의 우위를 선점하기 위해 개발 속도에 열을 올리고 있다. 이미 Generative AI는 다양한 부분에서 활용되고 있는데 음악, 문학, 미디어 등 새로운 창작물을 생성할 뿐만 아니라 향후 지식경제의 생산성을 획기적으로 향상시킬 것이다. 이런 Generative AI가 AI 휴먼 서비스 발전에 어떠한 영향을 미치는지에 대해 논하고자 한다.

I. 서론

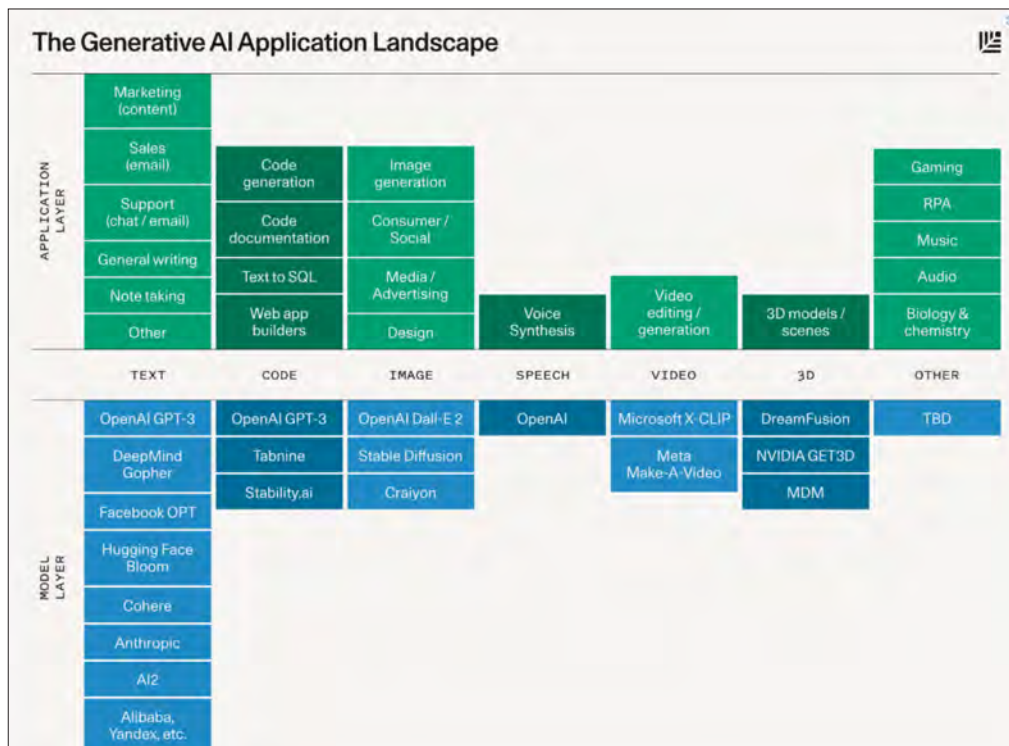
인공지능(AI, Artificial Intelligence)이라는 용어가 처음 등장한 때는 1956년에 미국 다트머스 대학교에서 마빈 민스키, 클로드 새넌 등 인공지능 및 정보 처리 이론에 지대한 공헌을 한 사람들이 개최한 학회에서 존 매카시가 이 용어를 사용하면서부터이지만 본격적으로 AI의 기술이 발전하게 된 시점은 GPGPU 공개 이후 2010년부터 Deep Learning 기반 인공지능 기술이 비약적으로 발전하

면서부터다. 딥러닝에 필요한 대량의 input data를 효율적으로 처리가 가능해지면서 발전하면서부터다. 이에 각종 연구 논문이 폭발적으로 늘어나게 되었고 Generative AI 관련 기술이 점점 고도화되면서 사람 고유의 영역이라고 생각했던 새로운 콘텐츠 창조, 창작마저도 AI가 할 수 있는 일이 되어가고 있다.

지금까지의 AI는 주어지는 데이터를 바탕으로 다음을 예측하거나 추천을 해 주는 AI 서비스가 다양한 산업에 반영되었다. 구글의 검색엔진이나 개인에 맞는 상품을 추천



<그림 1> AI 변화

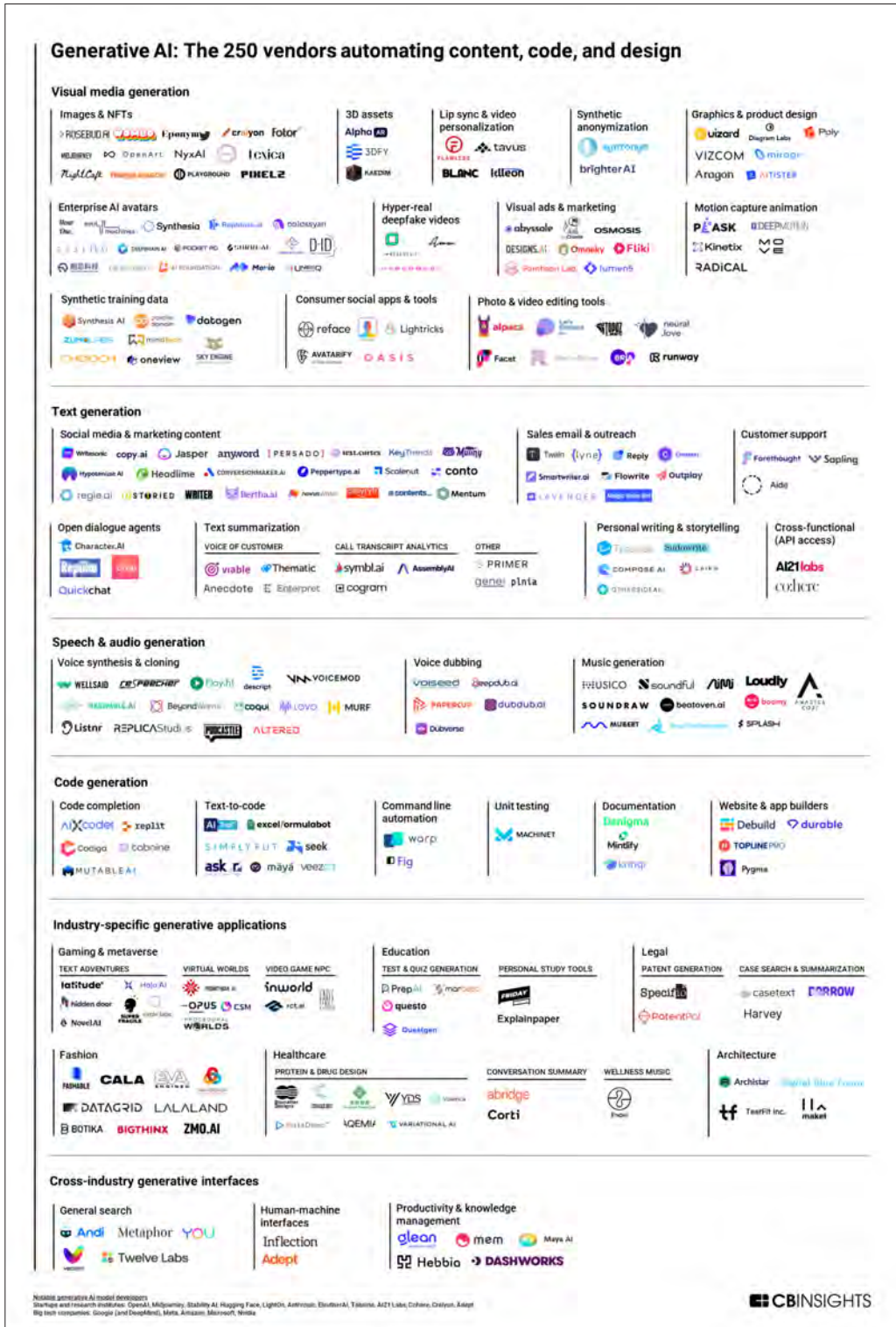


<그림 2> Generative AI Application Landscape

해 주는 금융서비스 등에서 쉽게 접해본 적이 있을 것이다. 2021년까지만 해도 AI는 단순·반복되는 업무처리나 많은 양의 데이터를 빠른 시간 내에 분석하는데 사용되는 등 사람을 보조하는 역할로 활용되었지, 사람의 ‘창작’ 행위까지 할 수 있을 것이라고 상상하기 어려웠다. 하지만 Generative AI는 몇 가지 키워드만으로 소설이나 논문을

작성하고, 텍스트 입력만으로 그림을 그리고 음성과 영상을 제작할 수 있는 것이다.

위 그림에서 보는 바와 같이 현재 다양한 포맷의 애플리케이션 영역에서 Generative AI 모델들이 등장했다. 대표적인 모델들로는 세계적으로 관심을 받고 있는 자연어 처리 및 대화가 가능한 GPT-3.5(ChatGPT)에 이어 GPT-4



<그림 3> 2023 Notable Generative AI model developers

가 출시되었다. GPT-4는 고급추론과 이미지와 글을 동시에 다룰 수 있는 멀티모달 모델로 이제 텍스트뿐만 아니라 이미지도 함께 제공해 맥락을 이해하고 창의적인 답변도 가능하다고 한다. 또 다른 모델인 Stable Diffusion이나 Dall-E2 모델은 텍스트를 자연어로 입력해도 그 문맥을 이해하여 이미지를 생성해 준다. 이처럼 Generative AI는 게임, 예술, 소설, 음성, 영상, SNS/비디오 등 이미 여러 분야에 현실로 나타나 적용되고 있고 대규모 투자 유치나 뚜렷한 사업 성과로 이어지는 케이스도 존재한다. 특히나 2022년 투자시장은 전 세계적으로 소극적인 투자가 이루어졌지만 Generative AI 영역에서의 투자는 OpenAI나 Jasper 등 유니콘 클럽이 탄생할 만큼 적

극적이었다.

글로벌 시장조사기관 CB인사이트가 발표한 '글로벌 250대 Generative AI 스타트업'에 따르면 국내에서는 3개의 한국 기업이 고성장 Generative AI 스타트업에 선정되었다. 이번 조사는 크게 시각미디어, 텍스트, 소리, 코드 생성, 기타산업, 산업융합 등 6개 분야에서 45개 세부항목으로 생성 AI의 활용 범위를 나누고 분야별로 주목받는 스타트업들을 선정했다. 선정된 국내 3개 기업은 모두 가상인간 등 시각미디어 분야로 이는 국내 시각미디어 분야의 기술 수준이 검증되었다는 것을 뜻한다. 특히 AI 휴먼 시장은 글로벌 시장조사업체 이머전리서치에서 글로벌 버추얼 휴먼 시장 규모를 2020년 100억 달러(약 12조7500억 원)에서 2030년 5275억8000만 달러(700조 원)로 10년 동안 약 53배 급성장할 것이라고 전망하기도 했다.



<그림 4> 글로벌 250대 Generative AI 스타트업 지역분포도

II. AI 휴먼 서비스

AI 휴먼은 실제 촬영한 원본 데이터의 음성과 영상을 머신러닝 기술인 GAN(Generative Adversarial Network)을 활용하여 실제 인간과 비슷한 가상의 인간을 제작하는 기술이다. AI 휴먼은 실제 인간을 모델로 하거나 여러 장의 사진을 조합하여 만든 가상 얼굴을 이용하기도 하며 2D, 3D, 캐릭터 등의 AI 휴먼이 있다. AI 휴먼 제작에 소요되는 시간은 제조사마다 제조하는 기술과 방식에 따라 차이가 나겠지만 몇 분부터 몇 주가 걸리기도 한다. AI 휴먼을 만드는 방법으로는 널리 알려진 기술인 3D VFX 기술을 활용하거나 앞서 말한 AI 기술인 GAN을 활용한다. 기술별로 장단점이 있기 때문에 어떤 기술이 우월하다고 보기는 어렵고 사용하는 영역별로 적절한 기술을 사용하고 있다. 엔터테인먼트 영역에서 주로 등장하는 AI 휴먼은 GAN보다는 VFX 기술을 활용하여 만들어지고 있다. VFX는 매우 정교한 애니메이션을 만드는 것과 같이 대역 모델의 촬영 프레임을 일일이 작업을 해야 하므로 자유로운 움직임, 표정, 제스처 등을 표현할 수 있지만 실시간 대

용이 어렵고 콘텐츠 제작에 많은 비용과 시간이 필요하다. GAN을 활용한 AI 휴먼은 초기 개발에 일정한 시간과 비용이 필요하지만, 일단 개발 후에는 수많은 콘텐츠를 생성해내는데 매우 합리적인 비용과 실시간에 가까운 속도로 빠르게 진행할 수 있다. AI 기술을 활용한 AI 휴먼은 금융, 미디어, 교육, 유통, 고객 상담 등의 다양한 영역에서 활용되고 있다.

1. AI 휴먼 제작 절차

AI 휴먼 서비스는 AI 휴먼을 제작하는 과정과 제작한 AI 휴먼을 서비스하는 과정으로 구분할 수 있다. AI 휴먼을 제작하는 여러 방법 중에서 가장 보편적이고 많은 기업이 사용하고 있는 GAN을 이용한 AI 휴먼 제작 절차에 대해 설명하고자 한다. GAN의 기술에 대한 설명은 이미 많은 참고 자료가 존재하므로 기술적인 설명보다는 AI 휴먼을 제작하는 과정과 GAN이 활용되는 시점 등에 대해 설명하겠다.

AI 휴먼을 제작하는 기업마다 방식에 차이는 있을 수 있지만 대부분 AI 휴먼을 제작하는 절차는 크게 4단계로 구분된다.

1) 모델 촬영 : AI 휴먼을 제작하기 위해선 AI 휴먼으로

만들 대역으로부터 딥러닝의 Input 데이터값을 확보해야 되므로 대역의 영상, 음성, 제스처 등을 촬영한다. 촬영 시 딥러닝에 방해되는, 예를 들어 반짝이는 액세서리 등의 요소들은 최소화하고 딥러닝에 필요한 충분한 양질의 데이터를 촬영한다. 물론 대역의 음성 이외에 다른 사람의 음성을 이용하기도 한다.

- 2) 데이터 전처리 : 촬영된 Input Data의 손실을 최소화하기 위해 음성과 영상 파일을 분리하여 전처리를 진행한다. 음성의 주파수 대역폭을 맞춰주는 작업 등이 진행된다.
- 3) 영상/음성 학습 : 전처리된 음성/영상 파일을 Input 값으로 딥러닝 학습을 진행한다. 이때 GAN을 활용하여 AI 휴먼을 제작한다. 비디오/오디오/제스처와 립싱크에 대한 딥러닝을 진행한다.
- 4) AI 모델 구현 : 딥러닝이 완료된 AI 휴먼을 텍스트 입력만으로 음성과 영상으로 합성되도록 솔루션에 탑재하거나 SDK/API로 제작하여 다양하게 서비스할 수 있도록 AI 휴먼을 모듈화한다.

AI 휴먼 기업들의 기술적인 노하우와 제작 방식에 따라 AI 휴먼의 품질(립싱크/해상도/제스처/TTS/성능)은 많은 차이가 날 수 있지만 AI 휴먼을 제작하는 과정은 비슷할 것이다.



<그림 5> AI 휴먼 제작 절차

2. AI 휴먼 서비스 유형

AI 휴먼 기술력은 유사하게 일정 수준 올라온 상황이다. 인간과 유사한 AI 휴먼을 제작할 수 있고 제작된 AI 휴먼의 품질, 립싱크, 합성 속도, 자연스러운 제스처, 다양한 언어 지원 등은 해외 기업뿐만 아니라 국내 기업들도 우열을 가리기가 힘들다. 즉 이제는 AI 휴먼을 어떻게든 활용하여 수익화할 서비스를 만들어 내는 것이 핵심이다. 기존에는 텍스트나 그래프 등을 통한 시각 위주의 정보 전달에서 이제는 AI 휴먼을 통한 시청각 정보 전달이 가능해진 것이다. 서비스 유형은 B2C 서비스, B2B 서비스로 구분할 수 있고 서비스 형태로는 구독 서비스와 구축 서비스로 분리할 수 있다. AI 휴먼을 적용한 단방향 콘텐츠를 손쉽게 제작하는 서비스와 AI 휴먼과 일반 사용자 간 양방향으로 Interaction이 가능한 서비스를 전개할 수도 있다. 해외나 국내 AI 휴먼 기업들은 이미 텍스트만 입력하면 AI 휴먼의 음성/영상 합성 서비스를 상용화하여 구독 서비스나 구축 서비스로 비즈니스를 전개하고 있다.

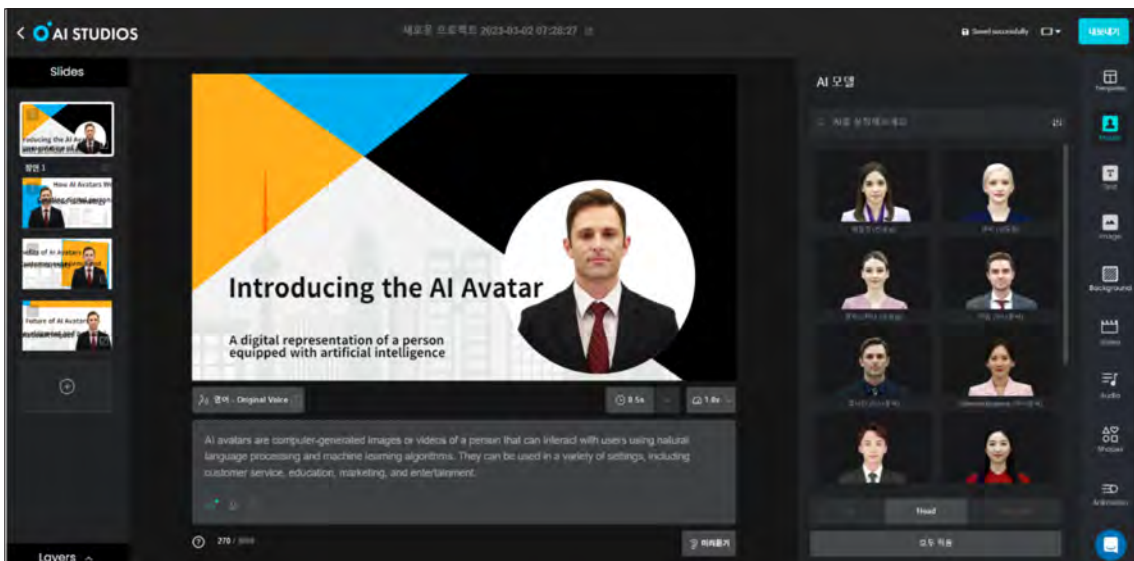
단방향 서비스를 위한 솔루션을 보유하고 있는 대표

적인 기업으로는 미국의 Synthesia, Soul Machines, 중국의 Iflytek, 한국의 딥브레인AI, 이스트소프트, 클레온 등이 이미 텍스트만 입력하면 AI 휴먼 영상을 합성할 수 있는 솔루션을 보유하고 있다. Generative AI의 다양한 활용 분야 중에서 Text To Video, Text to Speech, Text to 3D, ChatGPT to Text to Video 생성 등 많은 Application Layer에 활용되고 있는 것이다.

AI 휴먼을 활용한 단방향 서비스의 대표적인 사례들을 보면 다음과 같다.

- 1) 미디어 : 이미 많은 방송국의 주요 메인 앵커를 AI 휴먼으로 제작하여 실제 방송에서 활용하고 있다. AI 아나운서는 실제 아나운서보다 뉴스 제작 비용이 저렴하다. 뉴스 제작 시 들어가는 비용 항목 중 헤어, 의상, 메이크업, 기사, 촬영, 편집 등보다 실제 기사 작성과 편집만 하면 뉴스를 제작할 수 있다. 또한 시간과 장소에 구애받지 않으므로 속보 제작에 특화되어 있다.

자연스러운 AI 휴먼의 등장은 단방향 단순 콘텐츠 제작뿐만 아니라 고객과의 양방향 소통까지 가능하



<그림 6> ChatGPT to Text to Video 솔루션 예시



<그림 7> AI 아나운서 활용 사례

로 많은 서비스에 적용이 가능하다. 특히 COVID-19 이후에 많은 대면 활동이 비대면 활동으로 이동되었다. 기존 대면 활동의 주체인 ‘사람’의 일을 AI 휴먼의 등장으로 고객을 접하는 주체가 반드시 실제 ‘사람’일 필요가 없어진 것이다. AI 휴먼은 기업이 전달하고자 하는 메시지를 그 무엇보다 친숙하면서도 효율적으로 전달하고 있다. TV 광고 영상, 유명 인플루언서, K팝 아이돌, AI 은행원, AI 애널리스트, AI 아나운서, AI 튜터 등 고객과 친숙하면서 효율적으로 정보를 전달할 수 있다.

2) 엔터테인먼트 : 유명 인플루언서, 광고모델, 가수로 활동하고 있는 AI 휴먼이 계속해서 등장하고 있다. 이미 유명 “인싸”인 신한라이프의 ‘오로지’, SKT의 ‘나수아’, 20대의 윤여정, 가상 인간 가수 ‘이터니티’ 등 실제 연예인과 같은 팬덤을 형성하고 다양한 방

법으로 팬들과 의사소통을 하고 있다. 장점으로는 IP에 따른 재계약이 필요 없고, 아프지 않으며, 사건 사고도 없다. 또한 늙지도 않고 24시간 일을 할 수도 있다.

3) 교육 : 유명 강사를 AI 휴먼으로 만들어서 AI 휴먼과 실시간으로 소통하며 교육 커리큘럼을 만들어 간다.

4) 선거 : 2022년 대한민국 대통령 선거에서 대통령 후보를 AI 휴먼으로 만들어서 선거에 적극 활용했고 이에 많은 정치인들이 AI 휴먼 제작에 관심을 보이고 있다. AI 정치인을 활용한 선거 공약은 시공간에 제약을 받지 않고 지역별, 연령별, 성별 등 세밀하게 제시되었다.

5) 라이브 커머스 : 네이버나 틱톡 등의 커머스 플랫폼과 연계하여 AI 휴먼이 라이브로 커머스 활동도 진행하고 있다. 하지만 아직까지 고객과 양방향 소통을



<그림 8> AI 대통령 후보 사례

통한 라이브 커머스는 시기상조인 듯 하다.

- 6) 기타 : 그밖에 대표적인 단방향 AI 휴먼 서비스로는 금융권의 AI 애널리스트, AI 보험 설계사, AI 금융 비서 등이 있으며 기업의 홍보 영상 제작, 교육 영상 제작 등에 활용되고 있다.

AI 휴먼의 양방향 서비스는 해외에서는 찾아보기 힘든 케이스로 국내가 더 앞서 있다고 생각된다. AI 휴먼을 활용한 양방향 서비스의 대표적인 국내 사례로는 금융권이 가장 앞서 있다. KB은행, 신한은행, NH농협은행에서 키오스크에 AI 은행원을 탑재하여 각 매장에서 금융 상품을 소개하거나 고객과 대화를 통한 은행 업무 지원, 은행 창구 대기표 발급 등의 서비스를 진행하고 있다. 금융권 이외에는 무인점포에 AI 안내원이 매장의 상품 위치와 가격, 행사 제품 등을 고객의 질문에 답변하고 결제까지 진행해 준다. 이렇게 양방향 서비스가 가능하기 위해서는 AI 디바이스 + 챗봇 + AI 휴먼 성능(실시간 합성)이 갖춰져야 한다.

- 1) AI 디바이스 연계 : 스피커 모듈과 마이크 모듈로 입력되는 모든 음성/잡음/에코 신호를 최적의 음성 신호로 처리해 주는 음성 신호 전처리 개발 및 LM 학습을 적용하여 인식률이 높은 텍스트로 변경하는 STT를 통해서 음성이 텍스트로 변경되고 변경된 텍스트를 ChatGPT나 챗봇과 I/F를 통하여 전달하고 그 결과값을 텍스트로 받아 음성과 영상으로 합성한다. 합성할 때 텍스트를 음성으로 변경하는 TTS가 사용된다.
- 2) 챗봇 : 고객과 AI 휴먼 간 자연스러운 대화를 하기 위해선 고객 질문에 대한 답변이 필요하다. 고객이 어떤 질문을 할지, 적용해야 할 도메인이 무엇인지에 대한 말뭉치가 준비되어야 한다. ChatGPT처럼 다방면에서 결과를 가져올 수도 있지만 초거대 AI나 챗봇과 연계하여 금융, 호텔, 골프, 무인점포, 관광지 안내 등 특정 도메인에 대한 언어모델을 개발하고 대응할 수도 있다. 자연어 처리, 문맥 이해, 멀티턴 등이 반영된 챗봇이라면 실제 사람과 대면으로 대화하듯 자연스러움이 극대화될 수 있을 것이다. ChatGPT의



<그림 9> AI 키오스크를 이용한 양방향 서비스

등장으로 AI 휴먼 시장이 급성장할 것으로 기대되는 이유도 ChatGPT는 사실 관계에 관한 질문, 배경 지식이 필요하고 그걸 종합해서 추론하는 것이 필요한 질문 등 고차원적인 사고를 필요로 하는 질문에 매우 완성도 높은 답변을 주기 때문이다.

- 3) AI 휴먼 성능(실시간 합성) : 실제 고객과 원활하게 양방향 서비스를 하기 위해선 실시간으로 질의에 대한 답변을 AI 휴먼과 주고받아야 한다. 영상 합성 기술에 대한 엔진 성능, 하드웨어 사양 등 실시간 합성에 영향을 미친다. AI 휴먼에 대한 개발 과정이나 실시간 서비스를 하기 위해서는 GPU 서버가 사용되는데 동시 접속자 수가 많으면 GPU가 1000대를 갖추고 있어도 부족하다. 보통 금융권 대고객 서비스는 동시 접속자 수가 몇만 명일 수도 있는데 그렇다고 비싼 GPU를 그만큼 구매할 수도 없다. 이런 문제를 해결하기 위하여 앞서 언급한 도메인별 말뭉치를 미리 합성하여 캐시 서버에 올려두고 질문에 대한 답변이 있을 경우 캐시 서버를 통해 서비스된다. 물론 미리 합성된 AI 휴먼 영상이 없다면 챗봇에서 머신러닝을 통해 답변이 추가되고 추가된 답변을 텍스트로 받아서 합성한 이후에 다시 캐시 서버로 올려서 서비

스에 적용할 수 있다.

III. 결론

지금까지 AI 휴먼의 제작 과정과 AI 휴먼을 통한 비즈니스 유형에 대해 살펴보았다. 현재의 AI 휴먼 기술은 다양한 제스처와 감정 표현 등의 자연스러움과 주위 상황 인지와 상호작용까지는 구현이 가능하다. 이와 더불어 ChatGPT나 GPT-4와 같이 AI 휴먼의 두뇌를 담당할 Generative AI 모델이 빠르게 발전하여 인간과 이야기하는 것과 비슷한 수준까지 발전하고 있다. 이러한 자연어 처리 모델과 감정과 제스처까지 담아내는 음성/영상 합성 기술이 합쳐지면 사용자와 자연스러운 Interaction이 가능한 AI 휴먼이 등장할 것으로 기대된다. 과거 AI 휴먼이 “불쾌한 골짜기”였다면 지금 시대의 AI 휴먼은 “유패한 골짜기”가 될 수 있을 것이다. 이처럼 AI 휴먼의 기술 발전은 해외·국내 기업에 많은 기회를 제공할 것이고 어떤 Killer Service를 준비하느냐가 향후 기업의 미래를 결정할 것이다.

참 고 문 헌

- [1] <https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/>
- [2] <https://news.mt.co.kr/mtview.php?no=2023022118002957544&type=1>
- [3] <https://www.asiae.co.kr/article/2022090806340210950>
- [4] <https://news.mt.co.kr/mtview.php?no=2023021711092158020>
- [5] 인공지능 : AI 윤석열, 이재명 챗봇... 커지는 대선판 AI경쟁, <https://www.bbc.com/korean/news-59949885>
- [6] <http://www.mediatoday.co.kr/news/articleView.html?idxno=302369>
- [7] SPRI - 소프트웨어정책연구소, <https://spri.kr/>
- [8] CBINSIGHTS, <https://www.cbinsights.com/>
- [9] NIA 한국정보화진흥원 [IT&Future_Strategy_2020-6]_AI_스타트업_생태계_혁신을_위한_정책방향.pdf
- [10] 딥브레인AI의 AI Studios, <https://www.deepbrain.io/ko/aistudios>

저 자 소 개



한 종 호

- 2000년 : 전주대학교 산업공학과 전공 학사
- 2012년 ~ 2021년 : 주식회사 핑거 본부장
- 2021년 ~ 현재 : 딥브레인AI 부사장
- 2023년 ~ 현재 : 서울대학교 AICEO 수료 중
- 주관심분야 : AI 휴먼, 메타버스, 마케팅, Generative AI