creative commons
COMMONS DEED

저작자표시 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이차적 저작물을 작성할 수 있습니다.
- 이 저작물을 영리 목적으로 이용할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

연세대학교
YONSEI UNIVERSITY

생성형 인공지능 의료기기에 대한 주요 규제기관의 규제 동향 조사 및 대응전략

연세대학교 대학원

의료기기산업학과

김 보 람

연세대학교
YONSEI UNIVERSITY

생성형 인공지능 의료기기에 대한 주요 규제기관의 규제 동향 조사 및 대응전략

지도교수 구 성 욱 · 장 원 석

이 논문을 석사 학위논문으로 제출함

2024년 12월

연세대학교 대학원

의료기기산업학과

김 보 람

연세대학교
YONSEI UNIVERSITY

김보람의 석사 학위논문으로 인준함

심사위원 구성욱 (인)

심사위원 [illegible] (인)

심사위원 강진성 (인)

연세대학교 대학원

2024년 12월

연세대학교
YONSEI UNIVERSITY

차 례

차례 · i
그림 차례 · iii
표 차례 · iv
국문 요약 · v
제1장 서론 · 1
1.1. 연구의 배경 · 1
1.2. 연구의 목적 · 3
1.3. 연구 범위 및 방법 · 4
제2장 생성형 인공지능의 개요 · · · · · · · · · · · · · · · · · · · 5
2.1. 인공지능 기술 · 5
2.2 인공지능 기술 분류 · 6
2.3 생성형 인공지능 기술 · 8
2.4 생성형 인공지능 서비스 · · · · · · · · · · · · · · · · · · · 11
2.5. 생성형 인공지능의 의료분야 적용 · · · · · · · · · · · · · · · · 13
2.5.1. 적대적 생성 신경망 · · · · · 13
2.5.2 대규모 언어모델(LLM)과 대규모 멀티모달 모델(LMM) · · · · · 15
2.6. 생성형 인공지능 의료기기 개발 현황 · · · · · · · · · · · · · · · 17
제3장 생성형 인공지능 의료기기 관련 규제 동향 · · · · · · · · · · · · · 18
3.1. 주요 기관별 규제 동향 · 18
3.1.1. 한국 · 18
3.1.2. 미국 · 23
3.1.3. 영국 · 30
3.1.4. 유럽 · 32
3.1.5. 호주 · 40
3.1.6. 국제의료기기규제포럼(IMDRF) · · · · · · · · · · · · · · · 40

연세대학교
YONSEI UNIVERSITY

제4장 고찰 · 41
제5장 결론 · 47
참고 문헌 · 49
영문 요약 · 52

연세대학교
YONSEI UNIVERSITY

그림 차례

<그림 1> 헬스케어 시장에서의 생성형 인공지능 시장 성장 전망 ····· 1
<그림 2> 인공지능, 머신러닝, 인공신경망, 딥러닝, 자연어처리 관계도 ·· 6
<그림 3> 기존 인공지능과 생성형 인공지능의 차이점 ·· 8
<그림 4> 식의약 분야에 생성형 인공지능을 도입 적용할 수 있는 분야 예시 ·· 11
<그림 5> 최근 몇 년 동안 기존 대규모 언어 모델의 타임라인 ····· 12
<그림 6> 합성 PET 이미지 생성 모델 예시 ·· 14
<그림 7> Pre-Cert Program Approach throughout TPLC ····· 22
<그림 8> AI Lifecycle Considerations ·· 24
<그림 9> EU AI 법안의 위험 기반 접근 방식의 도식화 ····· 36

연세대학교
YONSEI UNIVERSITY

표 차례

<표 1> 파운데이션 모델 학습 방법 ····· 10
<표 2> 국내외 생성형 AI 관련 주요 서비스 ····· 12
<표 3> 디지털의료기기, 디지털 융합 의약품 및 디지털 의료·건강지원기기 정의 · 19
<표 4> '빅테크 기업 규제혁신 프로그램' 기 발굴 규제개선 과제 ···· 23
<표 5> 소프트웨어 기능별 미국 FDA의 규제적 접근 방식 구분 ···· 26
<표 6> AI 규제 활용을 위한 5가지 핵심 원칙 ····· 31
<표 7> 범용 AI모델 공급자의 의무 ····· 34
<표 8> AI 시스템 모델 공급자의 추가적인 의무사항 ····· 36
<표 9> EU AI 법안의 위험 기반구분 및 주요내용 ····· 37
<표 10> 대표적인 생성형 인공지능 언어모델의 평가지표 예시 ····· 43
<표 11> 생성형 인공지능을 활용한 의료기기 개발 시 고려사항 및 대응전략 · 45

연세대학교
YONSEI UNIVERSITY

국 문 요 약

생성형 인공지능 의료기기에 대한 주요 규제기관의 규제 동향 조사 및 대응전략

본 연구는 의료분야에 적용되는 생성형 인공지능 의료기기에 대한 주요 규제기관의 규제 동향을 조사하고 이를 분석하여 규제 설계 시 고려해야 할 주요 요소를 도출하는 데 목적을 두고 있다.

생성형 인공지능은 방대한 데이터를 기반으로 새로운 정보를 생성하거나 복잡한 문제를 해결할 수 있는 특징을 지니며, 진단 보조, 치료 계획 수립, 의료 교육 및 환자 데이터 분석 등 다양한 의료분야에서 혁신적인 역할을 수행할 수 있다.

그러나 이러한 기술의 고유 특성으로 인해 기존 의료기기 규제 체계만으로는 안전성과 신뢰성을 보장하기에 한계가 있으며, 새로운 규제 프레임워크의 필요성이 제기되고 있다.

생성형 인공지능 의료기기가 가진 특성과 이를 둘러싼 규제상의 도전 과제를 고찰하고, 미국, 유럽, 한국 등 주요 국가 및 국제 규제 조화기구(IMDRF)에서 제시하고 있는 생성형 인공지능 관련 규제 동향을 심층적으로 분석하였다.

주요 규제기관들은 생성형 인공지능의 잠재적인 위험 요소인 환각(hallucination), 데이터 편향 및 출력 불확실성을 완화하기 위해 위험 기반 접근법, 투명성 강화, 실사용 데이터(RWD/RWE) 활용, 성능 검증 및 사후 모니터링 강화 등을 강조하고 있다.

또한, 각국은 생성형 인공지능 의료기기의 안전성과 유효성을 보장하기 위해 별도의 인공지능 법안을 제정하거나 기존 법률을 보완하고 있으며, 한국 식약처는 세계 최초로 생성형 인공지능 의료기기에 특화된 허가·심사 가이드라인 제정을 준비하고 있다.

본 연구를 통해 생성형 인공지능 의료기기의 규제와 관련된 주요 고려사항을 사전에 파악하고 대응전략을 제시함으로써 제조사들의 효율적인 제품 개발과 인허가 과정을 지원하는 데 기여하고자 한다.

핵심되는 말: 인공지능, 생성형 인공지능, GenAI-enabled medical device

연세대학교
YONSEI UNIVERSITY

1. 서론

1.1. 연구의 배경

인공지능(AI) 기술의 급격한 발전으로 의료기기 산업에서 다양한 형태의 의료기기가 등장하고 있다. 특히, 미국 스타트업'오픈AI'가 출시한 챗GPT(Chat Generative Pre-trained Transformer)는 사람의 언어를 이해하고, 사람의 질문에 꼭 맞는 유용한 답을 하는 대화형·생성형AI로, 일반 대중에게AI의 일상화라는 경험을 제공하고 있으며 이를 의료현장에 접목하려는 연구가 활발히 진행되고 있다.[1)]

글로벌 시장조사 기관인 Precedence Research에 따르면 의료분야의 글로벌 생성 AI 시장 규모는 2022년에 10억 7천만 달러에서, 2032년까지 약 217억 4천만 달러를 돌파할 것으로 예상되며, 2023년에서 2032년까지의 기간 동안 35.14%의 CAGR로 성장할 것으로 예상 된다.[2)]

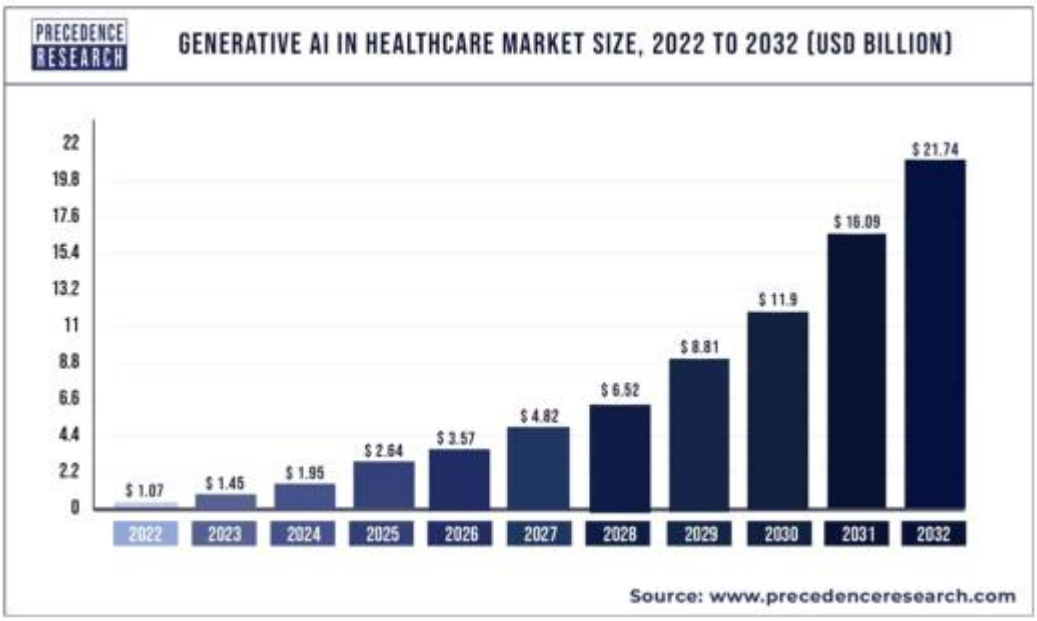

그림 1. 헬스케어 시장에서의 생성형 인공지능 시장 성장 전망[2)]

연세대학교
YONSEI UNIVERSITY

생성형 인공지능은 기존 데이터를 바탕으로 새로운 데이터를 생성하거나, 기존 데이터의 패턴을 학습하여 혁신적인 의료 솔루션을 제공할 수 있는 기술로, 진단, 치료, 그리고 환자 모니터링 분야 등 다양한 의료분야에 적용될 수 있다. 특히, 초거대 언어모델(LLM)을 기반으로 데이터를 스스로 생성하는 생성형AI가 주목을 받고 있으며 마이크로소프트, 구글, 아마존 등 글로벌 빅테크 업체들이 생성형 인공지능 기술이 적용된 서비스를 의료현장에 공급하는 것을 준비하고 있다. 국내에서는 카카오브레인이 초거대AI를 활용한 의료 영상 판독 서비스(KARA CXR)를 개발하고 있으며 이 제품은 각종 의료 데이터를 학습한 AI 모델로 흉부 엑스레이 사진에 대해 초안 판독문을 자동 생성해 의사가 빠르게 최종 판독을 내릴 수 있도록 돕는다.

그러나 이러한 생성형 인공지능 의료기기의 개발과 확산 속도에 비해, 이를 규제할 수 있는 법적, 제도적 장치는 아직 미비한 상황이다. 기존의 의료기기 규제 체계는 전통적인 의료기기를 대상으로 설계되었기 때문에 생성형 인공지능 기술이 가진 특성을 충분히 반영하지 못하고 있어 그대로 적용하기에는 한계가 있다. 실제로 현재까지 생성형 인공지능 의료기기에 대한 규제 가이드라인 및 기준은 아직 확립되지 않은 상황이며, 우리나라 식약처에서는 생성형 인공지능의 특성을 반영한 의료기기의 신속한 제품화를 지원하기 위한 ‘생성형AI 기반 의료기기 허가 심사 가이드라인’을 2025년 상반기에 제정할 계획에 있다.[3)]

의료기기 개발은 본질적으로 개발 과정에서 많은 시간과 비용이 소요되는 과정이다. 따라서 생성형 인공지능 의료기기에 적용될 규제 및 관리기준에 대해 고려를 하지 않은 상태로 제품을 개발하게 되면 인허가 과정에서 상당한 리스크를 갖게 된다.

따라서 본 연구에서는 생성형 인공지능 의료기기와 관련된 국내외 규제 현황 및 적용 가능한 관련 규제제도를 심층적으로 분석하여 현재 개발 중인 생성형 인공지능 기술이 적용된 의료기기의 개발 및 인허가와 관련된 고려사항들을 분석하고 이에 대한 대응전략을 제시하고자 한다.

연세대학교
YONSEI UNIVERSITY

1.2. 연구의 목적

본 연구의 목적은 생성형 인공지능 기술이 적용된 의료기기에 대한 규제 기준이 아직 명확히 마련되지 않은 상황에서, 이러한 기술이 실제로 의료기기로 분류될 수 있는 기준을 제시하고, 이를 바탕으로 예상되는 규제 대응 전략을 도출하는 데 있다. 특히, 국내외에서 생성형 인공지능 의료기기와 관련된 규제 현황을 분석함으로써 현재의 규제 동향에서 개발사들이 고려해야 할 주요 요소들을 파악하고 인허가 대응 전략 및 고려사항들을 제공하고자 한다.

생성형 인공지능 기술은 의료분야 전반에 걸쳐 적용될 수 있으나 서비스의 사용목적과 위해도에 따라 의료기기로의 관리가 필요하다. 일반 인공지능 기술을 적용한 서비스의 의료기기 판단 기준 및 의료기기 규제 고려사항은 식약처에서 발표한 '인공지능 의료기기의 허가□심사 가이드라인(2022년 5월)에서 정의가 되어 있지만 생성형 인공지능 기술의 특성에 대한 고려는 부재한 상황이다.

우리나라 식약처에서는 이러한 생성형 인공지능의 특성을 반영한 의료기기의 신속한 제품화를 지원하기 위한 '생성형AI 기반 의료기기 허가 심사 가이드라인'을 제정할 계획에 있다. 2024년 제25차 국제의료기기 규제기관포럼(IMDRF)에서도 생성형 인공지능 의료기기의 규제에 관한 논의가 이루어졌고, 현재 인공지능/머신러닝 워킹그룹에서 생성형 인공지능과 대규모 언어모델의 요구사항을 기존의 다른IMDRF 문서와 연결 및 일치하고자 하는 작업을 수행하고 있다고 언급 하였다.[4)]

본 연구는 생성형 인공지능 의료기기와 관련된 국내외 규제현황 및 적용 가능한 관련 규제제도 분석을 통해 의료기기 인허가를 위해 사전에 고려해야 할 사항 및 대응 전략을 제시하여, 생성형 인공지능 기술을 활용하여 의료기기를 제조하고자 하는 국내 업체들이 제품을 개발하는 단계에서부터 활용할 수 있도록 기여하고자 한다.

연세대학교
YONSEI UNIVERSITY

1.3. 연구 범위 및 방법

본 연구는 생성형 인공지능 의료기기의 특징을 명확히 정의하고 이에 대한 의료기기 판단 기준 및 사전 고려사항을 분석하고자 한다. 현재까지 생성형 인공지능 의료기기에 대한 명확한 규제기준 및 가이드라인이 부재함으로, 국내외 규제 동향을 파악함과 동시에 유럽 및 미국의 인공지능 규제제도에서 다루는 항목을 심층적으로 분석하여 생성형 인공지능 의료기기의 개발 단계부터 적용해야 하는 고려사항들을 제시하고자 한다.

(1) 생성형 인공지능의 정의 및 특성 분석

기존 의료기기 및 일반적인 인공지능 의료기기와 구분되는 생성형 인공지능 기술의 정의 및 특성을 정리한다. 특히, 대규모 언어 모델(LLM)과 같은 생성형 인공지능 기술이 의료기기에 적용될 때 발생할 수 있는 새로운 기술적·규제적 접근 방법을 도출한다.

(2) 국내외 규제 동향 분석

식품의약품안전처(MFDS), 미국 FDA, 유럽연합(EU), 영국 MHRA, 호주 TGA 등 주요 국가 및 국제의료기기규제포럼(IMDRF)의 규제 동향을 심층적으로 분석한다. 각국의 법률 및 가이드라인에서 다루는 생성형 인공지능 의료기기의 규제 적용 사례를 수집하고 공통되는 항목에 대해 요약한다. 또한 AI 법안과 같은 의료기기 규제와 별도의 규제체계를 조사하여 생성형 인공지능 의료기기에 적용되는 사항들을 분석한다.

(3) 규제 설계 및 고려사항 도출

생성형 인공지능 기술이 의료기기로 개발되고 규제 승인 과정을 거치는 데 있어 필요한 기술적, 임상적, 법적 고려사항을 정리하여 생성형 인공지능 의료기기를 개발하는 제조사에서 활용할 수 있도록 제시하고자 한다.

연세대학교
YONSEI UNIVERSITY

2. 생성형 인공지능의 개요

2.1. 인공지능 기술

인공지능(AI)은 의료 기술에 혁명을 일으켰으며 일반적으로 방대한 양의 데이터가 있지만 이론이 거의 없는 분야에서 많은 애플리케이션을 통해 복잡한 문제를 처리할 수 있는 컴퓨터 과학의 일부로 이해할 수 있다. 모바일 시대 이전에는 의료 기술이 주로 고전적인 의료 기기(예: 보철물, 스텐트, 임플란트)로 알려져 있었지만 스마트폰, 웨어러블, 센서 및 통신 시스템의 등장으로 매우 작은 크기에 인공지능(AI) 기반 도구(예: 애플리케이션)를 담을 수 있는 기능으로 의학에 혁명을 일으켰다.[5) 6)]

인공지능(AI)의 하위 분야인 머신 러닝은 수학적 접근 방식을 적용하여 경험을 통해 자동으로 개선되는 컴퓨터 알고리즘을 연구하는 것이며, 머신 러닝의 하위 집합인 딥 러닝은 생물학적 뇌의 뉴런을 모방하는 인공 신경망을 통해 입력 데이터를 처리하여 학습하는 알고리즘을 말한다.[6) 7)] 디지털 데이터의 폭발적 성장, 그래픽 처리 장치와 같은 하드웨어 기술 혁신으로 인한 컴퓨팅 파워 확장, 딥러닝을 통한 머신 러닝 알고리즘의 급속한 발전은 모두 헬스케어 분야에 상당한 영향을 미치고 있다. 또한 다양한 연구에 따르면 헬스케어에 인공지능(AI)을 적용하면 기존 기술보다 더 나은 결과를 얻을 수 있다는 결과가 있다. 이러한 연구에는 AI 기술을 이용하여 의료 영상을 분석하여 영상을 구별하고 치료에 사용하는 것, 다양한 의료 및 헬스케어 데이터를 통해 질병의 진행 과정을 예측하는 것, 치료나 진단 시 의사 결정을 지원할 수 있는 의료 기기를 개발하는 것, 의료 데이터를 암호화하는 것 등이 포함된다.[7)]

물리적 영역을 다루는 전형적인 기술과 달리 인공지능(AI)기술은 경험, 지능, 전문가의 판단과 같은 심리적 영역에 영향을 미치기 때문에 새로운 영역을 개척하고 있다. 특히 딥러닝 기술이 도입되면서 패턴 인식을 위한 머신러닝 알고리즘의 성능이 크게 향상되었고 AI 기술이 데이터 패턴을 분석하는 능력이 특정 작업(예: 이미지 인식 및 음성 인식)에 대한 평균적인 인간 능력과 유사해졌다. 딥러닝 알고리즘은 인간 뇌의 뉴런 네트워크를 닮은 인공 신경망을 기반으로 하며 매우 복잡한 비선형 관계를 학습할 수 있기 때문에 의료 데이터를 다루는 작업에 적극적으로 활용되고 있다.[10) 11)]

연세대학교
YONSEI UNIVERSITY

2.2 인공지능 기술 분류

인공지능(AI) 기술은 머신러닝(ML), 인공신경망(CNN), 딥러닝(DL), 자연어 처리(NLP), 언어모델(LM), 대규모 언어모델(LLM), 멀티모달 모델(MM) 등이 있으며 자세한 내용은 다음과 같다.[12)]

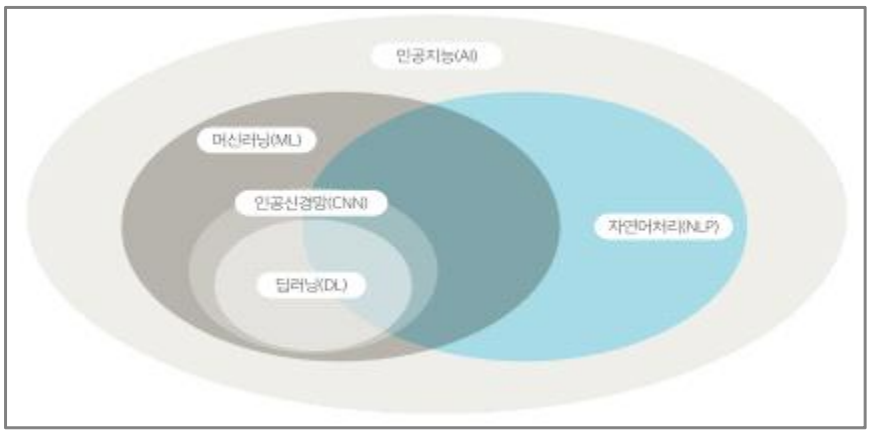

그림 2. 인공지능, 머신러닝, 인공신경망, 딥러닝, 자연어처리 관계도[12)]

- 머신러닝(Machine Learning, ML) : 컴퓨터가 스스로 데이터를 학습 및 예측 · 분류를 수행하여 인공지능의 성능을 향상시키는 기술이다. 학습데이터의 양과 질이 결과에 큰 영향을 미치며 주요 활용분야로는 이미지 및 음성 인식, 자연어 처리, 과거 데이터를 기반으로 한 결과 예측이 있다.

- 인공신경망(Artificial Neural Network)과 딥러닝(Deep Learning, DL) : 두 기술 모두 머신러닝의 일종으로 인간의 뇌신경세포(뉴런)의 컴퓨팅 구조를 모방한 기법인 인공신경망 방식으로 정보를 처리하며 인공신경망이 딥러닝보다 광의의 개념이며 대규모 데이터를 학습함으로써 성능을 향상 시킬 수 있다. 두 기술 모두 높은 정확도로 인해 현재 거의 모든 인공지능 분야에 사용되고 있으며, 이미지 데이터분석을 위한 CNN(Convolutional Neural Network),이미지 생성에 사용되는 GAN(Generative Adversarial Network), 자연어와 같은 시계열 데이터에 특화된 RNN(Recurrent Neural Network), ChatGPT의 기반 알고리즘으로서 자연어 처리 뿐만 아니라 이미지 분석에서도 널리 사용되는 트랜스포머(Transformer)등이 대표적이다.

연세대학교
YONSEI UNIVERSITY

- 자연어 처리(Natural Language Processing, NLP)와 언어모델(Language Model, LM) : 두 기술 모두 인공지능 하위 분야로 NLP는 텍스트 분류, 기계 번역, 질의응답분석 등 컴퓨터가 인간의 언어를 이해하고 처리하는데 초점을 둔 처리기술이다. LM은 NLP의 일종으로 인간의 언어를 학습하여 통계적 확률 및 인공신경망을 이용한 방법으로 가장 적절한 출력값을 출력하도록 만든 알고리즘 모델을 의미한다. 인간의 언어로부터 얻은 지식을 기반으로 텍스트와 다양한 콘텐츠를 인식하고 요약, 번역, 예측, 생성 작업을 할 수 있는 모델이다. 주요 LM에는 OpenAI의 GPT-4, 구글의 BERT, LaMDA-2, 메타의 LLaMA-3 등이 있다.

- 대규모 언어모델(Large Language Model, LLM) : LM의 일종으로 대규모 데이터 세트에서 훈련된 인공지능 언어모델을 의미한다. 주요 LLM으로는 OpenAI의 GPT-4, 구글의 BERT, LaMDA-2, 메타의 LLaMA-3, 네이버의 하이퍼클로바, 업스테이지의 솔라 등이 있으며 최근 LLM을 이용한 다양한 생성형 AI 서비스가 각광받고 있다.

- RAG(Retrieval Augmented Generation, 검색증강생성) 기술 : LLM은 방대한 데이터 기반으로 학습 하지만 간혹 정확한 정보를 제공하지 못하거나 일관된 답변을 얻지 못하는 한계점을 가지고 있다. RAG는 기존 LLM에 별도의 전문 도메인 지식 등을 연결함으로써, 해당 데이터를 기반으로 보다 정확하고 사실에 기반한 답변을 제공할 수 있도록 하는 새로운 자연어 처리 기술이다.

- 멀티모달 모델(Multimodal Model) : 텍스트, 이미지, 오디오, 비디오 등 다양한 유형의 데이터(모달리티)를 함께 고려하여 서로의 관계성을 학습 및 처리하는 기술로, 대규모 멀티모달 모델(Large Multimodal Model, LMM)의 예시로는 영상을 분석해 스크립트를 생성하는'구글'의 제미나이(Gemini), 이미지를 올리면 조리법을 분석해주는 'OpenAI'의 GPT-Vision 등이 있다.

연세대학교
YONSEI UNIVERSITY

2.3. 생성형 인공지능 기술

생성형 인공지능은 텍스트, 이미지, 음성 및 동영상 등 기존 콘텐츠를 활용한 학습 데이터를 기반으로 유사한 콘텐츠를 새롭게 만들어 내는 인공지능(AI)을 의미한다. 초거대 AI 등장으로 이를 기반으로 한 혁신적인 생성형 인공지능이 출시되고 있으며. 초거대 AI는 모델의 크기 측면에서 파라미터 수, 데이터의 규모와 학습 및 수행 능력을 강조하는 용어인데 비해, 생성형 인공지능은 이용자의 요청에 따라 새로운 것을 생성하는 기능을 강조하는 용어로 사용 된다.[13) 14)]

생성형 인공지능은 대규모 생성형 모형을 구축하고 엄청난 데이터를 학습시켜 실제 활용 가능한 수준의 창의적이고 고품질의 콘텐츠 생성이 가능하다. 자세하게는 생성형 인공지능의 구축을 위해서는 생성 능력을 학습하기 위한 특수한 모형 구조(예: GPT, BERT 등) 및 학습 알고리즘이 필요하다는 점, 대부분의 생성형 모형의 경우 수천억~수조 개의 파라미터로 구성되어 있으며 각 파라미터를 제대로 학습시키기 위해서는 대용량의 데이터가 필요하다는 특징이 있다.

기존의 인공지능도 결과물(output)을 생성하나 인공지능 기술의 발전으로 결과물의 창의성과 범용성이 실제 활용 가능한 수준(고화질 이미지, 영화 대본 등)에 도달함에 따라 '생성형 인공지능'이라는 용어가 탄생하였고 이는 기존의 인공지능 기술과의 가장 구분되는 차이점이다.[15)]

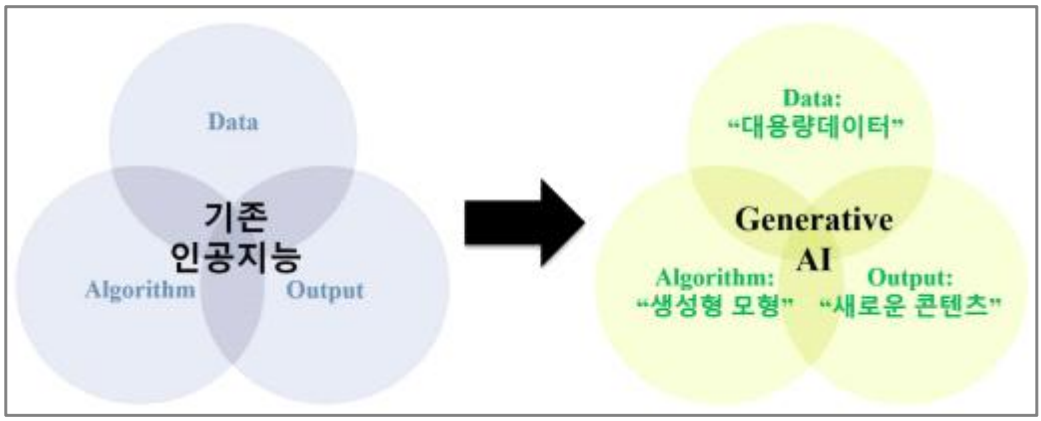

그림 3. 기존 인공지능과 생성형 인공지능의 차이점[15)]

연세대학교
YONSEI UNIVERSITY

2023년 생성형 인공지능 기술은 놀라운 수준으로 발전하며, 4월 의학저널인 미국의학협회지(JAMA)에 발표된 연구에 따르면 ChatGPT가 의학적 질문에 대해 내놓은 답변이 정보의 품질과 공감 측면에서 의사의 답변을 모두 능가하는 것으로 나타났다. 2024년에 대규모 언어모델의 학습 및 추론 비용을 현저히 낮출 획기적인 그래픽처리장치(GPU)인 GH200이 공개되기도 했다. 최근 들어서는 기존의 대규모 언어 모델(LLM)을 넘어 텍스트 뿐만 아니라 이미지와 음성까지도 활용할 수 있는 대규모 멀티모달 언어모델(LMM)이 부각되고 있다.

현재는 초거대 인공지능을 토대로 한 생성형 인공지능 모델 성능이 뛰어나, 초거대 인공지능이 생성형 인공지능이라는 인식이 확산되고 두 가지의 기술적인 융합·연결성으로 인해 '초거대 생성형 인공지능'과 같은 한 단어로 통용되기도 한다.[14)]

초거대 인공지능(Super-Giant AI, Hyperscale AI)는 인간 수준의 복잡한 작업을 수행하기 위하여 방대한 양의 파라미터(parameter)1)와 데이터로 학습한 대규모 AI 시스템 또는 서비스를 의미 한다.[16)] 일반적으로는 수천억 개 이상의 파라미터를 가지는 대규모 인공신경망(artificial neural network)을 사용할 때 초거대 인공지능이라고 한다. 초거대 인공지능은 대규모(거대) 언어모델(Large Language Model, LLM), 초거대 언어모델 등으로 혼용되어 사용되거나 이를 포함하는 개념이다.[17)] 이는, 대규모 언어모델은 사람들이 사용하는 언어(자연어)를 학습하여 실제 인간과 유사한 문장을 생성하기 위한 언어모델을 의미하는데 점차 규모가 커지면서 초거대 인공지능으로 진화하고 있기 때문이다.

초거대 인공지능 대부분이 대규모 언어모델이나 멀티모달형으로 진화하고 있으며 프롬프트(prompt)에 문장을 입력했을 때 문장만 생성하는 대부분의 언어모델과 달리 멀티모달 AI는 텍스트, 이미지, 음성, 영상 등을 동시에 처리할 수 있다는 특징이 있다.[17)] 초거대 인공지능 개발은 파운데이션(기반) 모델(Foundation Model, FM) 이라고 하는 초거대 인공지능 모델을 구축하고 파인튜닝(fine-tuning)등의 과정을 통해 여러 업무와 서비스에 적용을 가능케 하며 파운데이션 모델은 방대한 양의 데이터를 자기지도학습(self-supervised learning) 방식 등으로 사전학습이 된다.

연세대학교
YONSEI UNIVERSITY

표 1. 파운데이션 모델 학습 방법

학습 방법	특징
자기지도학습	정답이 주어지지 않은 데이터(예: 고양이 사진에 '고양이'라는 정답을 입력하지 않은 데이터)를 AI가 스스로 정의와 규칙을 찾아 분류할 수 있도록 학습시키는 방식
퓨샷 러닝(few-shot learning)	매우 적은 수의 데이터만 주어진 상황에서 AI를 학습 시키는 방식
제로샷 러닝(zero-shot learning)	추가 데이터 학습 없이 AI가 특정한 업무를 수행하도록 학습시키는 방식
지도학습	인간이 개입하여 문제와 정답을 모두 알려주고 학습시키는 방식
비지도학습	지도학습과 달리 정답을 알려주지 않고 AI가 유사한 것과 서로 다른 것을 구분하여 군집을 만들 수 있도록 하는 학습 방식
강화학습	보상 및 벌칙과 함께 여러 번의 시행착오를 거쳐 AI가 스스로 학습하게 하는 방식

생성형 인공지능은 다양한 형태 및 기술이 존재하나 본 연구에서는 헬스케어 특히 의료기기와 밀접한 산업분야에 주로 사용되는 기술을 중심으로 생성형 인공지능에 대한 개념을 정리하고자 한다.

연세대학교
YONSEI UNIVERSITY

2.4. 생성형 인공지능 서비스

생성형 인공지능은 텍스트, 이미지, 오디오, 합성 데이터 등 다양한 유형의 콘텐츠를 생성할 수 있는 기술로 딥러닝의 하위 분야이다. 또한 지도, 비지도, 준지도 학습 방법을 통해 레이블이 있는 데이터와 없는 데이터를 모두 처리할 수 있는 기술이다. 생성형 인공지능은 크게 판별형 AI (Discriminative AI)와 생성형 AI (Generative AI)로 설명할 수 있다.[12)]

그림 4. 식의약 분야에 생성형 인공지능을 도입 적용할 수 있는 분야 예시[12)]

많은 사람들이 기술과 서비스를 혼용하여 사용하고 있지만 개발자 등이 인공지능 기술을 이용하여 사람들이 활용할 수 있도록 만든것이 인공지능 서비스이다. 예를 들어 OpenAI사의 GPT-4는 기술명이고, ChatGPT는 이 기술을 활용한 서비스를 의미한다. 아래 그림에서는 최근 4년간 다양한 AI기술 · 서비스를 요약한 내용으로 LLM기술의 개발 타임라인을 나타낸다.

연세대학교
YONSEI UNIVERSITY

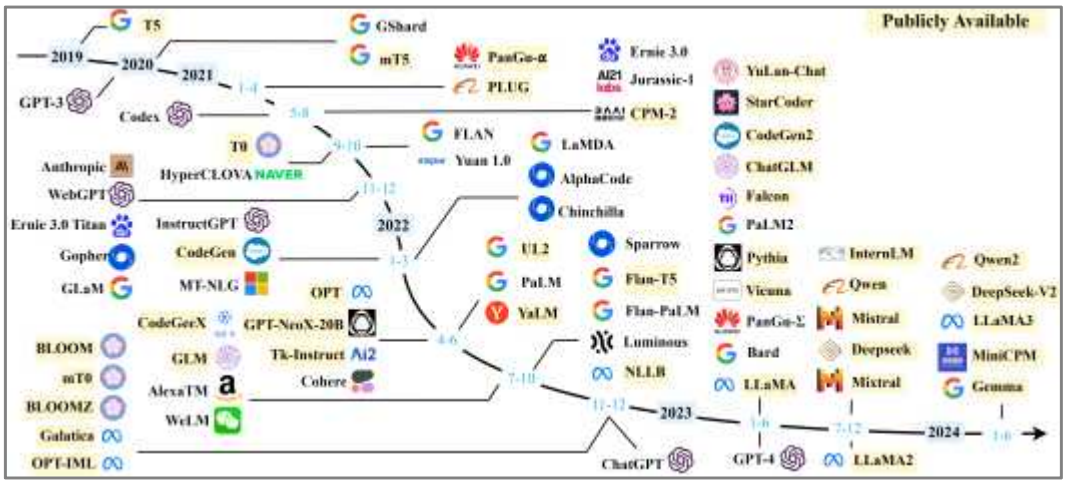

그림 5. 최근 몇 년 동안 기존 대규모 언어 모델(10B보다 큰 크기)의 타임라인[19)]

표 2. 국내외 생성형 인공지능 관련 주요 서비스[19)]

개발사	서비스명	서비스 내용
OpenAI	ChatGPT	여러 번의 검색 대신 질문에 대한 정확도 높은 답변 제공. 거짓답변 등에도 불구하고 검색시장 전반에 큰 변화 촉발
MS	코파일럿	GPT-4를 기반으로 개발된 코파일럿은 사용자 의도 및 업무 목적을 파악하여 문서 초안 작성, 문서 기반 PPT 파일 생성, 엑셀 데이터 분석 및 시각화 등 업무 효율 극대화
네이버	클로바X	한국형 거대언어모델(LLM) 인공지능인 하이퍼클로바를 기반으로 구축하였으며, 창작, 요약, 추론, 번역, 코딩 등을 할 수 있음
업스테이지	아숙업 (AskUp)	ChatGPT 기반으로 OCR 기술을 결합해 개발한 카카오톡 AI 챗봇으로 사용자가 문서를 촬영하거나 전송하면 내용을 읽고 이해하고 답변을 제공
LG	틸다	LG가 개발한 언어-이미지 양방향 데이터를 생성하는 AI

연세대학교
YONSEI UNIVERSITY

2.5. 생성형 인공지능의 의료분야 적용

의료분야에서 적대적 생성 신경망(GAN) 및 대규모 언어 모델(LLM), 대규모 멀티모달 언어모델(LMM)과 같이 생성형 AI 모델은 텍스트, 이미지, 비디오 등을 포함한 다양한 데이터 모달리티를 생성하는 데 사용되며, 특히 대규모 멀티모달 언어모델(LMM)은 의료 이미지(예: MRI 및 CT 스캔), 시계열 데이터(예: 웨어러블 장치의 센서 데이터 및 전자 건강 기록), 오디오 녹음(예: 심장 및 호흡음 및 환자 인터뷰), 텍스트(예: 임상 노트 및 연구논문), 비디오(예: 수술 절차) 및 오믹스 데이터(예: 유전체 및 단백질체)를 포함하는 다양한 데이터 소스의 정보를 동시에 통합하고 해석할 수 있다. 이를 통해 신약 개발, 의료 진단, 임상 문서화, 환자 교육, 맞춤형 의료, 의료 관리 및 의학 교육 등 다양한 시나리오에 적용된다.[20)]

2.5.1 적대적 생성 신경망(Generative Adversarial Network, GAN)

적대적 생성 신경망 GAN은 주로 이미지 생성에 사용되며 noise z를 입력값으로 받아 이미지를 생성해 내는 생성자(generator)와 생성된 이미지와 실제 이미지를 비교하여 주어진 이미지가 생성된 이미지인지 실제 이미지인지를 구별하는 판별자(discriminator)로 이루어져 있다. GAN의 생성자와 판별자를 설명할 때 흔히 사용되는 비유는 위조범과 탐정의 관계이다. 위조범 생성자는 탐정 판별자를 속이기 위해 사실적인 모조품을 만든다. 탐정은 주어진 그림이 모조품인지 실제 그림인지를 분간한다. 반복적인 경쟁관계를 통해 위조범은 점점 더 사실적인 모조품을 만들어내고, 탐정도 점점 더 사실적인 모조품과 실제 그림을 잘 분간해 내다가 어느 순간 너무 완벽한 모조품이 생기면 탐정은 이게 실제인지 모조품인지를 분간하지 못하고 확률에 이를 맡기게 되며 학습이 끝난다.

GAN은 기계학습의 하위 분야인 생성 모델링에서 중요한 역할을 한다. 생성 모델링은 주어진 데이터 분포를 학습하여 새로운 데이터를 생성하는 것을 목표로 하게 된다. GAN은 비지도 학습(Unsupervised Learning) 방법의 일종으로, 레이블이 없는 데이터로부터 학습할 수 있는 강력한 도구로 사용된다. 이는 특히 대량의 비레이블 데이터가 존재하는 상황에서 유용하며, 이미지 생성, 데이터 증강, 스타일 변환 등 다양한 분야에서 혁신적인 성과를 내고 있다. 아래는 GAN이 실제 의료분야에 적용될 수 있는 예시를 나열한 것이다.[21)]

연세대학교
YONSEI UNIVERSITY

(1) GAN 기반 합성 뇌 PET 이미지 생성:

적대적 생성 신경망(GAN)을 사용하여 합성 의료 이미지를 생성하는 접근방식이며 알츠하이머병의 세 가지 단계, 즉 정상 대조군(NC), 경미한 인지 장애(MCI), 알츠하이머병(AD)에 대한 뇌 PET 이미지를 생성할 수 있다.[22)]

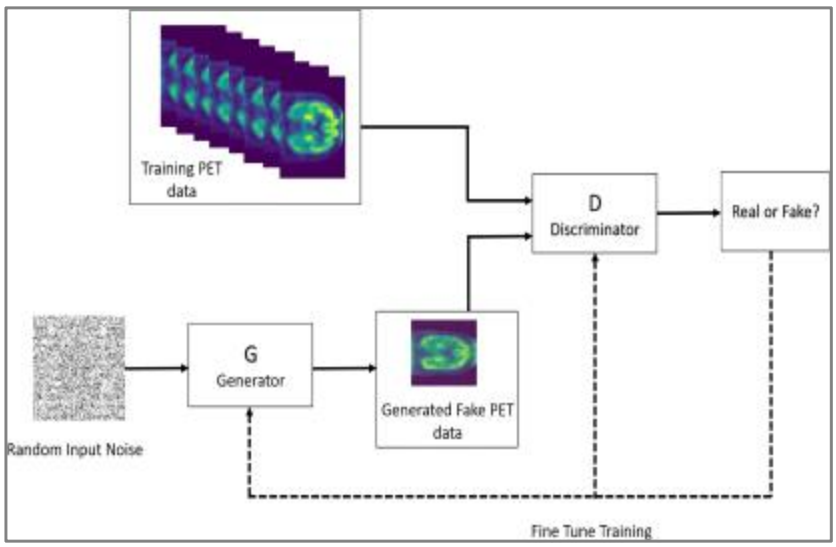

그림 6. 합성 PET 이미지 생성 모델 예시[22)]

(2) GAN을 사용한 MRI 잡음 제거:

MRI 잡음 감소를 위해 생성적 적대 신경망을 사용하는 방법으로 합성곱 인코더-디코더 네트워크 기반 생성기를 사용하여 잡음을 최대한 최소화할 수 있다. 이 방법을 통해 노이즈 제거와 해부학적 구조의 보존 측면에서 기존의 기술보다 더 나은 성능을 보여줄 수 있다.[23)]

(3) GAN을 사용한 흉부 CT 이미지의 자동 다중 장기 분할:

아틀라스 기반 방법은 자동 분할에 가장 일반적으로 사용되는 접근 방식이지만 환자의 다양한 해부학적 특징으로 인해 항상 정확한 결과를 제공하지 못할 수 있다. GAN을 사용하여 방사선 치료 계획을 위해 흉부 컴퓨터 단층 촬영(CT)에서 여러 흉부 OAR(Organ at Risks)을 자동으로 분할 할 수 있다.[24)]

연세대학교
YONSEI UNIVERSITY

2.5.2 대규모 언어모델(Large Language Model, LLM)과 대규모 멀티모달 모델 (Large Multimodal Model, LMM)

LLM은 주로 텍스트 데이터를 기반으로 하는 모델이다. 이 모델의 기본 목표는 자연어 처리를 통해 텍스트를 이해하고 생성하는 것으로 GPT-4와 같은 모델이 대표적인 예시이다. LLM은 방대한 양의 텍스트 데이터를 학습하여 언어 구조를 이해하고 자연스러운 문장을 생성할 수 있는 특징이 있다. LLM의 응용은 자연어 처리(NLP)에서 두드러지는데 대화형 AI, 텍스트 요약, 번역 시스템, 챗봇 등에서 활발히 활용되고 있다. 특히 GPT-4 같은 모델은 사용자의 명령어를 이해하고, 의미 있는 답변을 제공하거나 문장을 자연스럽게 이어 나가는 성능을 보여주고 있다.
LMM은 텍스트뿐만 아니라 이미지, 비디오, 오디오 등 다양한 형태의 데이터를 동시에 처리할 수 있는 모델로, 단일 데이터 유형에 국한되지 않고 서로 다른 형태의 데이터를 융합해 종합적인 분석과 결정을 내릴 수 있다는 특징이 있다.[20)]
아래는 실제 의료분야에 LLM과 LMM이 사용될 수 있는 예시를 나열한 것이다.

(1) EHR 및 환자 기록 분석:
LLM을 활용하여 비정형화 되고 양식이 다른 EHR(Electronic Health Record) 및 환자 기록을 분석하여 가독성 있는 내용으로 정리하거나 유의미한 의료정보를 추출하고 해석할 수 있다. 또한 검사 결과, 의사 메모 및 의료영상 보고서의 데이터를 상호 연결하여 환자의 건강에 대한 보다 전체적인 관점을 생성할 수 있고, 환자 케이스가 제시될 때, LLM은 설명된 증상을 기반으로 잠재적 진단 목록을 생성하고, 진단을 확인하기 위한 적절한 검사 및 치료 계획을 제안할 수 있다.

(2) 의료영상 판독문 작성 보조:
이미지를 분석하는 컴퓨터 비전 모델이 의료영상을 분석 후 LLM이 판독문 초안을 작성하여 의료진의 의사결정에 도움을 줄 수 있다. 기존 진단 AI에서는 제한된 병변 또는 이상 부위를 식별해 병변의 위치, 심각도 등의 정보만 제공했다면 LLM은 실제 판독문과 같은 형식의 정보를 생성할 수 있다. 이 과정에서 언어모델이 분석한 의료정보가 포함된다면 이미지 분석 모델과 언어모델이 결합 된 LMM의 형태라고 볼 수도 있다.

연세대학교
YONSEI UNIVERSITY

(3) 생체데이터 분석:

LMM은 시간 경과에 따라 기록된 환자의 심박수와 연속 혈당 모니터링(CGM)을 통해 패혈증 또는 심장 질환과 같은 임상 악화의 초기 징후를 감지할 수 있고, 음성 패턴과 호흡 소리를 분석하여 천식 또는 만성 폐쇄성 폐질환(COPD)과 같은 호흡기 질환을 발달 초기에 식별하고, 우울증, 불안 또는 스트레스를 나타낼 수 있는 음성 패턴, 정서적 어조 및 음성 톤의 미묘한 변화를 감지하는 데 사용할 수 있다.

연세대학교
YONSEI UNIVERSITY

2.6. 생성형 인공지능 의료기기 개발 현황

생성형 인공지능 기술을 다양한 의료분야에 적용하려는 노력은 광범위하게 이루어지고 있으나 현재까지 의료기기의 형태로 개발되어 규제 승인을 받은 제품은 존재하지 않고 있다. 실제로 2023년 10월까지 FDA에 승인된 AI/ML 기반 의료기기는 691개로, 그 중 108개(약 15%)가 2023년에 승인되었다. 연간 승인 건수 추세는 2016년부터 급속도로 증가하였으며 그 이후로 지속적으로 매년 승인된 기기 수가 증가해 왔다. 특히 ChatGPT, Llama 등과 같은 대규모언어모델(LLM)의 최근 증가에 따라 AI/ML 통합 의료 기기에 대한 관심이 급증했으나 보고된 자료에 따르면 2023년 10월까지 FDA에 승인된 AI/ML 기반 의료기기 691개 중 생성형 인공지능을 사용하는 의료기기는 승인을 받은 경우는 현재까지 존재하지 않는다.[25)]

비록 현재까지 규제기관에 승인을 받은 생성형 인공지능을 사용하는 의료기기 사례는 없지만 관련된 연구개발은 지속적으로 이루어지고 있다. 해외 빅테크 업체인 구글은 자사의 생성형 AI 모델 '제미나이'에 기반한 최첨단 의료용LLM '메드-제미나이'의 연구개발 과정을 공개하였다. '메드-제미나이'의 경우 헬스케어 분야에 최적화된 멀티모달(Multi-modal) 인공지능 모델로, 해당 연구에서는 흉부 엑스레이를 토대로 작성한 임상 결과 보고서에 대한 블라인드 테스트를 진행했으며 모델의 성능이 기존 의사가 작성한 보고서와 비슷하거나 더 우수하다는 평가가 72%에 달했다는 보고가 있다.[26)]

이와 유사하게 국내에서도 생성형 인공지능을 사용한 의료기기의 개발이 활발히 이루어지고 있다. 카카오브레인은 카라-CXR 제품에 생성형 인공지능을 사용하여 흉부 엑스레이 사진을 분석, 판독문 작성을 보조해 주는 제품을 개발 중에 있으며 국제학술지 '다이그노스틱스'에 게재한 논문에 따르면 카라-CXR의 흉부 엑스레이 사진 분석 정확도는 70% 수준으로 40% 중반에 불과한 GPT-4보다 25% 정도 높았다는 보고가 있다.[27)] 이와 유사하게 의료 인공지능 소프트웨어 개발 회사인 루닛 또한 흉부 엑스레이 사진을 분석하여 판독문 작성을 보조해 주는 생성형 인공지능 기술을 개발 중에 있다.

연세대학교
YONSEI UNIVERSITY

3. 생성형 인공지능 의료기기 관련 규제 동향

3.1. 주요 기관별 규제 동향

3.1.1. 한국

한국 식약처는 세계 최초로 디지털의료제품의 안전성·유효성의 효율적 평가를 위해 인공지능, 네트워크 연결 등 디지털 특성에 특화된 임상시험, 허가, 유통관리 등 전주기 규제 체계에 대한 내용이 담긴 디지털의료제품법을 제정하고 2025년 1월에 시행할 예정이다.

표 3. 디지털의료기기, 디지털 융합 의약품 및 디지털 의료·건강지원기기 정의

구분	정의
디지털의료제품	디지털의료기기, 디지털융합의약품 및 디지털의료·건강지원기기를 말함
디지털의료기기	지능정보기술, 로봇기술, 정보통신기술 등 총리령으로 정하는 첨단 기술(이하 “디지털기술”이라 한다)이 적용된 「의료기기법」 제2조제1항에 따른 의료기기(「체외진단의료기기법」 제2조제1호에 따른 체외진단의료기기를 포함한다) 또는 이와 디지털의료·건강지원기기가 조합된 제품으로서 다음 각 목의 어느 하나에 해당하는 제품을 말한다. 가. 질병의 진단·치료 또는 예후를 관찰하기 위한 목적으로 사용되는 제품 나. 질병의 치료 반응 및 치료 결과를 예측하기 위한 목적으로 사용되는 제품 다. 질병의 치료 효과 또는 부작용을 모니터링하기 위한 목적으로 사용되는 제품 라. 그 밖에 재활을 보조하는 목적으로 사용되는 제품으로서 식품의약품안전처장이 지정하는 제품

연세대학교
YONSEI UNIVERSITY

구분	정의
디지털융합의약품	「약사법」 제2조제4호에 따른 의약품(「첨단재생의료 및 첨단바이오의약품 안전 및 지원에 관한 법률」 제2조제5호에 따른 첨단바이오의약품을 포함한다. 이하 “의약품”이라 한다)과 디지털의료기기 또는 디지털의료·건강지원기기가 조합된 의약품을 말한다. 다만, 주된 기능이 디지털의료기기에 해당하는 경우는 제외한다.
디지털의료·건강지원기기	디지털의료기기에 해당하지 아니하나 의료의 지원 또는 건강의 유지·향상을 목적으로 생체신호를 모니터링·측정·수집 및 분석하거나, 생활습관을 기록·분석하여 식이·운동 등 건강관리 정보를 제공하는 목적으로 사용되는 디지털 기술이 적용된 기구·기계·장치·소프트웨어 또는 이와 유사한 제품으로서 식품의약품안전처장이 지정하는 제품을 말한다.

이 법안에서 디지털의료기기는 기존 하드웨어 중심의 규제 체계 적용이 어려워 새로운 규제로의 과학적 접근이 필요한 디지털 기술이 적용된 의료기기로 정의된다. 독립형 소프트웨어, AI 기반 지능형로봇, 디지털트윈 등의 기술을 활용한 제품들이 대부분이며 생성형 인공지능 기반 기기도 대상에 포함이 된다. 디지털의료제품법의 주요 내용은 아래와 같다.

(1) 임상시험

디지털의료제품법에서는 임상시험의 신속성을 높이기 위해 임상시험계획 승인 생략 제도를 도입했다. 이 제도는 인체에 미치는 위해도가 낮은 디지털의료기기에 한해, 특정 요건을 충족하는 경우 총리령으로 정하는 바에 따라 임상시험계획 승인을 생략할 수 있도록 허용한 것이다. 이를 통해 디지털의료기기에 대한 임상시험 절차를 간소화하고, 연구 개발이 더욱 신속하게 이루어질 수 있도록 지원하고자 하는 목적이다.

또한 기존의 의료기기 임상시험 체계는 지정된 임상시험 기관 내에서만 데이터를 수집하도록 제한되었으나, 디지털의료기기의 특성을 반영하여 웨어러블 기기와 같은 네트워크 기반 기기의 임상시험을 고려한 법적 근거가 마련되었다.

이를 통해 휴대용 기기를 활용해 지정된 장소가 아닌 외부 환경에서 수집된 데이터가 디지털 바이오마커로 활용될 경우, 적절한 감독 하에 데이터를 수집하고 임상시험

연세대학교
YONSEI UNIVERSITY

에 활용할 수 있도록 허용했다.

(2) 전자적 침해

디지털의료기기는 통신 및 네트워크와 결합 된 특성상 사이버 보안 위협에 노출될 가능성이 높으므로 이를 고려하여 법안에서는 전자적 침해 개념을 도입하였으며, 이를 방지하기 위한 기술적, 행정적 조치를 마련하였다.

- 결함 및 오류 수집 체계 도입: 디지털의료기기에서 발생할 수 있는 결함과 오류 데이터를 지속적으로 수집하고 이를 통해 제품의 안정성을 강화.
- 보안 지침 수립: 해킹, 바이러스 등 전자적 침해로부터 제품과 사용자를 보호하기 위한 체계적 지침 마련.
- 기술 지원 강화: 사이버 공격에 대한 방어 기술과 유지보수 지원 체계 구축.

이를 통해 디지털의료기기가 외부 위협에 대응할 수 있는 기반을 갖추도록 하고, 사용자 안전성을 확보할 수 있는 법적 근거를 마련한 것으로 판단된다.

(3) 실사용 평가

디지털의료기기의 실사용 데이터를 수집·평가하고 이를 법적 근거로 명확히 규정했다. 실사용 평가의 핵심은 다음과 같다.

- 의료 현장에서 사용 중 생성된 안전성·유효성 데이터를 수집하여 이를 허가, 인증, 신고 절차에서 반영.
- 빅데이터 및 기계 학습을 통해 디지털의료기기의 성능이 진화하는 특성을 규제 체계에 통합.

이는 미국 FDA 및 유럽 EMA 등 선진국의 실사용증거(Real-Wolrd Evidence, RWE)제도와 유사하게 디지털의료기기의 과학적 증거 기반을 강화하여 허가 심사 및 시판 후의 지속적인 안전성, 유효성을 검증할 수 있는 제도이다.

(4) 제품 품질관리

디지털의료기기 제품의 특성을 반영한 전주기 품질관리 평가 제도가 도입되었습니

연세대학교
YONSEI UNIVERSITY

다. 현재 허가체계는 변경할 때마다 규제당국의 변경 허가를 받도록 하는 원칙으로 인해 시판 후 축적된 데이터와 사용자 피드백을 기반으로 지속 개선되는 디지털제품의 특성을 충분히 반영하지 못하고 있다.

이와 달리, 디지털의료기기에 특화된 우수관리체계 인증을 통해 제조업체의 품질관리 역량을 평가하고, 일정 범위 내에서 자율적인 변경 및 품질관리를 가능하게 하는 제도가 도입되었다. 이 제도는 미국의 **소프트웨어 사전인증 파일럿 프로그램(Software Precertification (Pre-Cert) Pilot Program)**을 참고하여 설계 되었다.[28)]

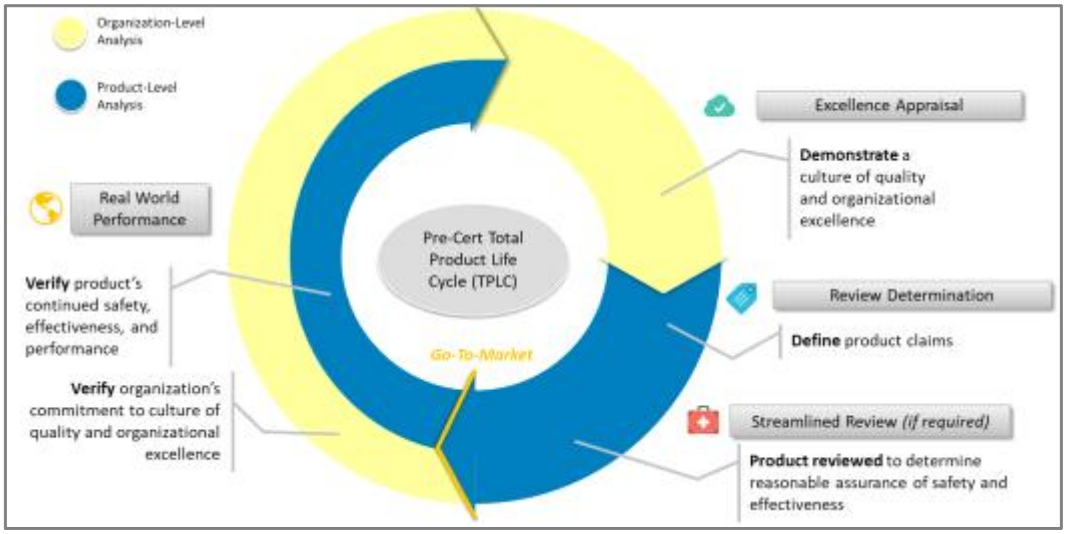

그림 7. Pre-Cert Program Approach throughout TPLC[28)]

또한 식약처에서는 '빅테크 기업 규제혁신 프로그램'을 통해 2023년부터 인공지능 의료기기 개발 업체들의 정확한 요구를 파악하고 인공지능 의료기기 분야 규제개선 과제 선정하여 규제개선을 추진하고 있다. 식약처는 이 프로그램을 통해 각 업체의 사업 진행 현황과 추진계획, 규제개선 과제 등을 발굴하여 보다 원활한 시장진입을 지원하고 있다.[3)]

연세대학교
YONSEI UNIVERSITY

표 4. '빅테크 기업 규제혁신 프로그램' 기 발굴 규제개선 과제[3)]

연번	과제명	결과물
1	(제안자: 카카오브레인) • 생성형 AI 의료기기(이하 AIMD) 특성을 고려한 심사기준 마련 - 생성형 AIMD 특성을 고려해 허가 시 활용할 수 있는 안전성·유효성 심사기준 마련 필요 ※ VIP, '디지털 권리장전('23.10)에 '생성형 AI 기반 의료기기 심사 규제 가이드라인 마련(18조-8번)' 과제 포함	생성형 AI 의료기기 허가·심사 방안 마련
2	(제안자: 카카오헬스케어) • AIMD의 신속 제품화 지원 - 관련 업체가 별도 식약처 질의 없이 디지털 의료기기 해당여부를 자체 판단할 수 있는 방안 강구 필요	① 한시품목분류운영 ② 온라인 대화형 민원안내시스템 고도화 ③ 유헬스케어 의료기기 FAQ 발간 ④ 디지털의료기기 해당여부 사례
3	(제안자: 카카오헬스케어, KT) • AIMD 해외시장 진출 임상·실증·인허가 컨설팅 지원 - 관계부처(과기부 등)와 협력하여 AIMD 해외 인허가 지원 방안 강구 필요	(한-싱) AIMD 임상시험 방법 설계 공동 가이드라인 개발
4	(제안자: 카카오헬스케어) • AIMD 국내 인·허가 시 RWD/RWE 활용 활성화 - 제품화 시간·비용절감을 위해 AI 의료기기 업체가 RWD/RWE를 능동적으로 활용할 수 있는 방안 강구 필요	① RWE 가이드라인 개정 ② RWE 기반 평가체계 마련 보고서 발간
5	(제안자: KT) • AIMD 임상기준 완화 - AIMD등 독립형 SW 특성에 맞도록 임상시험 기준 완화방안 강구	의료기기법 시행규칙 개정(독립형SW 임상승인 면제)

연세대학교
YONSEI UNIVERSITY

3.1.2. 미국

미국 FDA는 영국의 의약품 및 의료 제품 규제 기관(MHRA)과 캐나다의 Health Canada와 긴밀히 협력하여 의료기기의 인공지능 사용에 대한 공통 원칙을 수립하고 있다.

2024년 11월에 열린 FDA 디지털헬스자문위원회(THE DIGITAL HEATH ADVISORY COMMITTEE MEETING) 회의에서는 생성형 인공지능 의료기기에 대한 규제 접근 방식에 대한 논의가 이루어졌다. 이 미팅에서는 FD&C Act(The Federal Food, Drug, and Cosmetic Act)의 section 201(h)의 정의에 따라 생성형인공지능(GenAI)이 제품의 결과물이나 성능에 필수적인 기기를 지칭하기 위해 생성형인공지능 기반 기기(GenAI-enabled device)라는 용어를 사용하며, 생성형 인공지능 뿐만 아니라 일부 AI에 광범위하게 적용될 수 있는 위험과 TPLC(Total Product Life Cycle) 전반에 걸쳐 인공지능 및 생성형 인공지능 기반 기기의 규제에 대한 현재 과제에 대해 아래와 같이 설명 하였다.[29)]

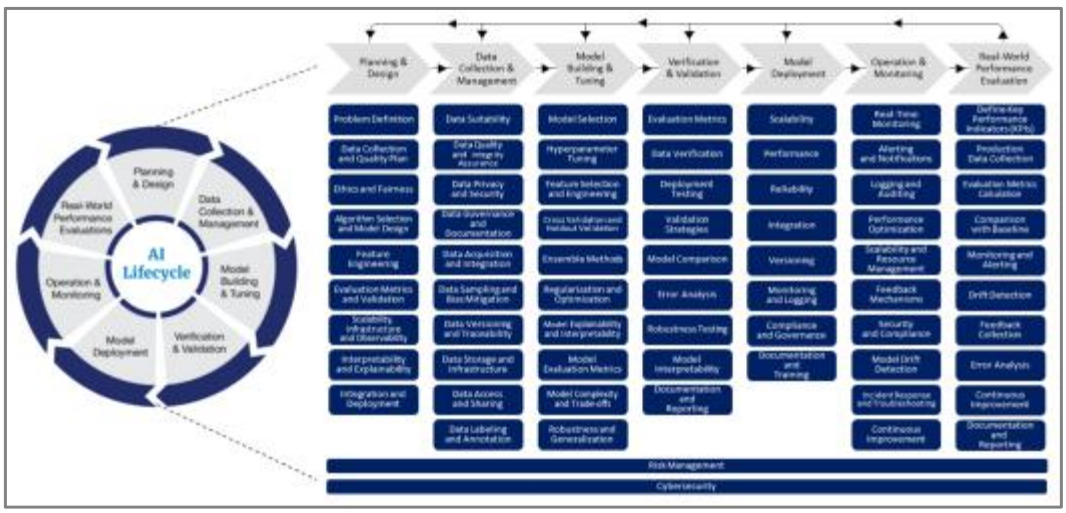

그림 8. AI Lifecycle Considerations

연세대학교
YONSEI UNIVERSITY

(1) 위험기반 접근

2024년 7월 미국국립표준기술 연구소에서는 생성형 인공지능의 위험을 식별과 관리에 대한 조치가 담긴 인공지능 위험관리 프레임워크:생성형인공지능 프로파일(NIST-AI-600-1, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile)을 발표 하였다.[30)]

FDA에서는 모든 기기와 마찬가지로 생성형 인공지능이 적용된 의료 제품의 안전성과 효과에 대한 합리적인 보증을 제공하기 위해 제품의 사용 목적과 기술적 특성을 고려한 위험 기반 접근 방식을 따른다고 규정하였다.

(2) 규제 대상이 되는 기기의 정의

FDA's guidance Policy for Device Software Functions and Mobile Medical Applications에 설명된 대로 FDA는 의료기기이며 기기가 의도한 대로 작동하지 않을 경우 환자의 안전에 위험을 초래할 수 있는 소프트웨어 기능에 대해 규제 감독을 적용한다. 이러한 정책은 생성형 인공지능 기반 기기에도 적용이 되나 생성형 인공지능 기반 기기는 다양한 형태로 개발이 되기 때문에 기기의 사용 목적 및 위험도에 따라 규제 감독 적용의 구분을 둘 수 있다고 본다.

예를 들어, 생성형 인공지능 기반 기기가 의료 전문가나 의료 교육을 위한 도구로 사용되는 경우 이 기기는 FDA에서 규정하는 정의에 해당되지 않게 된다. 또한, 특정 생성형 인공지능 기반 기기는 FDA 규제 대상 기기의 정의에 해당이 되긴 하지만 대중에게 미치는 위험이 낮기 때문에 FDA에서는 시행재량권을 행사할 수도 있다(시행재량권이란 FDA가 규제는 유지하되, 특정 상황에서는 해당 요구사항을 엄격히 적용하지 않겠다는 의미이다).

따라서 FDA에서 정의하는 규제 대상 생성형 인공지능 기반 기기는 환자별 분석을 수행하고 질병 또는 상태의 진단, 치료, 완화, 치료 또는 예방에 사용하기 위해 의료 전문가에게 특정 결과 또는 지침을 제공하거나, 환자별 분석을 수행하고 환자, 간병인 또는 의료 전문가가 아닌 기타 사용자에게 환자별 진단 또는 치료 권장 사항을 제공하는 기기로 규정된다.

연세대학교
YONSEI UNIVERSITY

표 5. 소프트웨어 기능별 미국 FDA의 규제적 접근 방식 구분[29)]

의료기기	시행재량권 대상	비의료기기
가. 기기간, 장치 간 의료 데이터 분석 및 연결하는 구성품에 해당하는 앱 - 의료기기 또는 웰니스 데이터를 환자의 의료 관리(식별, 관리 등) 등을 위해 분석 등 하거나, 의료기기를 제어, 변경 등 하는 경우 - 일반적인 비의료 목적 센서 등을 의료목적으로 활용하는 경우 - 임상의사결정시스템(CDS) 중 독립적인 검토 시간 없는(뇌졸등 환자센터 선별 등) 경우 나. 부착물, 센서 등을 통한 의료기기 기능 구사 다. 환자별 데이터 등 분석하여, 진단/치료/완화/예방 등을 목적으로 특정 결과, 지침 등을 제공하는 경우	가. 환자에세 질병이나 상태를 제공하지 않고 스스로 관리하도록 도움 - 만성질환자 등 환자지도, 체중유지, 영양섭취, 운동 및 건강 유지, 염분 섭취 또는 사전 결정된 약물 투여량 준수를 위한 전략 촉진 - 의사에게 전달할 이미지, 정보 등 캡쳐 및 전송(의사가 판단) - 우울증, 강박장애 환자 불안감 해소(오디오, 영상 등 제공) - 금연, 중독자에게 교육 정보, 알림 / 동기 부여 / 위험 위치 제공 등 - 천식 환자 등에게 위험 위치 제공(환경조건 알림) - 비디오 게임 기반 가정 물리치료보조 - 약초와 약물 상호관계 판멸 - 환자 연령, 성별, 행동 위험 요소 등 기반 의료	가.의학교과서 등 라이브러리 나.의료 교육을 위한 도구(수술훈련 비디오, 해부학 다이어그램, 심정지 시나리오 시뮬레이션 등) 다.일반적 환자 교육용 또는 환자 참조용 - 환자의 질병, 상태, 치료 또는 예정된 절차 관련 정보(의사와 소통 가능) - 질병, 상태 관심사 관련 식품 정보, 임상시험 정보 제공, 응급처치, 알약 모양 등 라. 일반 의료 분야 사무 업무 자동화(보험, 질병코드, 병원대기 등) 마. 일반 지원용 - 모바일 플랫폼 돋보기, 단순 통신용(이메일, 화상 등) 바. 개인 건강 관련 데이터 관리 등 - 식단, 운동 등 기록 관리, 정상적 아기수면, 수유 습관 관리, 운동 모니터링 등

연세대학교
YONSEI UNIVERSITY

의료기기	시행재량권 대상	비의료기기
	기관 상담 예방, 권장 사항 안내(일반적인 징후, 증상 기반 가능한 의학상태 목록 제공 의료전문가 상담 등 필요성 조언 제공) - 최초 미리 정한 의료진에게 알람제공(분석 X) - 복약순응도 향상을 위한 알림 - 활력징후(바이탈 사인) 과거 추세 및 비교 제공 - 혈압 데이터를 이메일 공유 및 추적/추세 파악 - 당뇨병 환자에게 식습관 개발 등 도움이 되는 지침 제공 - 중독 대상 약물 이미지, 기타 메시지 제공 - 단순 의료 분야 인구통계학적 데이터(연령, 성별, 진단명, 약물 등) 분석 나. 의료전문가를 위한 단순한 작업을 자동화 - 간단한 의학적 계산 : BMI, APGAR, 뇌졸중 척도 등	- 두뇌연령 점수, 동기부여, 스트레스 줄이기, 사. 인증된 EHR, PHR 데이터 전송, 저장, 형식 변환, 표시 등 아. 비의료 목적 CDS - 환자의 의료 정보를 기존 레퍼런스 데이터(환자 진단기록, 보고서, 진료지침, 동료문헌 등)와 비교해서 우선순위 등을 안내(의료인이 독립적으로 검토할 수 있는 충분한 시간 제공) - 약물 부작용 방지를 위한 상호작용, 금기 통지, 기존 처방집 등 토대 약물 추천 등 자. 환자 스스로 상태 등 이벤트 측정 및 관리 차. 임상 대화 녹음 기능 카. 천식환자 위치 및 환경 기록 등(흡입기 사용 정보 등) 약물 사용 정보 타. 데이터를 수정하지 않고 의료기기 데이터를 전송, 저장, 형식변환 및 표시하고 연결된 기기의 기능 변경을 하지 않는 기능(단순 저장전송)

연세대학교
YONSEI UNIVERSITY

(3) 생성형 인공지능 기반 기기에 위험 기반 접근 방식을 적용할 때 고려해야 할 사항

생성형 인공지능의 고유한 특성 중 하나는 생성형 인공지능 모델이 새로운 정보를 생성할 수 있다는 점이다. 일부의 정보는 창의적인 응답을 생성하기도 하나 많은 경우 환각(Hallucination)을 일으키기 쉬운 특성을 가지고 있다. 이러한 환각은 식별하거나 설명하기 어려울 수 있으며 기기의 의도된 사용 목적 및 성능에 영향을 줄 수 있다.

또한, 파운데이션모델(Foundation Model)과 같은 모델을 생성형 인공지능 기반 기기의 일부로 사용하는 경우 기기 제조사가 파운데이션 모델의 속성, 아키텍처, 훈련방법론, 데이터 세트 등 세부 정보를 얻는데 어려움을 겪을 수 있으며 이러한 모델은 편향(Bias)에 취약하고 기기의 개발자가 그로 인해 발생할 수 있는 문제를 식별하거나 완화하기 어려울 수 있다. 위험기반 접근 방식에서는 기기 제조사는 이와 같은 기성소프트웨어(Off-The-Shelf Software, OTS)에 대한 위험을 식별하고 적절한 완화조치를 수행해야 한다고 명시하고 있다.

FDA에서는 생성형 인공지능 기반 기기가 안전하고 효과적임을 보장하기 위한 위험 통제 조치로, 기기에 사용된 기초 파운데이션 모델이 제어범위를 벗어난 경우라도 생성형 인공지능 기반 기기에 적용된 모델을 어느 정도까지 통제할 수 있는지 고려하는 것이 중요하다고 언급했으며 그 외에도 적절한 거버넌스, 장치 안전성과 관련된 적절한 피드백 메커니즘, 그리고 실사용 환경에의 성능 평가 등이 포함될 수 있다고 언급했다.

(4) TPLC를 통해 GenAI 지원 디바이스에 대한 유효한 과학적 증거에 관한 고려사항

현재까지 생성형 인공지능 기반 기기의 안전성 및 유효성을 입증하기 위해서는 추가적인 과학적 근거가 필요하다. 특히, 개발형 입력 및 출력 형식(Unlock)을 사용하는 생성형 인공지능 기반 기기의 경우 필요한 과학적 증거를 수집하는데 어려움을 겪을 수 있다. 생성형 인공지능 모델은 대규모 데이터를 기반으로 설계되어 수십억개의 모델 파라미터를 가지고 있으며 매우 복잡한 구조를 가지고 있다. 이러한 이유 때문에

연세대학교
YONSEI UNIVERSITY

입력 데이터의 작은 변화가 출력에 큰 차이를 초래할 수 있는데 생성형 인공지능 모델은 이러한 개발형 입력을 허용하기 때문에 모든 입력을 평가하는 것은 현실적이지 않다. 이에 대해 FDA에서는 아래와 같은 생성형 인공지능 기반 기기 맞춤형 시판 전 증거 및 시판 후 시장 모니터링 전략을 제시하고 있다.

1) 현재 시판 전 증거 요구 사항

설계세부사양: 생성형 인공지능 기반 기기의 경우, 설계 세부 사항에는 특정 설계 사양(예: 어텐션 메커니즘, 모델 병합 등), 모델 매개변수(예: 온도, Top-K, Top-P 등), 내장된 프롬프트 전략(예: 인컨텍스트 학습, 제로샷 프롬프트 등), 그리고 최종 사용자에게 제공되는 프롬프트 기능과 같은 정보가 포함될 수 있다.

생성형 인공지능 기반 기기의 훈련과 관련하여, 초기 파운데이션 모델에 대한 데이터 관리 및 모델 개발과 관련된 정보를 가능한 한 합리적인 범위 내에서 제공해야 한다. 또한, 특정 생성형 인공지능 기반 기기에 맞춰 생성형 인공지능 모델을 세부적으로 조정(fine-tuning)하기 위한 데이터 관리 및 구체적인 세부 사항도 포함될 수 있다.

생성형 인공지능 기반 기기에 대한 다양한 투명한 정보를 사용자에게 제공할 수 있어야 한다. 예를 들어, 기기의 설계, 기기의 테스트 방법, 자율성 수준, 그리고 사용자와의 상호작용 방식(예: 프롬프트 엔지니어링)을 포함하며, 이러한 요소는 기기의 출력에 영향을 미칠 수 있다. 또한, 생성형 인공지능 기반 기기의 자율성 수준과 이를 통해 인간이 개입하는 방식(human-in-the-loop)이 포함되어 있는지, 그리고 최종 사용자에게 제공되는 제어 수준이 어떻게 이루어지는지 설명하는 것도 요구될 수 있다.

2) 시판 전 증거 요구를 위한 새로운 방법론

일부 생성형 인공지능 기반 기기의 경우, 현재 수준의 성능 평가 방법론이 여전히 적용될 수 있다. 추가적으로 프롬프트 엔지니어링, 입력 및/또는 출력 유형에 대한 품질관리, 특정 주제에 특화된 파운데이션 모델 사용과 같은 방법론과 전략을 활용하면 현재의 정량적 성능 평가 방법론을 특정 생성형 인공지능 기반 기기를 평가하는데 사용할 수 있다.

그러나 현재의 정량적 평가 접근법은 기기 성능에 대해 철저하거나 완전한 평가를 제공하지 못할 수 있으며, 추가적인 성능 지표로 보완되어야 할 수 있어야 한다. 예를

연세대학교
YONSEI UNIVERSITY

들어, 생성형 인공지능 기반 기기의 자율성 수준, 투명성, 설명 가능성과 관련된 파운데이션 생성형 인공지능 모델의 특성을 정확히 파악하는데 도움이 되는 질적 성능 평가 방법론이 적용될 수 있다.

3) 시판 후 증거 요구 사항

제조업체는 인공지능 기반 및 생성형 인공지능 기반 기기의 안전성과 성능을 선제적으로 모니터링하기 위한 강력한 사후 시장 모니터링 전략을 수립해야 한다. 이는 FDA 가이던스 Consideration of Uncertainty in Making Benefit-Risk Determinations in Medical Device Premarket Approvals, De Novo Classifications, and Humanitarian Device Exemptions에서 명시된 바와 같이 기기의 이득-위험 정보의 불확실성을 해결하기 위한 조치로 시판 후 시장 데이터 수집을 하고 시판 전 시장과 시판 후 시장 전반에 걸쳐 지속적이고 강력한 증거 생성을 통해 시판 전 증거를 보완할 수 있기 때문이다. 특히, 생성형 인공지능 기반 기기가 배포된 이후에도 기기의 정확성, 적합성, 신뢰성을 유지할 수 있도록 특정 의도된 사용 목적에 맞게 효과적으로 평가하고 모니터링 하는 방법을 고려해야 한다.

(5) 성능 지표

현재까지 생성형 인공지능 기반 기기에 대한 명확한 성능지표 및 평가 방법은 정해진 것이 없으나 생성형 인공지능 기술의 특성 및 의도된 사용목적에 따른 맞춤형 성능 지표를 설정해야 한다. 이러한 성능지표의 예로는 생성 텍스트의 경우(perplexity, quantitative comparison to reference text), 생성 이미지의 경우(Frechet Inception Distance(FID), Structural Similarity Index Measure(SSIM)), 또는 생성 오디오의 경우 (Log-Spectral Distance, Perceptual Evaluation of Speech Quality)와 같은 모달리티별 지표나, 생성형 인공지능 기반 기기의 오류 빈도와 오류 유형과 같은 기능 기반 지표가 포함될 수 있다.

연세대학교
YONSEI UNIVERSITY

3.1.3. 영국

2021년 9월에는 '영국을 세계적인 AI 대국으로 만들기' 위해 10개년 계획을 제시하고, 생성형 AI의 활용에도 긍정적인 자세를 보였다. 2023년 3월에는 AI 규제에 관한 화이트페이퍼가 발표되었고, 2024년 4월에는 의약품 및 의료 제품 규제 기관(MHRA)에서 해당 Impact of AI on the regulation of medical products 라는 화이트페이퍼를 발간하였다. 거기서는 안전성 · 보안성 · 견실성, 적절한 투명성과 설명 가능성, 공평성, 설명 책임과 거버넌스, 그리고 경쟁 가능성과 구제라는 5가지 원칙에 따라 AI 규제를 진행하되, EU와 같이 AI 전반에 대한 법률을 정하여 규제해야 한다고 하였다.[31)]

표 6. AI 규제 활용을 위한 5가지 핵심 원칙[32)]

원칙	요약
안전성 · 보안성 · 견실성	AI 시스템은 AI 수명 주기 전반에 걸쳐 견고하고 안전하며 안전한 방식으로 작동해야 하며 위험은 지속적으로 식별, 평가 및 관리되어야 한다.
적절한 투명성과 설명 가능성	AI 시스템은 적절히 투명해야 하며 설명 가능해야 한다.
공평성	AI 시스템은 개인이나 조직의 법적 권리를 훼손하거나, 개인을 부당하게 차별하거나, 불공정한 시장 결과를 만들어서는 안된다.
설명 책임과 거버넌스	AI 시스템의 공급 및 사용에 대한 효과적인 감독을 보장하기 위해 거버넌스 조치가 마련되어야 하며, AI 수명 주기 전반에 걸쳐 명확한 책임 라인이 확립되어야 한다.
경쟁 가능성과 구제	적절한 경우, 사용자, 영향을 받는 제3자 및 AI 수명 주기의 행위자는 해롭거나 실질적인 해를 끼칠 위험을 초래하는 AI 결정 또는 결과에 대해 이의를 제기할 수 있어야 한다.

연세대학교
YONSEI UNIVERSITY

생성형 인공지능 의료기기에 대한 규제는 기본적으로 영국 의료기기 규정 2002(UK Medical Devices Regulations 2002)와 의료기기로서의 소프트웨어 및 인공지능에 대한 가이드라인(Software and artificial intelligence (AI) as a medical device, 2024.06)에 따른다.

또한 의약품 및 의료 제품 규제 기관(MHRA)은 FDA 및 Health Canada와 긴밀히 협력하여 의료기기의 AI에 대한 공통 원칙을 수립하고 있다. 이러한 협력은 최근 출판된 기계 학습 지원 의료기기에 대한 투명성: 지침 원칙(Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles)에서 분명하게 드러난다. 영국은 자체 규제 프레임워크를 형성하고 있지만 미국의 미래 AI 규정과 공통 요소를 공유할 가능성이 높을 것으로 예상된다.

영국 MHRA의 공식 블로그에 따르면 현재 MHRA에서는 ChatGPT와 Bard와 같은 대규모언어모델(LLM)을 의료목적으로 사용하는 경우에 아래와 같은 고려사항들을 제시하였다.[33)]

(1) 구체적으로 AI 제품이 의도된 목적을 규정해야 한다.

(2) 적절한 임상적 증거를 확보하고 정상적인 사용 조건에서 기기가 의도한대로 작동 한다는 것을 입증해야 한다.

(3) 오픈소스 LLM은 알려지지 않은 소프트웨어(SOUP)로 간주될 가능성이 높으며, 제조사는 이에 대한 위험을 식별하고 완화 하려는 조치를 해야한다.

(4) 모델의 의사결정 프로세스를 이해하고 이에 대한 설계 제어 및 신뢰성이 보장되어야 한다.

(5) 모델의 훈련과정에 포함될 수 있는 편향을 식별하고 관리할 수 있어야 한다.

(6) 대규모 언어모델의 잘못된 출력(환각)을 완화하기 위한 조치를 수행하고 이를 검증해야 한다.

연세대학교
YONSEI UNIVERSITY

한편 영국의 의약품 및 의료 제품 규제 기관(MHRA)에서는 2024년 5월부터 의료기기로서의 AI(AIaMD) 기술에 대한 규제 샌드박스인 AI Airlock을 시행 중이다.[34)]

AI Airlock 제도는 AI 활용 의료기기 관련 규제 위험을 이해하고 완화하기 위한 협업 이니셔티브로, AI 기술이 의료분야에 혁신을 촉진하는 만큼 새로운 안전 규제 대응을 지원하기 위한 목적으로 설계됐다. 또한 NHS 시설을 통해 AI 의료기기를 배포하며, MHRA 및 승인 기관, NHS, 기타 주요 파트너를 통합하며, 개발된 AI 활용 의료기기에서 발생할 수 있는 규제 문제를 실증 테스트를 통해 식별하여 규제 준수 과정을 간소화한다.

3.1.4. 유럽

유럽 연합(EU)은 2024년 3월 13일에 27개 회원국이 AI법안(Artificial Intelligence Act ,이하 AIA)을 채택함으로써 인공지능(AI)을 규제하는 데 중요한 이정표를 세웠다. 2021년 4월 유럽연합 집행위원회에서 처음 제안된 AI법안은 인공지능의 혁신적인 잠재력을 인식하고, 관련된 위험과 윤리적 문제를 해결할 필요성에서 출발했다. 이 법안은 2018년 AI 공동 계획(Coordinated Plan on AI)과 2020년 AI 백서(White Paper on AI)를 포함한 이전의 EU 이니셔티브를 바탕으로 만들어졌다.[35)]

세계 최초의 인공지능(AI)에 관한 포괄적인 법적 프레임워크인 유럽의 AI 법안은 인간 중심의 신뢰할 수 있는 AI 개발을 촉진하는 동시에, AI 시스템의 잠재적인 위험 요소로부터 개인의 건강, 안전 및 기본 권리를 보호하는 것을 목표로 하고 있다. 이 법안은 AI 시스템의 시장 출시, 서비스 제공 및 사용에 대한 조화된 규칙을 규정하고 있으며, 2024년 8월에 발효가 되었으며 허용 불가한 위험을 초래하는 AI 시스템과 같은 일부 예외를 제외하고는 24개월 이후인 2026년 8월에 전면 적용될 예정이다.[36)]

이러한 유럽 AI법안(AIA)은 특히 MDR(Medical Device Regulation), IVDR(In Vitro Diagnostic Regulation)과 같은 기존 유럽 의료기기/체외진단의료기기 규제와 별도로 고려를 해야 하는데 이는 기존 의료기기 규제에서는 의료 AI 응용 프로그램을 명확히 다루지 않기 때문이다. 또한 의료행위와 관련이 없는 단순 임상의사결정지원 기능과 같이, 의료기기에서 제공하는 모든 AI 응용 프로그램이 반드시 MDR 또는 IVDR의 범위에 포함되는 것은 아니기 때문에 이 유럽 AI법안은 MDR, IVDR과 별도로 유럽시장에 진입하고자 하는 인공지능 프로그램에 별도 적용될 수 있다.

연세대학교
YONSEI UNIVERSITY

유럽의회의원(Member of the European Parliament)에서는 2023년 6월 기존의 유럽 AI법안의 제안에 더해 범용 AI 시스템(GPAI)과 생성형 인공지능 모델을 규제하는 새로운 조항인 제28조b항을 도입했다. 유럽 AI법안에서 새롭게 추가된 제28조b항에서는, 범용 AI 시스템(GPAI)을 시장에 출시되거나 서비스되는 방식과 관계없이(오픈 소스 소프트웨어 포함) 공급자가 이미지 및 음성 인식, 오디오 및 비디오 생성, 패턴 감지, 질문 답변, 번역 등 일반적으로 적용 가능한 기능을 수행하도록 의도한 AI 시스템으로 정의하는데 GPAI 모델 공급자는 아래와 같은 의무를 준수해야 한다고 명시하고 있다.[36)]

표 7. 범용 AI모델 공급자의 의무[37)]

구분	주요 내용
시장 출시 또는 서비스 제공 전 준수 의무	파운데이션 모델은 독립형 모델, AI 시스템 또는 제품에 포함된 모델, 오픈 소스 라이선스 또는 서비스 형태로 제공되는 경우 등 모든 배포 방식에 관계없이, 본 조항의 요구 사항을 준수해야 한다. (a) 위험 식별 및 완화 - 독립 전문가 참여 등 적절한 설계, 테스트, 분석을 통해 건강, 안전, 기본권, 환경, 민주주의 및 법치에 미치는 위험을 개발 전과 개발 중 지속적으로 식별, 완화하며 문서화 해야한다. (b) 데이터 처리 및 편향 관리 - 파운데이션 모델에 사용되는 데이터셋은 적절한 데이터 거버넌스 조치를 거쳐야 하며, 데이터 소스의 적합성과 편향 가능성을 평가하고 필요한 경우 완화 해야한다. (c) 성능 및 보안 설계 - 모델 개발 단계에서 성능, 예측 가능성, 해석 가능성, 수정 가능성, 안전성 및 사이버 보안을 보장하도록 설계하며, 독립 전문가 참여를 통해 평가 및 테스트를 수행해야 한다.

연세대학교
YONSEI UNIVERSITY

구분	주요 내용
	(d) 에너지 효율 및 환경 기준 준수 - 모델은 에너지 소비, 자원 사용 및 폐기물 발생 감소를 목표로 설계되며, 에너지 효율성을 높이도록 개발한다. - 시스템의 라이프사이클 전반에 걸쳐 에너지와 자원 소비를 측정하고 기록할 수 있는 기능을 갖춰야 한다. (e) 기술 문서 및 사용 지침 작성 - 하위 공급자가 AI 법안을 준수할 수 있도록 기술 문서와 사용 지침을 작성한다(제16조 및 제28조(1) 준수). (f) 품질 관리 시스템 구축 - 품질 관리 시스템을 수립하고, 해당 조항의 요구 사항을 충족하는 실험적 방법을 적용할 수 있도록 한다. (g) EU 데이터베이스 등록 - 파운데이션 모델은 제60조에 명시된 EU 데이터베이스에 등록하며, 부속서 VIII, C 항의 지침을 준수한다.
기술문서 보관 의무	파운데이션 모델이 시장에 출시된 후 10년 동안 기술 문서를 국가 당국에 제공할 수 있도록 보관한다.

이에 더해 생성형 인공지능 모델을, 광범위한 데이터를 기반으로 대규모로 훈련된 AI 시스템 모델로, 복잡한 텍스트, 이미지, 오디오 또는 비디오와 같은 콘텐츠를 다양한 수준의 자율성으로 생성하도록 특별히 고안된 AI 시스템으로 정의하고 모델 공급자는 3가지 추가적인 의무를 준수해야 한다고 명시하고 있다.

연세대학교
YONSEI UNIVERSITY

표 8. AI 시스템 모델 공급자의 추가적인 의무 사항

구분	주요 내용
투명성 의무 준수	• 제52조(1)에 따른 투명성 의무를 준수해야 한다.
불법 콘텐츠 생성을 방지하기 위한 설계 및 개발	• 최신 기술 수준에 따라 모델을 설계하고 개발하며, 표현의 자유를 포함한 기본권을 침해하지 않도록 한다.
저작권법에 따른 훈련 데이터 공개	• 저작권법에 의해 보호되는 훈련 데이터의 사용에 대한 내용을 충분하고 상세하게 문서화하고 공개해야 한다.

유럽의 MDR, IVDR 규제와 유사하게 유럽 AI법안도 위험 기반 접근 방식을 따르며, 주요 초점은 허용할 수 없는 위험을 가진 특정 AI의 금지와 고위험 AI 시스템 및 범용 AI(GPAI) 모델에 대한 분류와 의무에 있다.(제1조 2b - e항). 유럽 AI 법안의 위험 기반 접근 방식의 개요 및 분류체계는 [그림 9]와 [표 9]에 설명되어 있다.[38) 39)]

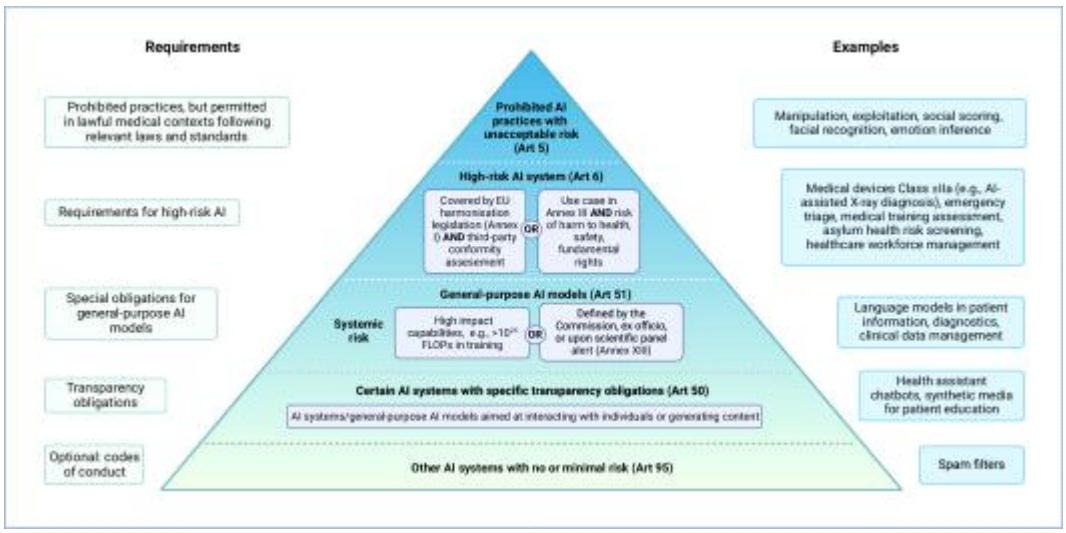

그림 9. EU AI 법안의 위험 기반 접근 방식의 도식화[39)]

연세대학교
YONSEI UNIVERSITY

표 9. EU AI 법안의 위험 기반 구분 및 주요 내용[38)]

구분	주요 내용
허용 불가 위험 (Unacceptable risk)	• 인간 기본권에 대한 명확한 위협으로 간주 되는 AI 시스템은 금지 - 사람의 인지 행동을 왜곡하는 AI 시스템, 정부나 기업이 시민에 대해 '사회적 점수평가(소셜 스코어링)'을 허용하는 시스템, 실시간 원격 생체 인식 시스템 등이 해당
고위험 (High risk)	• 고위험으로 분류된 AI 시스템은 위험 완화 시스템, 고품질 데이터 세트, 활동 기록, 상세 문서화, 명확한 사용자 정보, 인적 감독, 높은 수준의 견고성·정확성·보안성 등 엄격한 요건 준수 - 교통 등 시민의 생명·안전과 직결된 중요 인프라, 채용이나 대출 자격 여부 평가, 로봇수술의 AI 적용, 자율 로봇 실행, 수사 증거의 신뢰성 평가, 비자 신청 자동 심사 등 포함 • 모든 형태의 원격 생체 인식·식별 시스템은 고위험으로 분류되며, 법 집행의 목적으로 공공장소에서 원격 생체 인식 시스템을 사용하는 것도 원칙상 금지 - 실종 아동 수색, 테러 예방, 특수 범죄자/용의자 탐지 등과 같은 사안은 예외적으로, 사법 기관의 승인에 따라 제한적 조건 하에 사용 가능
투명성 위험 (Transparency risk)	• AI 사용의 투명성 부족과 관련된 위험을 의미하며, AI 법은 필요한 경우 사용자에게 정보를 제공하고 신뢰를 증진하기 위해 구체적인 투명성 의무를 도입 - 챗봇과 같은 AI 시스템 사용 시 사용자에게 기계와 상호 작용하고 있음을 명확하게 고지해야 하며, 서비스 제공자는 AI가 생성한 콘텐츠 여부를 식별 가능하도록 시스템을 설계 - 딥페이크 등 특정 AI가 생성한 콘텐츠에는 이러한 사실을 표기해야 하며, 감정인식이나 생체인식 기술이 사용될 경우 이러한 기술에 노출되는 사용자에게 고지 의무

연세대학교
YONSEI UNIVERSITY

구분	주요 내용
최소 위험 (Minimal risk)	• AI 기반 추천 시스템, 스팸 필터 등과 같이 현재 보편적으로 사용되는 AI 시스템은 대부분 이 범주에 해당하며, 시민의 권리와 안전에 대한 위험이 미미하므로 법적 의무가 부과되지 않음(기업들은 자발적으로 추가 행동강령 채택 가능)

위 표에 정리된 내용처럼 유럽 AI법안의 위험기반 분류체계의 기준은 AI 시스템이 초래할 수 있는 위험에 따라 구분된다.

특정 AI 시스템이 유럽 AI법안의 '고위험' 범주에 속할 수 있는 두 가지 방법이 있다고 해석된다. 첫 번째로, 제안된 유럽 AI법안의 부록 III에는 '고위험'으로 간주될 AI 시스템이 나열되어 있다. 대부분의 AI 시스템은 의료 AI와는 관련성이 적으며, 주로 고용, 법 집행, 사법 행정 등과 같은 분야에 집중하고 있다. 이러한 AI 시스템들은 '고위험' 용도로 사용될 수 있는 경우에 '고위험' AI 시스템 요구 사항을 준수해야 한다.

두 번째 방식으로는 유럽 AI법안의 위험 수준과 다른 NLF(신기술 조화 법률 프레임워크) 법률들 간의 연관성에 기반한다. 여기서 가장 중요한 NLF 법률은 **MDR(의료기기 규정)**이다. 간단히 말해, 유럽 AI법안은 특정 의료기기를 특히 위험한 것으로 간주하여 '고위험'으로 분류한다. 의료기기가 '고위험'인지 여부를 판단하기 위해, 유럽 AI법안은 MDR이 해당 기기에 대해 제3자 적합성 평가 절차가 필요한지 여부라는 하나의 특정 기준에 따라 해석할 수 있다.

'고위험'으로 분류되는 AI 시스템은 제안된 유럽 AI법안의 III장에서 규정된 일련의 실질적 요구사항을 따라야 하지만, 낮은 위험으로 분류된 AI 시스템은 최소한의 투명성 요구사항만 적용되거나 유럽 AI법안에 의해 규제되지 않는다. 실제로 이는 유럽 AI법안의 가장 중요한 요구 사항들이 '고위험'으로 분류된 AI 시스템에만 적용된다는 것을 의미한다고 볼 수 있다. 따라서, 의료 AI 시스템을 포함한 모든 AI 시스템의 위험 분류를 결정하는 것이 중요하며, '고위험'의 범주에 속하는지 여부를 판단해야 한다. 만약 '고위험'의 범주에 속하는 의료 AI 시스템일 경우 AIA의 요구사항이 적용되며, MDR(의료기기 규정)과 같은 다른 규정의 요구사항과 함께 적용된다고 판단할 수 있다.

연세대학교
YONSEI UNIVERSITY

두 법안은 유럽 AI법안 제6조(1)에서 정의된 '고위험' AI 시스템이라는 정의를 통해 상호 연결된다고 볼 수 있다. 간단히 말해, AI 시스템이 의료기기이거나(SaMD) 그 안전 구성 요소인 경우(SiMD), 그 의료기기가 MDR 하에서 제3자 적합성 평가를 받는 경우에만 유럽 AI법안 상의 '고위험'으로 간주 되기 때문이다. 예를 들어 MDR의 분류체계에 따라 Class I과 같이 저위험으로 분류되는 의료기기에 포함된 AI 시스템은 유럽 AI법안의 해석에서 '고위험' AI시스템으로 간주될 수 없다. 반대로 유럽 AI법안에서 '고위험'으로 분류된 의료 AI 시스템은 MDR의 분류체계에서 반드시 제 3자 적합성 평가를 받아야 하는 Class II 이상으로 분류된다고 볼 수 있기 때문이다. 따라서 의료 AI 시스템에 대해는 MDR의 위험 분류 기준을 자세히 평가해서 제품이 어떤 등급분류로 구분되는지 먼저 정의를 할 필요가 있다.

MDR에서 인공지능 또는 생성형 인공지능 기술에 대한 별도의 규제 사항은 특별히 명시되어 있지 않다. 다만 MDR의 Annex VIII 중에서 Rule 11과 Rule 22과 같이 몇몇 Rule은 생성형 AI가 적용된 의료기기의 위험도에 따른 분류를 결정하는 것과 관련이 있다.

<MDR Annex VII Rule 11>
진단 또는 치료 목적의 결정을 내리는 데 사용되는 정보를 제공하는 소프트웨어는 다음과 같이 분류된다.

- IIa 등급: 위 사용 목적을 사진 소프트웨어는 기본적으로 이 범주에 포함.
- IIb 등급: 결정이 개인의 건강 상태에 심각한 악화를 초래하거나 수술적 개입이 필요한 경우.
- III 등급: 해당 결정이 사망 또는 건강 상태의 비가역적 악화를 초래할 수 있는 경우.

생리적 과정을 모니터링하는 목적으로 설계된 소프트웨어는 기본적으로 IIa 등급에 속한다. 다만, 생명에 중요한 생리적 파라미터를 모니터링하며, 해당 파라미터의 변동이 즉각적인 위험을 초래할 수 있는 경우에는 IIb 등급으로 분류된다.
그 외 모든 소프트웨어는 I 등급으로 분류된다.

연세대학교
YONSEI UNIVERSITY

이와 같은 기준에 따르면, 대규모 언어모델(LLM)을 적용하여 의료영상을 분석하여 질병의 유무와 같은 결과가 포함된 판독문을 자동으로 생성해 주는 의료기기의 경우, 인공지능의 산출물이 의료진의 진단 또는 치료 목적의 결정을 내리는데 사용이 되므로 Class IIa에 속한다고 볼 수 있다. 앞서 논의 되었던 유럽 AI법안과 MDR의 위험 분류 정의에 따라 Class IIa에 속하는 의료기기는 유럽 AI법안에 따라 '고위험' AI 시스템으로 분류되며 법안에서 명시하는 요구사항을 준수해야 함을 알 수 있다.

그러나 모든 AI 시스템에 MDR의 위험 분류 규칙을 적용하기 전에, 해당 의료 AI 시스템이 반드시 의료기기로 분류되는 것은 아니라는 점을 확인해야 할 필요가 있다. 의료 AI 시스템이 의료기기가 아닌 경우, 유럽 AI법안의 부록 III에 있는 목록에 포함되어 있어야 '고위험'으로 분류될 수 있다. 부록 III에서 '고위험'으로 분류된 AI 시스템 목록에 해당 되는 의료 AI 시스템은 매우 적을 것으로 예상되므로, 앞서 언급한 유럽 AI법안과 MDR 간의 연관성이 가장 중요한 기준이 된다고 볼 수 있다.

따라서 대부분의 경우, AI 의료 시스템이 유럽 AI법안에서 '고위험'으로 분류되는지를 결정하는 실질적인 기준은 해당 시스템이 MDR 적합성 평가 절차에서 제3자의 적합성 심사를 받는지의 여부와 관계 있다는 주장이 있다.

일반적으로 제조업체는 Class I 기기에 대해 자체 평가를 할 수 있지만, 고위험 등급에 해당하는 기기들은 인증 기관의 참여가 포함된 적합성 평가 절차를 거쳐야 한다. 따라서 Class IIa, Class IIb 또는 Class III에 속하는 기기들은 항상 인증 기관이 평가에 참여하게 된다.[39)]

AI 의료기기의 경우, Class I 기기에 대해 제3자 인증 기관이 적합성 평가에 참여하지 않는다는 일반적인 규칙에 따라 이러한 AI 의료기기들은 유럽 AI법안 상에서 '비-고위험' AI 시스템으로 간주 된다고 볼 수 있다. 반면, MDR Class IIa, IIb 또는 III에 속하는 AI 의료기기들은 유럽 AI법안 상에서 '고위험' AI 시스템으로 분류되며, 제공자는 AIA 제III장에서 규정된 요구사항을 준수해야 한다고 판단할 수 있다.

연세대학교
YONSEI UNIVERSITY

3.1.5. 호주

호주 TGA에서는 2024년 5월 공식 홈페이지에 Artificial Intelligence (AI) and medical device software에 대한 TGA의 규제 관점이 담긴 내용을 제공하였다.
제공된 규제관점에 따르면, 최근 활발하게 개발 중인 ChatGPT, GPT-4, Gemini(구 Bard), Claude 등과 같은 텍스트 기반 AI 제품을 언급하며 생성형 AI를 포함하는 소프트웨어(대규모 언어 모델(LLM), 텍스트 생성기, 멀티모달 생성형 AI 등)는 의료기기 정의에 부합하는 경우 의료기기로 규제된다고 정의하고 있다.

또한 임상 및 기술에 대한 요구사항으로 생성형 AI(예: LLM)를 사용하는 제품은 다른 의료기기와 동일한 수준의 안전성, 신뢰성 및 성능을 입증해야 하며 고위험 제품의 경우에는 더 엄격한 임상 및 기술 근거가 요구된다고 명시했다.
생성형 AI(LLM 등)의 사용이 제품과 관련된 의료 목적 또는 주장이 없는 경우 또는 해당 제품이 1989년 치료용품법(Therapeutic Goods Act 1989) 제41BD항에 정의된 의료기기의 정의를 충족하지 않는 경우에는 의료기기로 분류되지 않는다고 설명하였다.[40)]

3.1.6. 국제의료기기규제기관포럼(IMDRF)

IMDRF에서는 영국, 캐나다, 미국의 규제기관이 개발한 초기 3자 문서를 기반으로 Good Machine Learning Practice(GMLP)에 대한 지침 원칙 초안을 작성중에 있다. 또한 AIWG(AI Working Group)이 생성형 인공지능과 대규모 언어모델의 요구사항을 다루고 있고 이를 다른 IMDRF의 문서와의 조화시키기 위한 작업을 진행하고 있다고 25차 정기 회의에서 밝힌 바 있다.[4)]

연세대학교
YONSEI UNIVERSITY

4. 고찰

본 연구에서는 생성형 인공지능을 의료분야에 적용 시 고려해야 하는 사항들을 각국 주요 기관의 규제 관점을 중심으로 분석해 보았다.

생성형 인공지능은 때때로 환각(Hallucination)이 포함된 결과를 제공하는데, 이러한 환각은, 입력 데이터 또는 사실 정보에 근거하지 않은 출력 결과를 생성하는 것을 말한다. 이러한 잘못된 정보는 진단, 치료 또는 권장 검사와 관련에도 잘못된 영향을 미칠 수도 있다. 임상적으로 풍부한 경험이 없는 사람에게는 이러한 잘못된 출력이 높은 수준의 확신처럼 전달이 되거나 진실로 쉽게 받아들일 수 있는데, 이는 치명적인 위험을 발생시킬 수 있는 잠재적인 위험 요소이다.

또 다른 문제는 생성형 인공지능을 사용하는 동안 의료지식에 대한 편견이 임상적 의사결정, 환자 결과 및 의료 형평성에 영향을 미칠 수 있다는 것이다. 특정 인구통계적 특성이 반영되지 못한 데이터, 특정 치료에 대한 과도한 강조 또는 구식 의료 관행과 같은 편향이 훈련 데이터에 포함되어 있는 경우, 생성형 인공지능은 이러한 편향을 실수로 학습하여 출력에서 생성할 수 있다. 이러한 편향된 출력은 잘못된 진단 또는 최적이 아닌 치료 권장 사항으로 이어져 환자에게 해를 끼치거나 적절한 치료를 지연시키는 부작용을 초래할 수 있다.

생성형 인공지능의 의료분야 적용은 윤리적 문제를 야기 하며, 이에 대한 규제 프레임워크가 필요할 수 있다. 생성형 인공지능 모델을 학습하고 의료기기로 사용하기 전에 투명성, 책임성, 공정성과 같은 문제들이 사전에 다루어져야 하며, 윤리적 결함을 예방할 수 있어야 한다. 예를 들어, 의료 전문가와 환자들은 인공지능이 의사 결정 과정에 참여하고 있다는 사실을 인지해야 하며, 인공지능의 권고 사항에 대한 설명을 제공 받아야 한다.

생성형 인공지능 기술의 특성 및 의도된 사용 목적에 따른 맞춤형 성능지표를 설정하고 성능에 대한 검증이 이루어져야 한다. 이러한 성능지표는 기존 의료제품 또는 인공지능 의료기기의 성능지표를 참고할 수는 있으나 생성형 인공지능 기술의 특성이 맞는 별도의 성능지표의 설정이 필요하다.

연세대학교
YONSEI UNIVERSITY

표 10. 대표적인 생성형 인공지능 언어모델의 평가지표 예시[12)]

모델	평가지표	지표설명
언어모델	PPL(Perplexity)	• PPL은 당혹감, 혼란의 의미로 모델이 정답을 결정할 때 얼마나 헷갈렸는가를 나타내는 지표로 문장이 완성될 때까지 선택된 토큰들의 누적된 확률을 기반으로 계산한 값 • 값이 낮을수록 모델이 덜 헷갈린 상태로 확신을 가지고 답을 냈다는 의미
	BLEU(Bilingual Evaluation Understudy)	• 문장의 길이와 단어의 중복을 고려하여 정답문장과 예측문장 사이의 겹치는 정도를 계산하는 지표로 사용 • BLEU은 1에 가까울수록 그 정확도가 높음
	SSA(Sensibleness and Specificity Average)	• 구글에서 발표한 자율 발화 모델에 대한 평가지표로 사람이 직접 자율 발화 모델과 대화를 하며 점수를 평가 • Sensibleness는 사람의 말에 대한 봇의 답변이 합리적인지를, Specificity는 해당 답변이 단답형이 아닌 구체적인 답변인지를 나타냄
	ROUGE(Recall-Oriented Understudy for Gisting Evaluation)	• 주로 문서 요약(Text Summarization) 의 품질을 평가하는 데 사용되는 지표로 단어나 구의 재현율 측정 1에 가까울수록 그 정확도가 높음
	휴먼평가지표	• AI 답변의 정확성 등을 평가하기 위해 별도의 평가지표를 만들어 사람이 직접 평가하는 방식
	기타	• (자체보유 성과지표 솔루션) AI 업체 등에서 자체 보유하고 있는 성과지표가 있는 경우 사용 • (국내외 AI평가지표)국내외 논문, 각종 AI리더보드 대회(H6) 등에서 사용하는 객관적이고 공신력 있는 지표 제시하여 사용 • 분류모델에서 사용하는 여러 가지 지표

연세대학교
YONSEI UNIVERSITY

또한 유럽의 MDR, IVDR과 같이 기존 의료기기의 규제는 각 규제기관의 의료기기법과 관련된 규제 의무가 요구되었다면 한국의 디지털의료제품법, 유럽의 AI법안(AIA)과 같이 인공지능 및 생성형 인공지능 제품에 적용되는 별도의 법안이 하나둘씩 만들어지고 있다. 이에 따라 생성형 의료기기를 공급하고자 하는 제조사는 별도의 인공지능 법안의 요구사항 또한 준수해야 하며 이는 기존 의료기기법과 관련된 규제 의무사항과 별도의 요구사항이 될 수도 있다.

이처럼 제조사는 생성형 인공지능이 의료기기로 개발됨에 따라 예상할 수 있는 여러 고려사항을 제품 개발 시 고려해야하며, 기존 규제제도로는 충분히 다룰 수 없는 생성형 인공지능의 특성으로 인해 각국에서는 기존의 인공지능 규제 방안과는 별도의 규제 프레임을 수립하고자 하는 노력을 하고 있으며 각 규제기관 및 단체에서 언급한 내용을 정리한 표는 아래와 같다.

연세대학교
YONSEI UNIVERSITY

표 11. 생성형 인공지능을 활용한 의료기기 개발 시 고려사항 및 대응전략

고려사항		세부내용 및 대응전략
생성형 인공지능의 의료기기 규제 대상 여부 확인		생성형 인공지능이 1) 환자별 분석을 수행하고 질병 또는 상태의 진단, 치료, 완화, 치료 또는 예방에 사용하기 위해 의료 전문가에게 특정 결과 또는 지침을 제공하거나, 2) 환자별 분석을 수행하고 환자, 간병인 또는 의료 전문가가 아닌 기타 사용자에게 환자별 진단 또는 치료 권장 사항을 제공하는 것과 같이 생성형 인공지능이 기기의 결과물이나 성능이 필수적인 역할을 하는 경우 의료기기 규제 대상이 됨을 고려
위험기반 접근 방식을 적용한 시판 전, 시판 후 요구사항	위험기반 접근 방식의 목적	생성형 인공지능의 고유한 특성인 환각(Hallucination)의 발생 위험과 가능성, 파운데이션모델(Foundation Model)에 대한 정보획득 및 통제의 어려움과 같이, 생성형 인공지능의 사용으로 인해 발생할 수 있는 위험을 완화하기 위하여 생성형 인공지능 기기의 설계 초기 단계부터 위험 통제 전략을 통합하고, 실시간 모니터링 및 피드백 시스템을 통해 출력과 성능을 지속적으로 관리하기 위함.
	모델 아키텍처와 설계 제어	1. 설계 단계에서의 통제: - 모델 개발 초기부터 설계 사양에 대한 상세 문서화 필요. - 모델의 매개변수와 출력 통제 전략 포함하여 모델이 작동할 수 있는 범위를 명확히 정의. - 모델의 훈련과정에 포함될 수 있는 편향(Bias)를 식별하고 관리.
		2. 사용 시 제어 가능성 강화: - 모델이 특정 작업(예: 진단 보조, 치료 계획)만 수행하도록 기능을 제한하는 추가 레이어 설계. - 인간 개입 허용 수준 (Human-in-the-Loop)을 모델과 함께 설정하여, 모든 결정 과정에서 의료 전문가가 확인 및 조정할 수 있도록 설계.

연세대학교
YONSEI UNIVERSITY

고려사항	세부내용 및 대응전략
데이터 사용 및 훈련 관리	1. 데이터 거버넌스 강화: - 모델 훈련에 사용된 데이터의 출처, 품질, 편향 여부를 사전에 평가. - 데이터 라벨링 및 검증을 통해 모델 출력의 예상 범위를 구체화.
	2. 훈련 데이터의 적합성 검토: - 훈련 데이터의 다중 검증 과정을 도입하여, 모델이 훈련 데이터의 특성으로 인해 예기치 못한 행동을 보이지 않도록 관리. - 의료 도메인에 특화된 추가적인 훈련(Fine-tuning)을 통해 불필요한 일반화 방지.
실시간 모니터링과 동적 제어	1. 실시간 추적 및 모니터링: - 모델의 작동 중 데이터를 지속적으로 모니터링하여, 예외적인 출력(환각, 편향 등)을 자동으로 탐지하는 시스템을 구축. - 모델의 비정상적인 작동이 감지되면 자동으로 출력을 차단하거나 제한하는 출력 오류 자동 탐지 및 차단 알고리즘 도입.
	2. 동적 출력 관리: - 출력 데이터를 단계별로 필터링하는 알고리즘을 구현하여, 기기 사용자(의료 전문가)에게 제공되기 전에 불확실한 정보를 제거. - 특정 상황에서 모델 출력의 신뢰도를 평가하는 불확실성 측정 시스템 적용.
피드백 메커니즘	- 의료 전문가 및 사용자가 모델의 오류나 비정상적인 출력을 보고할 수 있는 피드백 메커니즘을 제공. - 피드백을 통해 모델의 취약성을 파악하고, 업데이트 및 개선 과정에 반영.

연세대학교
YONSEI UNIVERSITY

고려사항		세부내용 및 대응전략
	성능 평가	1. 성능 평가: - 모델의 성능을 다양한 테스트 환경에서 평가하여 제어범위를 사전에 정의. - 기술의 특성 및 의도된 사용목적에 따른 맞춤형 성능지표를 설정하여 평가 (출력의 신뢰도, 예측 가능성, 반복 가능성 등을 측정하는 지표(예: Perplexity, BLEU 점수 등)를 설정하고 모니터링)
		2. 추가적인 성능지표 적용 - 자율성 수준, 투명성, 설명 가능성과 관련된 파운데이션 생성형 인공지능 모델의 특성을 정확히 파악하는 데 도움이 되는 질적 성능 평가 방법론을 적용.
	시판 후 모니터링	- 시판 후 실사용 데이터를 기반으로 모델의 안전성과 효과를 평가하는 Post-market Surveillance 시스템 도입.
생성형 인공지능 제품의 품질관리 평가제도		미국의 Pre-Cert, 한국 디지털의료제품 법의 우수관리체계 인증과 같이, 제조업체의 품질관리 역량을 평가하고, 일정 범위 내에서 자율적인 변경 및 품질 관리를 가능하게 하는 제도의 도입에 대한 고려
대규모 학습 데이터 소스 또는 모델 등록		파운데이션 모델과 같이 대규모 학습데이터가 필요한 생성형 인공지능 모델을 규제기관의 데이터베이스에 등록하여 관리하는 것을 고려
기술문서 및 사용지침		생성형 인공지능 모델 공급자는 하위 공급자가 의료기기 규제 및 별도 인공지능 법안의 요구사항을 준수할 수 있도록 기술문서와 사용지침을 제공하는 것을 고려
별도 인공지능 법안 적용		한국의 디지털의료제품법, 유럽의 AI Act와 같이 인공지능 및 생성형 인공지능 기술에 대한 별도 규제 법안의 적용여부를 고려

연세대학교
YONSEI UNIVERSITY

5. 결론

생성형 인공지능은 진단 보조, 치료 계획 수립, 의료 교육, 환자 데이터 분석 등 다양한 의료분야에서 기존의 한계를 극복하여 새로운 의료적 개입을 할 수 있으며, 생성형 인공지능 기반 의료기기는 방대한 데이터를 학습하여 복잡한 문제를 해결하거나 새로운 정보를 생성할 수 있는 특징을 지니며, 의료 현장에서의 효율성을 크게 향상시킬 것으로 기대하고 있다. 그러나 기술의 발전 속도가 빠르고 적용 범위가 방대한 이유로 생성형 인공지능 기술을 의료분야에 적용할 때 기존의 의료기기 규제체계로는 적용에 한계가 있다. 생성형 인공지능은 대규모 언어 모델(LLM)과 같은 파운데이션 모델을 기반으로 설계되며, 높은 자율성과 복잡성을 지니는 특성상 기존의 하드웨어 중심 의료기기 규제 체계로는 안전성과 신뢰성을 충분히 보장하기 어렵다. 특히, 생성형 인공지능의 출력은 불확실성, 편향, 환각(hallucination) 등의 문제를 동반할 가능성이 있으며, 이는 환자의 건강과 안전에 직접적인 영향을 미칠 수 있다. 또한, 이러한 기기는 훈련 데이터의 품질, 알고리즘의 설계, 실사용 환경에서의 지속적 업데이트 및 성능 진화와 같은 특수성을 지니기 때문에 새로운 규제 접근 방식이 요구된다.

현재 미국, 유럽을 비롯한 각 규제기관에서는 생성형 인공지능 의료기기의 안전성과 유효성을 보장하기 위해 위험 기반 접근법, 실사용 데이터 활용, 투명성 및 설명가능성 강화, 별도의 인공지능 규제 법안 등의 규제 방안을 제시하고 있으며 기존 법률을 생성형 인공지능 의료기기의 적용에 맞춰 보완하고 있다. 또한 IMDRF와 같은 국제 규제 조화기구를 통해 통합된 규제 프레임워크를 제시하려는 작업을 진행중에 있다.

본 연구에서는 생성형 인공지능 의료기기의 특성과 이를 둘러싼 규제상의 도전 과제를 고찰하고, 주요국의 규제 동향을 분석하여 규제 설계 시 고려해야 할 주요 요소를 도출하고자 한다. 이를 통해 생성형 인공지능 기반 의료기기의 안전성과 유효성을 보장하면서도 제조사의 제품 개발 과정을 촉진할 수 있는 규제 프레임워크를 제안하는 데 목적이 있다. 그러나 현재까지 생성형 인공지능에 대한 각 규제기관별 세부 요구사항 또는 가이드라인이 명확하게 수립되지 않은 관계로 규제기관별 비교 분석을 수행 하는데에는 한계가 있었다. 또한 조사된 각 주요기관들의 규제 프레임워크 내용이 다소 방법론적인 수준에 머물러 있어 실제 생성형 인공지능 의료기기가 시장에 출시되기 위해서는 많은 도전적인 과제들을 맞닥뜨리게 될 것으로 보인다.

연세대학교
YONSEI UNIVERSITY

이에 대해 한국 식약처에서는 세계 최초로 생성형 인공지능 의료기기에 대한 구체적인 지침(생성형 AI 기반 디지털 의료기기 허가·심사 규제 가이드라인)을 2025년 상반기에 마련하여 국내 생성형 인공지능 의료기기의 규제 한계에 대응할 예정이다. 이를 필두로 각 주요기관들의 구체적인 지침이 마련될 것으로 기대하며 지속적인 규제 현황 모니터링을 통하여 생성형 인공지능 의료기기의 개발 과정에 도움이 될 수 있도록 관련 연구를 지속적으로 수행할 예정이다.

본 연구를 통해 생성형 인공지능 기술을 활용하여 의료기기를 제조하고자 하는 국내 업체들이 제품의 개발단계에서부터 생성형 인공지능 관련 규제 고려 사항을 적용하여 제품개발 과정을 지원할 수 있기를 기대 한다.

연세대학교
YONSEI UNIVERSITY

참고 문헌

1. 공공경제_2023 Vol.15. AI산업 경쟁력 강화를 위한 생성형AI 규제 방향(전문가 ViewII)
2. Precedence Research. 2023.07. Generative AI In Healthcare Market Size, Share, and Trends 2024 to 2034
3. 식품의약품안전처 보도자료. 2024.03. 식약처,'빅테크 기업 규제혁신 프로그램 2024' 착수
4. RAPS.org. 25th session of the IMDRF: Regulators offer updates on new working group documents
5. Peng Y, Zhang Y, Wang L. Artificial intelligence in biomedical engineering and informatics: an introduction and review. Artif Intell Med. (2010) 48:71-3. doi:10.1016/j.artmed.2009.07.007
6. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Trans Med. (2015) 7:283rv3. doi:10.1126/scitranslmed.aaa3487
7. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med. 2019; 380(14):1347-1358. PMID: 30943338.
8. Shimizu H, Nakayama KI. Artificial intelligence in oncology. Cancer Sci. 2020; 111(5):1452-1460. PMID: 32133724.
9. Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. Cancer Inform. 2007; 2:59-77. PMID: 19458758.
10. Erickson BJ, Korfiatis P, Akkus Z, Kline TL. Machine learning for medical imaging. Radiographics. 2017; 37(2):505-515. PMID: 28212054.
11. Hu W, Cai B, Zhang A, Calhoun VD, Wang YP. Deep collaborative learning with application to the study of multimodal brain development. IEEE Trans Biomed Eng. 2019; 66(12):3346-3359. PMID: 30872216.
12. 식품의약품안전처. 식의약 분야 인공지능(AI) 도입 · 적용을 위한 AI 사업 실무가이드.
13. 한국저작권위원회. 2023. 생성형 인공지능(Generative AI) 산업 현황 보고서.
14. TTA저널, 207, 37. 2023.6.30. 초거대 AI와 생성형 인공지능.

연세대학교
YONSEI UNIVERSITY

15. 한국수출입은행 해외경제연구소. VOL.2023-이슈-9. 생성형 인공지능(Generative AI)으로 인한 인공지능 혁명 및 산업 변화.
16. 국회도서관. FACT BOOK 2023-5호 통권 제105호. 초거대 AI 한눈에 보기.
17. 소프트웨어정책연구소. 2023.02. 초거대언어모델의 부상과 주요이슈 : ChatGPT의 기술적 특징과 사회적·산업적 시사점.
18. 한국지능정보사회진흥원. 2023.04. 대규모 언어모델 기반의 공공분야 초거대 AI 도입방향.
19. Wayne Xin Zhao, Kun Zhou, Junyi Li. 2023.03. arXiv:2303.18223v15. A Survey of Large Language Models.
20. 한국바이오결제연구센터. 브리프193. 2024.10. 생성형 AI, 헬스케어 산업의 미래.
21. Aysen Çeliktaş. Medium post. 2024.02. Overview of GAN Use in The Medical Field
22. Islam J, Zhang Y. 2020. Brain informatics 7. “GAN-based synthetic brain PET image generation.”
23. Tian M, Song K. 2021. IEEE Access 9. “Boosting magnetic resonance image denoising with generative adversarial networks.”
24. Dong X, Lei Y, Wang T, Thomas M, Tang L, Curran WJ, Yang X. 2019. Medical physics 46.5. “Automatic multiorgan segmentation in thorax CT images using U-net-GAN.
25. FDA. FDA-Approved Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices: An Updated Landscape.
26. Lin Yang, Shawn Xu, Andrew Sellergren. 2024.05. arXiv:2405.03162v1. Advancing Multimodal Medical Capabilities of Gemini.
27. Kyu Hong Lee, Ro Woon Lee, Ye Eun Kwon. 2023.12. MDPI. Validation of a Deep Learning Chest X-ray Interpretation Model: Integrating Large-Scale AI and Large Language Models for Comparative Analysis with ChatGPT.
28. FDA. The Software Precertification (Pre-Cert) Pilot Program: Tailored Total Product Lifecycle Approaches and Key Findings.
29. FDA EXECUTIVE SUMMARY FOR THE DIGITAL HEATH ADVISORY COMMITTEE MEETING. 2024.11.20.-21. Total Product Lifecycle

연세대학교
YONSEI UNIVERSITY

Considerations for Generative AIEnabled Devices.
30. NIST-AI-600-1. 2024. Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile.
31. GOV.UK. 2021. National AI Strategy.
32. GOV.UK. 2024. Impact of AI on the regulation of medical products.
33. GOV.UK. 2023. MedRegs. Large Language Models and software as a medical device.
34. GOV.UK. 2024. MHRA launches AI Airlock to address challenges for regulating medical devices that use Artificial Intelligence.
35. European Parliament. 2024. European Union Artificial Intelligence Act
36. KISTEP. 2024. EU 인공지능(AI) 규제 현황과 시사점
37. European Commission. 2024. EU AI Act Overview.
38. KISTEP. 2024. 과학기술&ICT 정책 기술 동향
39. Felix Busch, Jakob Nikolas Kather, Christian Johner. 2024. npj;digital medicine. Navigating the European Union Artificial Intelligence Act for Healthcare.
40. TGA. 2024. Artificial Intelligence (AI) and medical device software

연세대학교
YONSEI UNIVERSITY

ABSTRACT

Regulatory Trends and Response Strategies for Generative AI-Based Medical Devices Across Major Regulatory Agencies

This study aims to investigate the regulatory trends of major regulatory agencies for generative AI-based medical devices applied in the healthcare field and analyze them to derive key considerations for regulatory framework design.

Generative AI, characterized by its ability to generate new information and solve complex problems based on vast datasets, can play a transformative role in various medical fields, including diagnostic assistance, treatment planning, medical education, and patient data analysis. However, due to the unique characteristics of this technology, existing medical device regulatory frameworks have limitations in ensuring safety and reliability, necessitating the development of new regulatory frameworks.

This research examines the unique features and regulatory challenges of generative AI-based medical devices and provides an in-depth analysis of regulatory trends presented by major countries, including the United States, Europe, and South Korea, as well as by international regulatory harmonization bodies such as IMDRF.

Key regulatory agencies emphasize approaches such as risk-based frameworks, enhanced transparency, the utilization of real-world data (RWD/RWE), performance validation, and reinforced post-market monitoring to mitigate potential risks associated with generative AI, such as hallucination, data bias, and output uncertainty.

Additionally, countries are enacting separate AI-related legislation or amending existing laws to ensure the safety and efficacy of generative AI-based medical devices. Notably, South Korea's Ministry of Food and Drug Safety (MFDS) is preparing the world’s first regulatory guidance specialized for the approval and review of generative AI-based medical devices.

연세대학교
YONSEI UNIVERSITY

This study aims to assist domestic manufacturers of generative AI-based medical devices in identifying key regulatory considerations during the development phase, thereby supporting efficient product development and regulatory approval processes.

Furthermore, by proposing a regulatory framework that ensures the safety and efficacy of generative AI-based medical devices while promoting technological innovation, this study is expected to contribute to the sustainable development of the digital healthcare industry.

Key words : Artificial Intelligence, Generative Artificial Intelligence