

ISSUE REPORT

연합학습 기반 신약개발 플랫폼 FDD



- 신약개발 패러다임 변화
- K-MELLODDY 사업
- K-MELLODDY 추진 전략 및 체계
- 성과 활용 및 기대효과
- 맺음말

글로벌 주요 동향

- ◆ 유럽, 실사용데이터(RWD) 전자 카탈로그 출시
- ◆ 유럽, 의약품 접근성 개선 규정 승인
- ◆ 호주, 첨단 기술에 대한 신규 의료기술평가(HTA) 수립 권고
- ◆ 덴마크, 임상시험·체외진단연구 동시 승인 프로세스 개시
- ◆ 영국, AI 규제 유연성 확대
- ◆ 프랑스, 의약품 부족 대응 신규 로드맵 발표
- ◆ 파키스탄, 의약품 임상시험 관리기준 지침 변경 검토
- ◆ ICH, 세포·유전자 치료제 및 RWD 사용 지침 개발계획

※ KPBMA FOCUS는 제약바이오 산업계의 최근 이슈를 발굴하고, 전문가의 시선으로 주요 동향을 깊이 있게 분석하였습니다.

(본 호부터는 제호를 글로벌이슈파노라마에서 KPBMA FOCUS로 변경하여 발간합니다.)

연합학습 기반 신약개발 플랫폼 FDD

김화종 한국제약바이오협회 AI신약융합연구원장

■ 신약개발 패러다임 변화

- 2023년 3월 글로벌 회사인 엔비디아(NVIDIA) 社가 신약개발에 뛰어들면서 신약개발을 위한 생성형 AI 플랫폼 'BioNeMo'를 소개한 이후, 이와 유사한 것들을 모두 다 적용하기 위해 미국의 주요 제약 회사들이 진행 중이던 파이프라인들을 멈추고 AI가 도출하는 결과를 보고 파이프라인 순위를 바꾸려고 할 정도로 AI의 파급력이 커지고 있음
- AI 기반의 신약개발 모델이 성공하기 위해서는 충분한 학습데이터가 확보되어야 함
 - 제약바이오 분야의 데이터는 이미지나 텍스트 같은 방대한 공개 데이터와 달리, 실험으로 확보할 수 있는 학습데이터 수가 적고, 특히 다른 기관에서 생산한 데이터를 공동 활용하기가 어려움
- 2024년 3월 11일 한국제약바이오협회 AI신약융합연구원은 "연합학습 기반 신약개발(FDD, Federated Learning based Drug Discovery) 가속화 프로젝트" 사업에 최종 선정되어 4월부터 세부 과제 기획, 공모·선정 등 본격적으로 추진할 예정임¹⁾
- FDD의 최종 목표는 "신약개발을 위한 AI 모델 학습에 필요한 데이터를 기관 간에 안전하게 공유할 수 있는 플랫폼을 확보"하는 것임. 본 고에서는 FDD의 핵심 개념을 소개하고자 함

■ K-MELLODDY 사업

▶ 사업 개요

- K-MELLODDY 사업은 2019~2022년 수행된 EU의 MELLODDY(Machine Learning Ledger for Drug Discovery) 사업을 모델로 기획됨
 - EU MELLODDY 사업은 암젠, 아스텔라스, 아스트라제네카, 바이엘, 안센, 머크, 노바티스 등 10개 제약 회사와 5개 IT 회사, 2개 대학이 참여하여 각 제약 회사의 실험 데이터를 연합학습(Federated Learning)²⁾을 이용하여 안전하게 공유하는 것을 실증한 사업임

1) 보건복지부 보도자료, "연합학습기반 신약개발 가속화 프로젝트 사업단장에 한국제약바이오협회 김화종 원장 선정", 2024.3.11

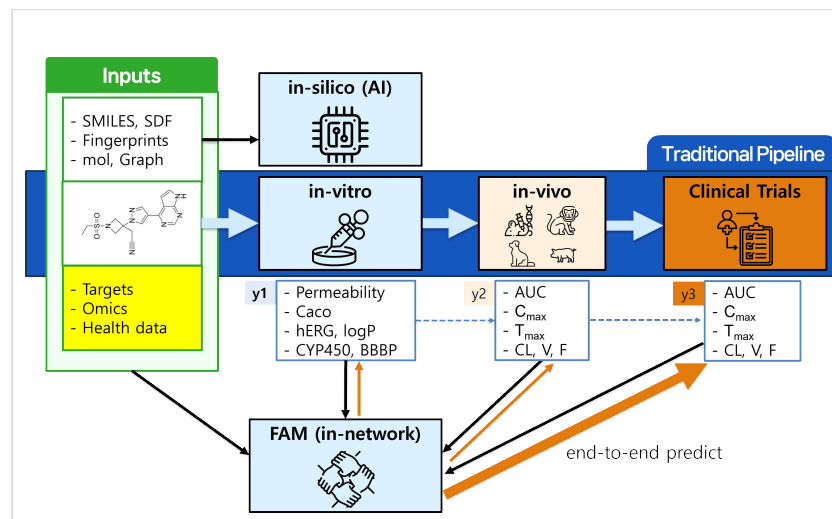
2) 연합학습(Federated Learning): 여러 참가자가 각자의 데이터를 직접 공개 및 공유하지 않고, 각자의 로컬 AI 모델 학습 결과 파라미터(가중치)만 공유하여 참가자 전체의 AI 모델의 성능을 공동으로 향상시키는 방법으로 구글 社가 2017년에 제안

- 본 사업을 통해 신약개발 과정에서 유용하게 사용할 수 있는 연합학습 기반 신약개발 FDD (Federated Drug Discovery) 플랫폼을 구축하고 약물 흡수, 분포, 대사, 배설, 독성(ADMET, Absorption, Distribution, Metabolism, Excretion, Toxicity) 예측 솔루션을 만들 것이며 이 예측 솔루션을 "FAM (Federated ADMET Model)"이라고 명명함
- 이 사업에서는 제약 회사와 같은 ADMET 데이터 소유자 20개 내외, 플랫폼 개발자 1개, AI 모델 공급자 15개 내외 등 산·학·연·병이 과제에 참여하도록 지원할 계획임

▶ FAM 솔루션

- 신약 후보 물질의 첫 검증 과정인 시험관내(in-vitro) 실험을 통해 후보물질의 ADMET 값을 측정하고 있으나, 이 결과만으로 생체내(in-vivo) 및 임상시험 통과를 보장하지 못하는 경우가 많으며 이는 in-vitro 실험 방식의 한계임
- 신약개발 가속화 연구보고서에 의하면 북미·유럽의 신약개발 R&D 비용 중 22%를 ADMET 검증이 차지³⁾하고 있으며 ADMET는 임상시험의 성공률을 결정짓는 가장 중요한 요소임
- 최근 in-vitro 실험에 필요한 시간과 비용을 줄이기 위해서 가상(in-silico) 실험으로 대체하는 노력을 하고 있으나, 현재 AI 성능과 적용 수준의 한계로 임상시험 통과를 보장하기 어려움
- FAM 솔루션은 현재의 AI 솔루션과 달리, End-to-End⁴⁾ 예측 모델로 동작하며 최초 측정된 in-vitro 데이터를 일련의 시그널(신호)로 임상시험 결과를 직접 예측하는 모델임
- FAM 솔루션은 in-vitro 실험의 목적변수⁵⁾(그림1, Y1)를 예측하는 것이 최종 목표가 아니라, 미래 임상시험 결과(그림1, Y3)를 예측하는 모델임

<그림 1> FAM: 연합학습 기반 ADMET 예측 솔루션



출처: 한국제약바이오협회 AI신약융합연구원

3) Kraljevic S, Stambrook PJ, Pavelic K, EMBO Reports, Accelerating drug discovery, 2004.9.
 4) End-to-End: 입력에서 출력까지 도달하기 위한 복잡한 파이프라인 없이 하나의 신경망으로 입력받은 값에 대한 출력값을 찾아낼 수 있는 모델을 의미
 5) 목적변수(target variable): 독립변수(independent variable)들의 영향으로 나온 결과값, 기계학습에 있어서 예측하고 싶은 정보

- FAM 솔루션은 신약개발 과정에서 생산된 in-vitro, in-vivo, 임상시험(in-human Clinical Trials) 데이터를 모두 사용하는 플랫폼 기반 솔루션이며, 이에 필요한 각 기관의 데이터를 공개하지 않아도 되도록 연합학습 방법을 도입함
- FAM 솔루션을 구현하기 위해서는 제약 회사, 연구소, 병원, 규제 기관, 임상시험수탁기관(Contract Research Organization, CRO), 대학 등 다양한 기관에 분산 저장된 데이터 연합이 필요함
 - 향후 in-vitro 실험값뿐 아니라, 타겟 정보, 개인 정보를 추가로 사용함으로써 약물의 특정 타겟에 대한 약효나 적용 환자 그룹별 ADMET 상세 예측 모델로 확장이 예상됨

■ K-MELLODDY 추진 전략 및 체계

▶ 추진 전략

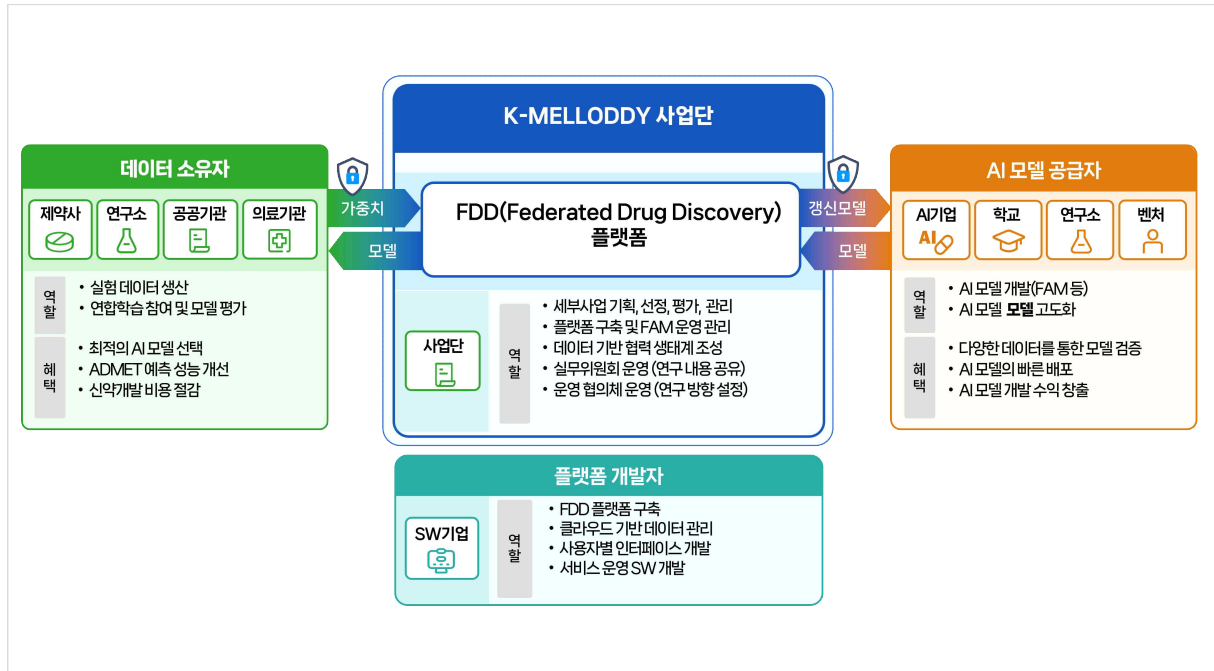
- (연합학습 실용성 입증) K-MELLODDY 사업에서 가장 중요한 목표는 연합학습의 실제 현장 적용을 확인하고 향후 확대 가능성을 입증하는 것임
 - 1차년도에는 세부 사업별 구체적인 목표 설정과 연구개발 범위를 상세하게 설계함
 - 2차년도 이후에는 FDD 플랫폼을 구축하고 FAM 솔루션 동작을 확인하며 이의 고도화 및 확산을 추진함
- (데이터 기여도 평가) 연합학습 실제 도입 시 가장 큰 이슈는 자신이 보유한 모델 성능도 개선 되겠지만 공동의 모델 성능개선을 위해서 자신의 데이터를 활용하는 것에 대한 거부감임(즉, 자신의 귀중한 데이터로 남 좋은 일을 하는 것이라는 생각)
 - 이러한 문제를 피하기 위해서는 데이터 소유자에게 인센티브를 제공해야 하며, 예를 들어 활용된 데이터의 양과 질에 기반한 'AI 모델 개선 기여도'를 측정하고 이를 향후 정부 지원 사업이나 우수 AI 회사 발굴에 활용하는 전략이 필요함
- (데이터의 다양성 확보) AI 모델의 성능을 높이기 위해서는 다양한 데이터가 필요하며 이를 위해 여러 기관의 참여와 협력이 중요함
 - 제약 회사가 약물 후보물질 발견 단계에서 수행한 in-vitro 데이터뿐 아니라, in-vivo, 임상 진행 중 측정된 약동학 데이터 보유기관과의 협력 전략이 필요함

▶ 추진 체계

- 데이터 소유자는 분석 목적 태스크(세부 ADMET 예측 등)를 협의하여 선정하며, FAM 솔루션을 구성하는 베이스라인 AI 모델 개발에 필요한 데이터를 공급함
- 모델 공급자는 FAM 솔루션을 구성하는 베이스라인 AI 솔루션을 개발하고, 이의 성능을 FDD 플랫폼으로 검증하며 고도화함
- 플랫폼 개발자는 안전한 데이터 활용과 편리한 사용자 인터페이스를 개발하고, 연합학습 참여와 기여도를 산정하는 시스템을 구축함

- K-MELLODDY 사업단은 한국제약바이오협회에서 담당하며 데이터 소유자, 모델 공급자, 플랫폼 개발자를 선정하고 사업의 최종 목표를 달성할 수 있도록 사업평가 및 관리를 수행함

<그림 2> K-MELLODDY 추진 체계 및 참여자의 역할과 혜택



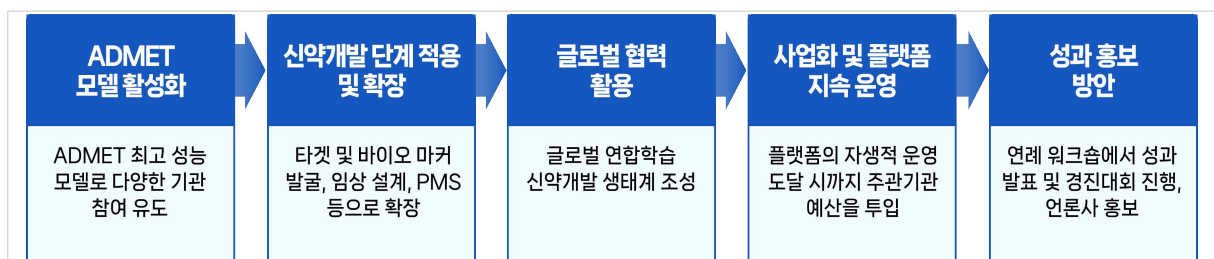
출처: 한국제약바이오협회 AI신약융합연구원

■ 성과 활용 및 기대효과

▶ 활용 방안

- (ADMET 모델 활성화) 현장에서 유용하게 사용할 수 있는 우수한 성능을 가진 ADMET 예측 모델을 개발하고 참여기관을 확대함
- (신약개발 단계 적용 및 확장) 본 사업을 통해 구축되는 플랫폼은 향후 신약개발의 여러 단계 즉, 타겟 발굴, 후보물질 탐색, 선도물질 최적화, 임상 설계, 시판 후 조사 등에 적용될 수 있음
- (글로벌 협력 활용) 글로벌 협력을 통해 K-MELLODDY 참여자를 확대하고 플랫폼의 성능을 높이도록 함

<그림 3> K-MELLODDY 사업 성과 활용 방안



출처: 한국제약바이오협회 AI신약융합연구원

▶ 기대효과

- 성능이 개선된 ADMET 예측 모델 개발을 통하여 시험 분석 비용을 절감함
 - 2022년 국내 상장 제약 회사의 민간 신약개발 R&D 투자 비용은 4조 3,894억원⁶⁾이며, R&D 비용 중 ADMET 시험 분석에 약 9,656억원이 소요됨(북미·유럽 비율 22% 적용 시)
 - 개선된 모델을 통해 20%의 비용을 절감하는 경우, 2022년 기준 R&D 비용 연간 1,900억원 이상 절감이 예상됨
- 이외에 본 사업의 기대효과를 기술, 제도·법규, 신약개발 산업, 생태계 영역별로 구분하였음

<그림 4> K-MELLODDY 사업(FDD 플랫폼 구축) 기대효과



출처: 한국제약바이오협회 AI신약융합연구원

■ 맺음말

- 최근 AI 기술의 발전과 보편화로 AI의 중요성에 대한 인식은 높아졌으나 AI의 성능을 높일 수 있는 다양한 데이터 확보에 대한 대책이 부재함
- K-MELLODDY 사업을 통하여 연습학습 기술을 현업 문제 해결에 적용되는 국내 최초의 성공 사례를 만들고자 하며 구체적으로는 최고의 성능을 얻는 네트워크 기반의 ADMET 예측 솔루션, FAM 솔루션을 구축하고자 함
- FAM 솔루션을 기반으로 AI 신약개발의 여러 단계인 타겟 발굴, 타겟 검증, 바이오마커 발굴, 유효물질 발굴, 선도 물질 최적화, 임상 설계, 효과 분석 등에서 FDD가 사용될 것으로 기대함
- 이와 관련된 산업계, 학계, 연구소, 병원 등 연구자들의 많은 관심과 참여가 필요함

6) 한국제약바이오협회, 2023 제약바이오산업 Databook, 2023.12

글로벌 주요 동향

◆ 유럽, 실사용데이터(RWD) 전자 카탈로그 출시

- 유럽의약품청(EMA)과 의약품 규제기관정상회의(HMA)는 연구와 실사용데이터(RWD, Real World Data) 소스 간 상호 운용성 향상 등 다양한 개선사항을 도입한 전자 카탈로그*를 출시

* HMA-EMA 카탈로그는 실사용데이터(RWD) 소스 및 RWD 연구에서 수집된 데이터 저장소

- EMA는 카탈로그가 의약품의 사용, 안전성 및 효과 조사에 적합한 데이터 식별, 데이터 소스의 적합성 평가, 사용자 보안 강화, 데이터 기반의 의약품 규제 촉진 등에 도움을 줄 것으로 기대

Pink Sheet, "EU Regulators Make Finding & Using RWD Easier", 2024.2.15.

EMA, "HMA-EMA Catalogues of real-world data sources and studies" (<https://catalogues.ema.europa.eu/>)

◆ 유럽, 의약품 접근성 개선 규정 승인

- 유럽의회 법무위원회(JURI)는 보건 비상사태 발생 시, 의약품 접근성 개선을 위해 유럽 집행위원회(EC)가 '강제실시권(CL, Compulsory Licenses)'을 발급하도록 하는 EU 규정 초안에 승인

* 강제실시(強制實施, compulsory license): 정부가 특정 지식재산권의 효력을 일시 중지시키거나 특허권자의 허락 없이 강제로 특허를 사용하도록 하는 규정

Pink Sheet, "EU Committee OKs More Pharma-Friendly Compulsory Licensing Plan", 2024.2.15.

◆ 호주, 첨단기술에 대한 신규 의료기술평가(HTA) 수립 권고

- 호주 정부는 의료기술평가(HTA, Health Technology Assessment) 검토 보고서에서 기존 시스템이 항체-약물 접합체, 마이크로바이옴 등 신기술 평가에 부적합하다고 결론짓고, 세포 및 유전자 치료제 등 최신 의료기술을 위한 새로운 HTA 프로세스 수립을 권고

Pink Sheet, "Australia Identifies HTA Challenges Posed By Emerging Technologies", 2024.2.16.

◆ 덴마크, 임상시험·체외진단연구 동시 승인 프로세스 개시

- 덴마크 의약품청(DNA)은 제약사가 체외진단(IVD, In Vitro Diagnostic) 성능 연구와 병행하여 의약품의 임상시험을 포함하는 이른바 결합 연구(combined studies)의 동시 승인을 신청할 수 있도록 하는 새로운 프로세스를 개시

Pink Sheet, "Denmark Leads The Way On Speeding Up Approvals For Drug/IVD Combination Trials", 2024.2.22.

◆ 영국, AI 규제 유연성 확대

- 영국 정부는 당분간 AI 시스템을 관리하는 법안을 도입하지 않기로 결정. AI 규제에 대한 비입법적인 '유연한 접근 방식'이 혁신을 촉진하고 산업 생태계 변화에 신속하게 대응할 수 있게 해줄 것이라고 밝힘
- 정부 발표에 따라 영국 의약품규제청(MHRA)의 제약업계를 위한 안전한 AI 사용 관련 지침 개발 예상

Pink Sheet, "What Does The UK's 'Flexible' Approach To AI Regulation Mean For Pharma?", 2024.2.21.

◆ 프랑스, 의약품 부족 대응 신규 로드맵 발표

- 프랑스 정부는 의약품 부족 문제에 대처하기 위해 2024년부터 2027년까지의 새로운 로드맵을 발표
- 로드맵에는 공급부족 예측 강화, 주요치료목적의약품(MITM, Médicaments d'Intérêt Thérapeutique Majeur)* 목록 최적화, 약사를 위한 부족 의약품 대체 품목 조제 규칙 등이 포함

* 주요치료목적의약품(MITM): 연중 2개월(특정 대상의 경우 4개월) 분량의 안전 재고를 확보해야 하는 의약품
Pink Sheet, "France: Repatriating Drug Production Among Actions In 'Roadmap' For Tackling Shortages", 2024.2.26.
France Government, Feuille de route 2024-2027: Garantir la disponibilité des médicaments et assurer à plus long terme une souveraineté Industrielle, 2024.2.21.

◆ 파키스탄, 의약품 임상시험 관리기준 지침 변경 검토

- 파키스탄 의약품규제청(DRAP)은 국제 표준 및 모범 사례에 부합하도록 임상시험 관리기준(GCP, Good Clinical Practice), 의약품 리콜, 의약품 사업부 설립에 관한 지침을 변경하기 위해 이해관계자의 의견수렴 절차 진행

RAPS, "Asia-Pacific Roundup: Pakistan Floats Guideline Changes On GCP Inspections and More", 2024.2.26.

◆ ICH, 세포·유전자 치료제 및 RWD 사용 지침 개발계획

- 국제의약품 규제조화위원회(ICH) 지역회의에서 세포 및 유전자 치료제와 실사용 데이터(RWD)를 사용한 약물역학 연구 설계에 대한 국제적으로 조화된 지침 개발 계획을 발표

RAPS, "ICH Targets Cell and Gene therapies, Real-world Data as Topics ripe for harmonization", 2024.2.23.