

# 사업명 : 보건의료데이터 상호운용성 지원 기술개발(R&D)

RFP

## 보건의료데이터 표준스펙 적용 자동화 기술개발 및 실증

사업명	보건의료데이터 상호운용성 지원 기술개발		사업유형 해당여부	<input type="checkbox"/> 혁신도전형 R&D <input checked="" type="checkbox"/> 다부처 협업 패키지 R&D <input checked="" type="checkbox"/> 연구데이터 관리계획 제출 대상	
제안요청서명	보건의료데이터 표준스펙 적용 자동화 기술개발 및 실증				
공모유형	품목지정형	TRL 단계 (시작~종료) * TRL 단계 공모안내서 참고	해당사항 없음	기술료 납부대상	X

### ▶ 사업목적

- 의료현장에서의 보건의료데이터 정제, 표준화 변환, 가명처리 등의 작업을 AI 등 최신기술을 활용해 자동화함으로써 업무 효율화, 정확성 제고 및 상호운용성 기반 마련

### ▶ 지원대상

- 주관연구개발기관은 「의료법」 제3조제2항제3호에 근거한 병원급 의료기관이어야 함
- 주관연구책임자는 소속 의료기관의 최고정보관리책임자, CIO(Chief Information Officer) 또는 이에 준하는 권한을 가진 책임자이어야 함
- 공동연구개발기관으로 산·학·연·병 구성은 가능하나, 기업은 반드시 참여해야 함
  - 과제 신청 시 EMR 시스템 개발기업은 공동연구개발기관으로 1개 과제에만 참여할 수 있으며, 중복 참여는 불가함
- 일반적인 사항은 공고안내서의 '신청요건' 부분 참고

### ▶ 지원규모

지원분야	지원기간	연간 연구개발비 (1차년도)	선정 예정 과제수
보건의료데이터 표준스펙 적용 자동화 기술개발 및 실증	5년 이내(3+2)	1,500백만원 이내 (1,125백만원 이내)	3개

※ 1차년도 연구개발기간 9개월 이내, 연구개발비는 9개월 이내 기준으로 지급

※ 연도별 예산 상황에 따라 연간 연구개발비 변동 가능

### ▶ 연구내용

지원분야	연구내용
보건의료데이터 표준스펙 적용 자동화 기술개발 및 실증	<p>&lt; 사업 개념도(안) &gt;</p>

- AI 기술 기반 SW의 데이터 변환 대상은 2내역(보건의료데이터 표준 통합 운영체계 개발 및 실증)에서 제시하는 표준 데이터모델 구조 및 제시 항목을 충실히 반영하여야 하며, 개발한 SW는 2내역 ‘표준 통합 운영체계’에 포함하여 패키지” 배포 가능하여야 함
    - \* 본 사업에서 수행하고자 하는 연구 수요가 높은 항목 및 용어체계를 연구자 추가 제시
    - \*\* (패키지) 의료기관 종별과 EMR 유형을 고려하여 의원, 병원·종합병원, 상급 종합병원 등 3종 이상으로 제공
- ※ 표준 데이터모델은 「보건의료데이터 용어 및 전송 표준」 고시(개정사항 포함)를 반드시 포함해야 하며, 그 외 전자의무기록에서 교류 가능한 최대항목을 선정하되 의료데이터 기반사업\* 및 임상연구에 활용이 가능하여야 함

\* 진료정보교류 표준 고시, 건강정보고속도로 및 EMR인증제 항목 등 검토 필수
- 의료기관 종별(의원·병원·종합병원 등)로 AI 기반 SW 개발을 위해 데이터 수집, 정제, 라벨링, 의료기관 기존 보유 현황 및 모델 재학습 전략 등 데이터 활용 연구
  - 의료기관 내에서 활용될 수 있도록 AI 기술 기반 ① 비정형 데이터 구조화 SW, ② 표준용어 매핑 SW, ③ 가명처리 SW 개발 및 실증
    - ① 비정형 데이터 구조화 SW 개발
      - AI 등 기술을 활용하여 Freetext, 이미지 등 비정형 데이터를 정형화·구조화하는 자동 변환 기술개발
    - ② 표준용어 매핑 SW 개발
      - AI 등 기술을 활용하여 의료기관마다 다르게 사용하는 의료용어를 국가 용어 표준(KR CD) 및 표준용어체계(SNOMED CT, LOINC, ICD, RxNorm 등의 매핑 자동화를 지원하는 기술개발
    - ③ 가명처리 SW 개발
      - ‘보건의료데이터 활용 가이드라인’을 준수하여, 다기관 연구·협력 시 개인정보 유출을 최소화하기 위한 AI 기반 가명처리 기술개발
      - \* 식별 정보(이름, 주민등록번호, 질병 진단기록 등) 처리기준을 연구개발계획서 내 제시
    - ④ 보건의료데이터 표준스펙 적용 자동화 SW(①~③) 성능평가
      - (1단계) 과제별 참여기관을 중심으로 의료기관 종별·EMR 유형별 AI SW 성능 평가 및 개선(정확도 80% 이상)
      - (2단계) AI SW의 범용성 확보 및 성능개선을 위해 3내역사업 참여의료기관 전체를 대상으로 성능 검증 및 배포용 SW 테스트(정확도 90% 이상)
  - 자동화 AI SW를 설치한 의료기관에서 해당 SW의 성과와 오류를 확인하고, 이를 ‘표준 통합 관리 시스템’에서 모니터링할 수 있도록 개발
  - 보건의료데이터 표준 통합 운영체계 기반 임상 현장 실증
    - 표준스펙 적용 자동화 SW(비정형 데이터 구조화 SW, 표준용어 매핑 SW, 가명처리 SW)를 2내역(보건의료데이터 표준 통합 운영체계 개발 및 실증)의 ‘표준 통합 운영체계’에 통합 적용 및 성능평가
    - 최종 통합된 ‘표준 통합 운영체계’를 2내역과 협업하여 현장 실증 수행
    - ※ 현장 실증 절차 및 방안은 1내역(총괄 코디네이션 센터)과 협의
    - ※ 실증의료기관은 총 12개 이상으로 의원급 4개 이상(EMR 3종, 진료과 3개 포함), 병원급 4개 이상, 종합병원급 4개 이상(상급종합병원 2개 포함)
    - \* 종합병원급 이상은 국가 통합 바이오 빅데이터 등 국가 데이터 교류 관련 사업 참여병원으로 구성해야 함

## ▶ 성과목표·지표

- (1단계(’25~’27년)) 보건의료데이터 표준스펙 적용 자동화 기술개발 및 검증

지원분야	성과지표	달성 목표	비고
보건의료데이터 표준스펙 적용 자동화 기술개발 및 실증	<ul style="list-style-type: none"> <li>▪ SCI(E) 논문 게재</li> </ul>	<ul style="list-style-type: none"> <li>▪ 1건</li> </ul>	<ul style="list-style-type: none"> <li>▪ JCR(Journal Citation Report) 카테고리 기준 IF(Impact Factor) 상위 20% 이내 논문만 인정</li> </ul>

지원분야	성과지표	달성 목표	비고
	▪ 비정형 데이터 구조화 SW 개발	▪ 1건	▪ 성능평가 기준: 정확도(Accuracy) 80% 이상, 일관성(Consistency)* 90% 이상 * 반복 테스트, 교차 검증을 통한 동일 결과 생성 비율 ▪ SW 연구성과물 등록 필요 (한국저작권위원회 또는 정보통신산업진흥원) ▪ 사용자/운영자 매뉴얼
	▪ 표준 용어매핑 SW 개발	▪ 1건	
	▪ 가명처리 SW 개발	▪ 1건	

○ (2단계('28~'29년)) 보건의료데이터 표준스펙 적용 자동화 기술 실증 및 확산

지원분야	성과지표	달성 목표	비고
보건의료데이터 표준스펙 적용 자동화 기술개발 및 실증	▪ SCI(E) 논문 게재	▪ 1건	▪ JCR(Journal Citation Report) 카테고리 기준 IF(Impact Factor) 상위 20% 이내 논문만 인정
	▪ 비정형 데이터 구조화 SW 고도화	▪ 1건	▪ 성능평가 기준: 정확도(Accuracy) 90% 이상, 일관성(Consistency) 95% 이상 ▪ 사용자/운영자 매뉴얼
	▪ 표준 용어매핑 SW 고도화	▪ 1건	
	▪ 가명처리 SW 고도화	▪ 1건	▪ 성능평가 기준: 정확도(Accuracy)·일관성(Consistency)·재현율(Recall) 95% 이상 ▪ 사용자/운영자 매뉴얼
	▪ AI 자동화 기술 성과보고서	▪ 1건	▪ 처리 데이터 규모, 자동화 처리 증가율, 오픈소스 보급 건수, 가명처리 적절성 등 기술 확산 성과 내용 포함
	▪ SW별 검증 보고서	▪ 3건	▪ 비정형 데이터 구조화/표준용어 매핑/가명처리 SW에 대한 기능 성능 보안 사용자 수용성 등 검증 결과 포함 ▪ 공인된 시험성적 기관(한국정보통신기술협회 등) 인증서 첨부
	▪ 표준 통합 운영체계(2내역) 적용 결과 및 성능평가 보고서	▪ 1건	▪ 의료기관 종별 및 EMR 유형별 패키지 성능평가 결과 포함
	▪ 현장 실증보고서	▪ 1건	▪ 의료기관 종별 및 EMR 유형별 패키지 실증 결과 포함

※ 상기 성과목표를 연구개발계획서에 포함하여 제출하여야 하며, 제시된 성과목표 이외 성과목표 추가 가능

▶ 특기사항

- 본 사업은 과제선정 이후 선정된 주관연구개발기관이 모두 참여하는 하나의 컨소시엄 추진체계를 구성하여 운영(붙임 1. 사업 추진체계 참조)
- 주관연구책임자는 컨소시엄 운영위원회에 반드시 참여해야 함
- 과제 선정 이후 수정계획서 제출 시 컨소시엄 운영위원회 검토의견을 반영하여야 함
- 연구책임자는 컨소시엄 운영위원회를 통하여 합의된 내용을 차년도 연차보고서에 반영하여야 함
- 컨소시엄 운영위원회 운영 결과를 전문기관(한국보건산업진흥원)에 보고하여야 함
- 본 사업은 바이오헬스 다부처 협업 R&D로 다부처 협업 사업\* 간 연계·협력에 적극 참여하여야 함
- \* 국가 통합 바이오빅데이터 구축 사업, 첨단바이오 특화연구소 데이터 플랫폼 사업 등

- 연구개발계획서 작성 관련
  - 의료기관 종별(의원·병원·종합병원 등) 및 다양한 EMR 시스템 유형을 포함하여 개발하고 실증될 수 있도록 실증에 참여하는 의료기관(12개 기관 이상)과 EMR 시스템 개발기업을 연구개발계획서에 명시하여야 함
  - AI SW의 안정적 연구개발을 위하여 AI 등 최신 기술을 보유하거나 활용할 수 있는 개발기업을 연구개발계획서에 명시하여야 함
- 연구데이터의 생산·보존·관리 및 공동활용 등에 관한 계획을 검토하기 위해 과제 신청 시 연구데이터 관리계획(Data Management Plan, DMP)을 필수 제출하여야 함(계획서 첨부 양식)
- 본 사업을 통해 구축된 데이터셋은 보건 의료 연구자 정보센터(CODA), 한국보건의료정보원 등 기탁에 대한 구체적이고 실질적인 협력을 추진하여 컨소시엄 운영위원회 보고 후 연차보고서에 반영해야 함
- 개발된 성과물은 병원 및 산업계, 타 사업 등에서 연계·활용될 수 있도록 오픈 소스로 보급(MIT 라이선스 적용)
  - \* (MIT 라이선스) 소스 코드의 복제·배포·수정이 가능하고 상업적 이용을 허용하는 활용이 자유로운 라이선스
- 본 과제에 참여하는 기업은 「국가연구개발사업 연구개발비 사용 기준」 제65조 제4항 제3호, 제4호, 제5호에 근거하여 참여연구자 인건비 현금 계상이 가능함

#### ▶ 일반사항

- 연구개발계획서 작성시 주요사항
  - 연구목표 달성을 위한 공동연구과제 및 참여연구진 간의 구체적인 역할 배분 및 협력방안 제시
  - 총 연구기간 동안 연차별 마일스톤(정량지표)을 제시
    - ※ 마일스톤은 연구개발 단계별로 달성해야만 하는 주요한 기술적인 실적으로 평가를 통해 실적 달성 여부를 판단 시 주요 기준으로 활용 예정
- 일반적인 사항은 「보건의료기술 연구개발사업 가이드라인」 참고

#### ▶ 선정평가 기준

적용가점		해당사항 없음
구분	평가항목(배점)	
	대 항목	소 항목
서면·구두 평가	1. 연구개발 계획 (40)	<ul style="list-style-type: none"> <li>○ 사업목적에 대한 이해도(5)               <ul style="list-style-type: none"> <li>- 제안요청서(RFP)의 목표와 지원내용에 부합함</li> <li>- 사업목적에 대한 이해도가 높음</li> </ul> </li> <li>○ 연구목표의 구체성 및 실현가능성(10)               <ul style="list-style-type: none"> <li>- 제시한 연구목표가 구체적이며 타당함</li> <li>- 보유기술 완성도 및 현장적용 가능성이 타당함</li> </ul> </li> <li>○ 연구개발과제 수행 계획의 충실성(20)               <ul style="list-style-type: none"> <li>- 연구개발 수행 계획이 구체적이며 충실함</li> <li>- 실증 추진전략, 절차 등이 체계적이고 적절함</li> <li>- 과제간 연계·협업 계획이 체계적이고 구체적임</li> </ul> </li> <li>○ 과제구성 및 추진일정의 적절성(5)               <ul style="list-style-type: none"> <li>- 과제구성 및 추진일정이 적절하고 효율적임</li> </ul> </li> </ul>
	2. 연구개발 역량 (20)	<ul style="list-style-type: none"> <li>○ 연구책임자의 전문성 및 연구실적의 우수성(10)               <ul style="list-style-type: none"> <li>- 연구책임자가 해당 연구를 수행에 필요한 전문성과 연구경력을 갖추고 있음</li> </ul> </li> <li>○ 연구개발기관 연구수행 능력의 우수성(10)               <ul style="list-style-type: none"> <li>- 연구개발기관(의료기관, EMR·AI 기업 등)이 해당 연구 수행에 필요한 역량과 인프라를 갖추고 있음</li> </ul> </li> </ul>
	3. 연구개발 성과 (40)	<ul style="list-style-type: none"> <li>○ 연구의 성공가능성(30)               <ul style="list-style-type: none"> <li>- 연구개발을 통한 기술·경제·사회적 가치 창출 가능성이 높음</li> <li>- 연구개발 성과물의 활용 및 확산 계획이 충분히 고려됨</li> </ul> </li> <li>○ 연구결과의 파급효과(10)               <ul style="list-style-type: none"> <li>- 연구결과가 국민건강증진, 질병극복, 공공복지실현에 기여함</li> </ul> </li> </ul>

※ 선정평가 계획 수립 시 일부 평가항목(배점) 및 내용이 달라질 수 있음

## 붙임 1 사업 추진체계

○ 사업 목표 달성을 위해 컨소시엄 체계로 운영

