

연합학습 기반 신약개발 가속화 프로젝트 (K-MELLODDY)

사업개요

K-MELLODDY(Korea Machine Learning Ledger Orchestration for Drug Discovery)는 제약사, 병원, 대학, 정부 연구기관 등이 보유한 민감한 실험 데이터의 프라이버시를 보호하면서 신약개발에 필요한 AI 모델을 공동으로 구축하는 것을 목표로 함

☑ EU에서 추진한 MELLODDY 프로젝트는 Merck, Janssen, Novartis, AstraZeneca, Bayer 등 10개 제약사가 연합학습(Federated Learning, FL) 기반으로 AI 모델을 공동 구축하는 프로젝트

예산

348억원 (보건복지부 : 199억원 / 과학기술정보통신부 : 149억원)

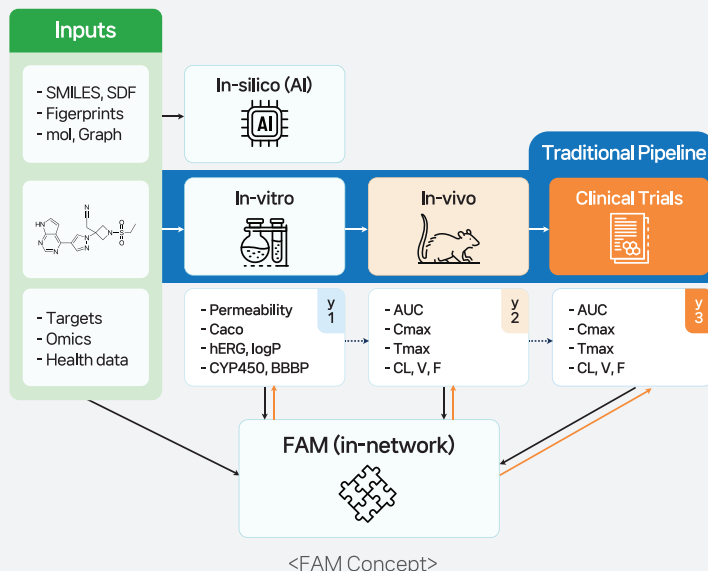
사업기간

2024.4 ~ 2028.12 (5년)

사업목표

- 연합학습 기반 신약개발(FDD, Federated Drug Discovery) 플랫폼 구축
- FDD 플랫폼 기반 ADMET/PK 파라미터 예측 모델(FAM, Federated ADMET Model) 개발

- ☑ 연합학습: 데이터를 기관 외부로 이동하지 않고 AI 모델 파라미터만 공유하여 참여 기관들이 성능이 우수한 AI 모델을 공동으로 만드는 기술
- ☑ FAM: 연합학습을 통해 in-vitro, in-vivo 및 임상시험(in-human) 데이터를 연계하여 endpoint로 ADMET 및 임상시험 PK 파라미터까지 예측하는 AI 모델로 데이터가 추가되면서 성능이 연속적으로 개선됨



세부과제

1세부 : 플랫폼 구축 (1개 과제)	2세부 : 데이터 공급·활용 (20개 과제)	3세부 : AI 모델 개발 (15개 과제)
연합학습 기반 신약개발 플랫폼 FDD를 구축하고 FAM 솔루션을 운영	제약사, 병원, 연구소 등 각 기관이 보유한 실험 데이터 공급 및 FAM 활용	FAM 솔루션 소프트웨어 개발 (1차년부터 3차년도까지 매년 5개 선정)

기대효과

- 데이터 기반 오픈 이노베이션 생태계 구축
 - 경쟁 관계의 기관들도 각자 보유한 데이터를 기반으로 협력할 수 있는 오픈 이노베이션 생태계 구축
 - AI 모델 개발자는 기존 데이터 제공자와의 1:1 매칭 방식과 달리 플랫폼 기반의 1:N 매칭이 가능해지므로 빠른 모델 검증 및 배포 가능
- AI 기반 신약개발 선도국진입
 - ADMET/PK의 전 주기적인 통합 예측 모델 확보로 신약개발 R&D 비용과 시간 절감
 - FAM을 기반으로 신약개발 파이프라인의 다양한 AI 모델 확장 개발 가능
 - FDD 플랫폼 및 FAM을 통한 글로벌 협력 확대



Korea Machine Learning Ledger Orchestration for Drug Discovery(K-MELLODDY) Project

Overview The K-MELLODDY project aims to build and operate an AI-driven drug discovery platform that allows pharmaceutical companies, hospitals, universities, and government research institutions to collaboratively develop AI models while preserving the privacy of sensitive experimental data.

- ☑ The EU-led MELLODDY project is a collaborative initiative in which 10 pharmaceutical companies—including Merck, Janssen, Novartis, AstraZeneca, and Bayer—jointly developed AI models based on Federated Learning (FL).

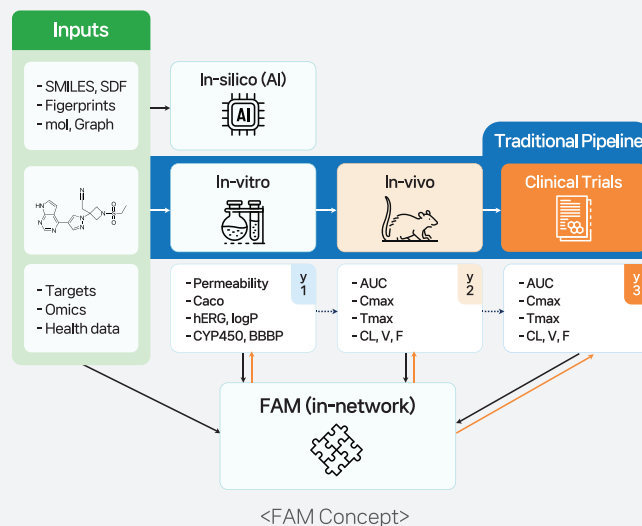
Budget/Duration

KRW 34.8 billion / April 2024 – December 2028 (5 years)

Goals

- Build a Federated Drug Discovery (FDD) acceleration platform
- Develop a Federated ADMET Model (FAM)

- ☑ Federated Learning: A technology that allows participating institutions to collaboratively develop a high-performance AI model by sharing only model parameters, without transferring their data outside their organizations.
- ☑ FAM: An AI model that links in-vitro, in-vivo, and in-human (clinical trial) data through federated learning to predict ADMET and clinical pharmacokinetic(PK) parameters as endpoints. The model continuously improves in performance as additional data is integrated.



Specific tasks

Platform Development (1 project)	Data Provision & Utilization (20 projects)	AI Model Development (15 projects)
Build a Federated Drug Discovery (FDD) platform and operate the Federated ADMET Model (FAM) solution.	Participating institutions—including pharmaceutical companies, hospitals, and research institutes—will supply experimental data and utilize the FAM solution.	Develop the FAM solution software (Select 5 institutions annually from Year 1 to Year 3)

Expected Outcomes

- Establishing a Data-driven Open Innovation Ecosystem
 - Foster an open innovation ecosystem where even competing institutions can collaborate based on their proprietary data.
 - Enable AI Model Developers to transition from traditional 1:1 data matching to a platform-based 1:N model, accelerating model validation and deployment.
- Advancing as a Global Leader in AI-Based Drug Development
 - Reduce R&D costs and time by developing integrated ADMET/PK prediction models that span the entire drug development cycle.
 - Enable the development of a wide range of AI models across the drug development pipeline, based on FAM.
 - Promote global collaboration through the FDD platform and FAM.

