

「초정밀 헬스케어 인공지능 연구개발 중장기 로드맵 2035」 수립

- 유전체·영상 데이터와 실시간 건강정보 등 헬스케어 데이터가 결합한 차세대 초정밀 헬스케어 연구개발 청사진 제시
- 코호트 기반 한국형 인공지능(AI) 학습용 데이터 모음 45종 100만명분 구축
- ‘데이터가 치료가 되는 시대, 헬스케어 인공지능 강국 선도’ 비전 발표

질병관리청(청장 임승관) 국립보건연구원(원장 남재환)은 빅데이터·인공지능 기반 미래 보건의료 기술 개발 방향을 담은 「초정밀 헬스케어* 인공지능 연구개발 중장기 로드맵 2035」를 수립했다.

* 개인의 유전체와 건강정보 및 일상의 연속적 생체·행동 데이터 등을 통합 분석하여 초개인화된 맞춤형 건강관리 솔루션을 실시간 응답형(adaptive)으로 제공하는 차세대 헬스케어 기술과 서비스

이번 로드맵은 한국형 인공지능 학습용 데이터를 구축하고 개인 맞춤형 건강관리 기술을 개발하기 위한 중장기 연구개발 전략이다.

이번 로드맵은 급속한 고령화와 만성질환 증가, 디지털 헬스케어 및 정밀의료 시장 확대 등 변화하는 보건의료 환경에 대응하기 위해 마련됐다. 국민 개개인의 건강상태와 유전적·환경적 특성을 종합적으로 고려하는 초정밀 헬스케어 체계 구축을 목표로 한다.

국립보건연구원은 그동안 구축한 대규모 코호트와 바이오뱅크를 기반으로 임상·역학, 의료영상, 라이프로그, 유전체 등 바이오빅데이터가 통합·연계된 세계 최고 수준의 한국형 인공지능 학습용 통합 건강정보 데이터 45종(100만 명분)을 구축할 예정이다.

또한 이번 로드맵은 ‘데이터가 치료가 되는 시대, 헬스케어 인공지능 강국 선도’를 비전으로 제시했다.

신뢰, 혁신 체감, 통합의 4대 핵심 가치를 바탕으로 국민이 일상에서 체감할 수 있는 헬스케어 인공지능을 개발하고 산·학·연·병이 함께하는 협력 생태계를 구축하고자 한다.

국립보건연구원은 이를 위해 ▲코호트 기반 대규모 한국형 멀티모달 인공지능 학습용 데이터 확보 ▲목적형 헬스케어 멀티·옴니모달* 데이터셋 구축 ▲초정밀 헬스케어 인공지능 기술 개발 및 실증 연구 ▲공유·개방을 통한 초정밀 헬스케어 AI 생태계 조성 등 4대 핵심 추진 전략을 마련하였다.

- * 멀티모달 : 동일인 유래 임상·역학, 의료영상, 라이프로그, 멀티오믹스 등 서로 다른 유형의 데이터를 특정 시점으로 연계·통합한 데이터
- * 옴니모달 : 개인의 다양한 멀티모달 데이터를 시간의 흐름에 따라 연속적으로 통합·축적하여 개인의 상태 변화를 시계열적으로 반영하는 데이터

로드맵은 기술 성숙도와 헬스케어 현장의 수요를 고려하여 2035년까지 추진단계를 3단계*로 구성하고, 우선 2032년까지 향후 6년간에 걸쳐 데이터 자원화 및 모델 지능화를 위한 핵심 과제들을 추진할 예정이다.

- * 1단계('27~'29, 도약기): 데이터 자원화, 2단계('30~'32, 가속기): 모델 지능화, 3단계('33~'35, 완성기) : 가치실현

2035년에는 한국인의 건강·질병 특성을 반영한 글로벌 수준의 한국형 인공지능 학습용 데이터를 구축하고, 헬스케어 분야 인공지능 및 유전체 기술 혁신을 통해 국민 개개인에게 초개인화된 실시간 맞춤형 건강관리와 질병 예측·예방 서비스를 제공할 수 있을 것으로 기대된다.

남재환 국립보건연구원장은 “초정밀 헬스케어 인공지능은 개인별 건강위험을 더 정밀하게 예측하고 예방과 치료의 정확도를 높이는 핵심 기술이며, 앞으로 국민이 체감할 수 있는 AI 기반 미래의료 기술 혁신을 위하여 본 로드맵을 차질 없이 추진하겠다.”고 밝혔다.

임승관 질병관리청장은 “세계 최고 수준의 바이오 빅데이터와 초정밀 헬스케어 인공지능 기술을 융합하여 국민 모두가 차별 없이 혜택을 누릴 수 있는 건강 및 질병 관리 서비스 개발에도 최선을 다하겠다.”고 밝혔다.

<붙임> 초정밀 헬스케어 인공지능 연구개발 중장기 로드맵 2035

담당 부서	국립보건연구원 헬스케어인공지능연구과	책임자	과 장	김상철 (043-719-8850)
		담당자	연 구 관	이수연 (043-719-8861)
			연 구 사	공진화 (043-719-8853)

□ **사업 개요**

- **(목적)** 초정밀 헬스케어 실현을 위한 한국형 대규모 인공지능 학습용 데이터셋 구축 및 헬스케어 AI 기술 활용 체계 구축
- **(사업 내용)**
 - **(코호트 기반 멀티모달 인공지능 학습용 데이터 확보)** 기 구축 코호트 기반 멀티모달 데이터 확대 수집(디지털 코호트) 및 바이오뱅크 인체 자원으로부터 바이오데이터 생산(데이터 자산화)
 - * (KAIR) Korea AI-Ready 데이터: 임상유전체와 연계된 뇌영상, 라이프로그, 멀티오믹스, 심전도, Audio(음성, 숨소리 등), Video(걸음걸이, 표정, 의료 동영상등)
 - **(목적형 헬스케어 멀티·오니모달 인공지능 데이터셋 구축)** 코호트 기반 시계열 멀티모달(오니모달) 레이블링, 주석 가공 등을 통해 헬스케어 AI 연구에 최적화된 플래그쉽 AI-Ready 데이터셋 구축
 - * (PAIR) Purpose-driven AI-Ready.
 - PAIR-Ref(참조용 데이터), PAIR-Edu(교육용 데이터), PAIR-Val(검증평가용 데이터), PAIR-Syn(합성데이터), PAIR-Omni (범용 헬스케어 AI 솔루션 연구개발용 데이터)
 - **(초정밀 헬스케어 인공지능 활용 기술 개발)** 고품질 멀티모달 데이터를 활용한 개인 맞춤형 건강관리 및 질병 예후예측 모델링 등 초정밀 헬스케어 인공지능 실증 연구
 - * 코호트 특화 AI Co-scientists 및 오니모달 헬스케어 AI 기본모델 개발
 - **(초정밀 헬스케어 인공지능 연구지원)** AI 학습용 데이터 공유·개방 체계 구축, 성과확산을 위한 참여형 프로그램 운영 및 신뢰형 헬스케어 AI R&D 생태계 조성
 - * 데이터셋(45종) 구축 대비 공개율 단계별 확대: 1단계('27~'29) 18종 40%, 2단계('30~'32) 35종 80%, 3단계('33~'35) 45종 100%



<초정밀 헬스케어 인공지능 연구개발 중장기 로드맵 2035 추진 개요도>

□ 사업 비전 체계도



- **초정밀 의료** : 멀티모달 데이터 기반 헬스케어 특화 인공지능 모델을 활용하여 개인 (N-of-1)의 건강상태를 정밀하게 분석·예측하고, 전 생애주기에 걸친 선제적·지능형 건강 관리 및 맞춤형 의료서비스를 제공하는 체계
- **초정밀 헬스케어** : 개인의 유전체와 건강정보 및 일상의 연속적 생체·행동 데이터 등을 통합분석하여 초개인화된 맞춤형 건강관리 솔루션을 실시간 응답형(adaptive)으로 제공하는 차세대 헬스케어 기술과 서비스

구분	정밀의료 (Precision Medicine)	초정밀의료 (Hyper Precision Medicine)
중점 적용 대상	환자군 단위 분류 (Sub-population)	개인 단위 분류 (Individual, N-of-1)
활용데이터	유전체(Genomics), 임상 기록	멀티오믹스, 라이프로그(웨어러블환경, 행동식이 등)
접근 방식	유사한 유전적 특성을 가진 환자군에 대한 치료 전략 적용	실시간 예측-개입-피드백이 순환되는 폐쇄형 순환구조 (closed-loop) 기반 개인 맞춤 치료 전략 최적화
적용 범위	질병의 진단 및 치료 중심	예측·예방·치료·사후 관리를 포함한 전 생애주기 건강 관리
의료 개입 시점	증상 발생 이후 개입	증상 이전 단계의 선제적·자동 개입
주요 기술 수단	통계 기반 머신러닝	AI 에이전트(LLM 기반), 디지털 휴먼 트윈, 실시간·연속 학습 및 의사결정 시스템
의료 패러다임	반응형 의료 (Reactive Care)	적응형·자율형의료 (Adaptive & Autonomous Care)

- **옴니모달 데이터** : 개인의 다양한 멀티모달 데이터를 시간의 흐름에 따라 연속적으로 통합·축적하여 개인의 상태 변화를 시계열적으로 반영하는 데이터

구분	멀티모달 데이터 (Multimodal Data)	옴니모달 데이터 (Omnimodal Data)
데이터 구조	단면적(cross-sectional) 데이터 결합	연속적(longitudinal)·시계열 데이터 통합
시간적 특성	특정 시점 또는 제한된 기간의 데이터	전 생애주기 기반 연속·실시간 데이터
데이터 수집 방식	개인의 다양한 데이터 모달을 특정 시점 기준으로 수집	개인의 생체·임상·생활환경 데이터를 시간 축을 따라 지속적으로 수집
각 데이터별 차이점	임상·역학 진단 시점의 검사 결과, 처방, 병력 멀티오믹스 유전체, 전사체 등 특정시점 데이터 의료영상 CT, MRI 등 단일 시점 의료 영상 라이프로그 설문 기반 생활습관, 단기 측정 데이터	임상·역학 진료 이력 전체 치료 반응경과 데이터 멀티오믹스 다중 시점별 시계열 멀티오믹스 데이터 의료영상 질병 진행에 따른 반복·추적 시계열 의료 영상 라이프로그 웨어러블 기반 심박·활동·수면 등 시계열 데이터
데이터 통합 방식	사전 정의된 모달 간 특징 융합	상황·맥락 기반 동적 데이터 통합
AI 모델 역할	분류·예측 중심 모델	상태 예측-개입-피드백 순환 모델
모델 예측 결과	시점별 정보 융합을 통한 상태 추정	상태 변화 추적을 통한 예측적응 학습
헬스케어 인공지능 모델 개발 적용	정밀의료 분석에 활용	초정밀의료(Hyper-Precision Medicine) 구현 핵심 데이터 구조