

# 의료데이터의 활용과 정책 시사점<sup>1)</sup>

## 요약

우리는 데이터 경제 시대에 살고 있다. 데이터는 삶의 질을 향상시키는 새로운 서비스 창출 및 신산업의 핵심 요인으로 작용하고 있다. 정부는 데이터가 가진 가치에 주목하고 그 가치를 극대화하기 위해 산업의 육성과 관련 규제를 개선·혁신하는 계획을 발표하고 있다.

수많은 데이터 중에서도 의료데이터는 정보의 민감성, 데이터의 수집과 생산이 이루어지는 영역, 데이터의 정형화 정도 등의 특성으로 인하여 분절되어 수집·관리되고 있어 분석과 활용이 쉽지 않다. 데이터를 연계하여 활용하기 위하여 국가 차원에서 플랫폼을 구축하고 연계 사업을 수행하고 있으나 기업들은 제도적 제약과 폐쇄적인 유통구조로 인하여 고품질의 데이터 수집에 어려움을 느끼고 있어 산업적 활용이 저조한 것으로 보인다.

따라서 데이터의 활용을 활성화하기 위해서는 양질의 데이터를 수집하고 안전하게 유통·관리하여 데이터 간 결합을 통해 의미있는 가치를 창출할 수 있도록 데이터 가치사슬 전(全) 단계에 대한 중장기적인 정책 지원 제도가 요구된다.

1) 본고는 산업연구원(2022), 「서비스산업의 데이터 활용 생태계 활성화 전략」의 주요 내용을 요약 및 수정하였다.

## 1. 서론

보건의료산업에서 데이터가 가지는 의미는 매우 크다. 보건의료의 목적인 건강 보호 및 증진, 질병 치료를 위한 접근 방식이 기술과 데이터의 결합으로 빠르게 변하고 있다. 데이터를 분석하여 개인 맞춤형 의료를 제공하고, 의료기관의 프로세스를 최적화하며, 고품질의 의료 기기 및 서비스를 개발하고 있다. 점점 많은 의료기관과 관련 기업들이 데이터를 이용하여 서비스 질을 향상시키기 위해 노력하고 있으나, 금융이나 유통 등 타 업종과 비교하면 데이터의 수집부터 거래, 활용까지의 역량이 다소 낮다. 이에 정부는 보건의료데이터가 가진 본질적 가치의 측면뿐만 아니라 산업적 가치에 주목하고 데이터 역량 확보를 위해 적극적으로 힘을 쏟고 있다. 의료데이터를 안전하게 가공하고

활용하는 가이드라인을 만들어 데이터 활용을 위한 초석을 마련하였으며, 보건의료빅데이터의 플랫폼사업, 국가 통합 바이오 빅데이터 구축 사업 등 공공데이터를 통합하여 활용할 수 있도록 국가 주도 데이터 구축·활용 사업을 수행하고 있다. 또한 의료데이터 중심병원사업 등을 통해 민간 의료기관의 데이터를 공유하기 위한 노력도 진행 중이다. 그동안의 노력들로 보건의료 분야에 대한 데이터 활용 인식이 확산되었지만 해결해야 할 문제가 여전히 남아있다. 따라서 본고에서는 현재 국내 의료데이터의 수집과 유통, 활용 현황을 살펴보고 국민의 삶의 질 향상과 산업·사회 현안에 대한 대응 방안으로써 안전하게 의료데이터를 활용하기 위한 정책 시사점을 제시하고자 한다.

## 2. 보건의료데이터의 수집-유통-활용 현황

### (1) 의료데이터의 종류와 수집

보건의료기본법에서는 보건의료 정보를 보건 의료와 관련한 지식 또는 부호·숫자·문자·음성·음향·영상 등으로 표현된 모든 종류의 자료로 정의하고 있다. 의료데이터는 의료기관에서 환자를 치료하는 과정에서 가장 많이 발생하고 있으며, 최근에는 생체리듬을 측정할 수 있는 웨어러블 기기의 보급이 확산되면서 혈압, 혈당, 호흡, 체온 등 개인의 활력징후 데이터도 생산·축적되고 있

다. 오믹스(omics)<sup>2)</sup> 기술이 고도화되면서 엄청난 양의 유전체, 마이크로바이옴 등의 오믹스 데이터도 생성되고 있다. 또한 신체 측정 및 검사를 통해 획득하는 정보는 아니지만 앱이나 소셜미디어에서 얻을 수 있는 의료 지식을 담고 있는 건강 정

2) 오믹스(omics)란 전체를 뜻하는 말인 옴(-ome)과 학문을 뜻하는 접미사 익스(-ics)가 결합된 말로 유전자(gene), 전사물(transcript), 단백질(protein), 대사물(metabolite) 등 개별 연구에 대비되는 총체적인 개념의 데이터 세트를 바탕으로 하는 생물학 분야를 말한다. <https://terms.naver.com/entry.naver?docId=5141548&cid=60266&categoryId=60266>(검색일: 2023. 5. 23).

보 등도 광의에서 의료데이터에 포함될 수 있다.

의료데이터는 생산 혹은 수집을 담당하는 영역에 따라 공공영역에서 축적되는 데이터와 민간영역에서 축적되는 데이터로 구분할 수 있다. 공공영역의 데이터는 의료기관의 청구·행정을 위해 수집된 데이터, 각종 통계조사 데이터, 정부 주도로 수집된 데이터 등이 포함되는데, 우리나라는 단일 건강보험체계를 갖고 있기 때문에 모든 국민의 건강보험 자격, 보험료, 의료급여 등의 데이터가 공공기관에 축적되어 있다. 민간 의료데이터는 의료기관에서 가지고 있는 환자의 진료 정보, 검사 정보 등이 가장 대표적인데 병원정보시스템 등을 통해 체계적으로 수집·관리되고 있다.

의료데이터는 정형화 정도에 따라 정형데이터, 비정형데이터, 반정형데이터로 구분할 수 있는데, 정형 정도에 따라 활용 가능성에 차이가 있다. 정형데이터는 정해진 형태가 있고 규칙에 따라 저장된 데이터를 말하는 것으로 진료 기록에 있는 진단명, 처방 코드 등이 포함된다. 비정형데이터는 말 그대로 형태가 없는 연산이 불가능한 데이터로 의료 기록, 영상 이미지 등이 포함되며 의료데이터의 80~90%가 비정형데이터라고 할 수 있다.<sup>3)</sup> 전자의무기록(EMR, Electronic Medical Record)은 언뜻 보면 전자 형태로 작성되고 저장되기 때문에 정형데이터라고 생각할 수 있지만 진단명, 처방, 검사 수치 등 정형화된 형태 이외에 간호 기록, 검사 판독 기록 등은 자유 기술의 형태이기 때문에 비정형데이터의 비중이 매우 높은 정보이다. 특히 자유기술 형태의 기록들은 의

료 전문 용어, 줄임 용어, 영문·국문의 혼용 등 그 형태가 매우 다양하고 진료과마다 그리고 병원마다 사용 패턴이 상이하여 분석을 위해 정형화하기가 매우 어렵다.<sup>4)</sup> 영상 이미지의 경우 형식 자체는 의료영상저장전송시스템(PACS, Picture Archiving Communication System)에 의해 규격화되어 있지만, 분석에 필요한 정보들이 정해지지 않은 형태로 이미지에 존재하기 때문에 비정형데이터로 분류하는 것이 일반적이다. 반정형데이터는 형식이 있지는 않으나 데이터 작성자 나름의 규칙을 가지고 작성되어진 데이터로 영상 검사에 대한 의료진의 판독 소견 등이 이에 해당된다. 수집되고 있는 의료데이터에서 비정형데이터가 차지하는 비중이 매우 크어도 불구하고 활용을 위한 비식별조치 등의 가이드라인이 정형데이터에 초점이 맞추어져 있어 의료데이터가 가진 가치를 온전히 활용하기 어려운 실정이다.<sup>5)</sup>

## (2) 의료데이터의 유통

보건의료데이터는 데이터의 생성 목적 및 기관에 따라 개별적으로 수집되고 있으며, 의료데이터에 대한 소유권과 책임 소재의 문제들로 인하여 데이터의 거래 및 연계가 미흡하다. 정부는 이러한 상황에서 의료데이터의 활용을 촉진하기 위하여 데이터 플랫폼을 구축하고 공공의료데이터를 개방하고 있다. 보건의료빅데이터 플랫폼에서는

3) 이코노믹리뷰(2018), "의료데이터 80% 이상 활용 불가...공공체계 현실화 절실", 12월 10일.

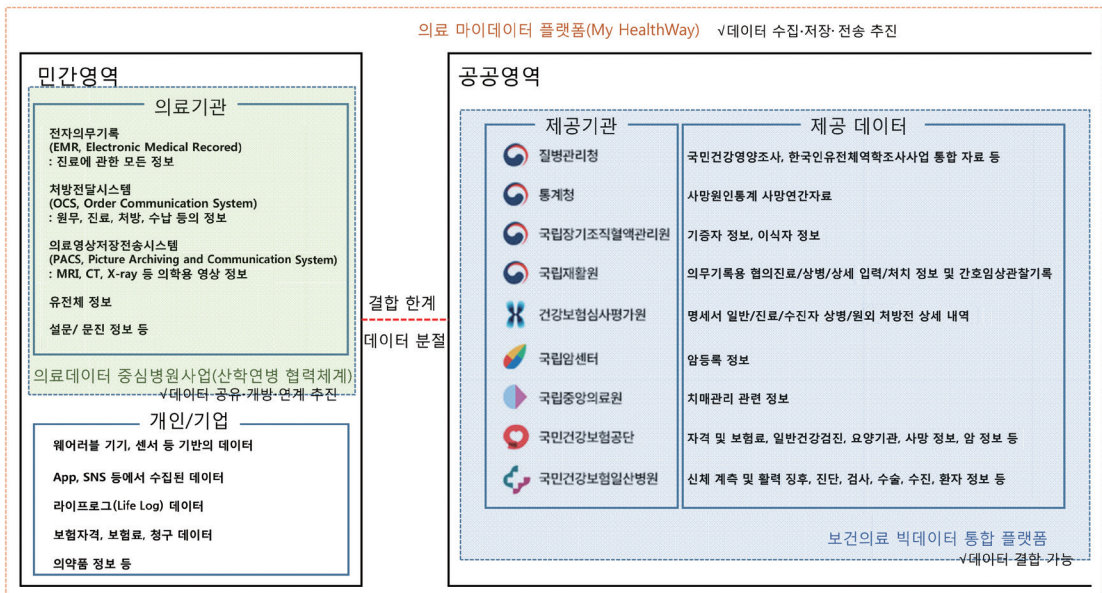
4) 한창호 외(2022), "전자의무기록 데이터 분석 접근법", 『Journal of Health Informatics and Statistics』.

5) 이선진 외(2022), "의료 데이터 산업을 위한 비정형 데이터 비식별화 정책에 관한 연구", 『융합보안논문지』 제22권 제4호, 10월.

질병관리청, 건강보험심사평가원, 국민건강보험공단 등을 포함하여 9개 공공기관이 보유한 의료 데이터를 개방하고 있다. 연구자가 데이터 이용 신청을 하면 연구의 사회적 기여도와 데이터 제공에 대한 적정성을 연구평가위원회 및 제공기관이 심의하고 데이터의 결합과 추가적인 비식별처리를 실시하여 반출이 승인된 결합데이터를 한국보건의료정보원, 건강보험심사평가원, 국민건강보험공단에서 운영하고 있는 폐쇄망 분석센터를 통해 열람할 수 있다. 아쉬운 점은 플랫폼 내 공공기관 데이터는 결합과 분석이 가능하지만 민간영역 의료데이터와는 연계가 불가능하다는 점이다. 민간영역 의료데이터는 개인정보보호법 개정에 따라 가명 조치를 통한 제3자 제공 등의 활용을 위한 데이터 제공의 법적 근거가 마련되었으나, 데이터의 소유권과 문제 발생 시 책임소재, 데이터

제공을 위한 인프라 구축 비용 등으로 인하여 공공데이터만큼 유통 및 개방이 쉽지 않다. 민간영역 데이터 활용을 위해 의료데이터 중심병원사업을 추진하여 각 의료기관마다 상이한 의료데이터를 표준화하고 공유하고 있으나 현재까지는 목적에 따라 컨소시엄을 맺은 기관 사이에서만 진행되고 있다. 최근에는 개인이 자신의 의료데이터에 대한 통제권을 지니면서 데이터를 관리하고 경제적 활용에 동의하는 의료데이터 거래 모델이 데이터 활용 촉진 방안으로 등장하기 시작했는데, 정부가 추진하고 있는 마이 헬스웨이 플랫폼 사업이 바로 그것이다. 여러 곳에 흩어져 있는 의료데이터를 개인 중심으로 통합하여 개인의 선택에 따라 데이터를 활용 기관으로 전송하고 의료서비스 질 제고에 활용토록 하는 것이 마이헬스웨이 플랫폼 구축 사업의 목적이다.

<그림 1> 공공 및 민간영역 데이터 현황



자료: 저자 작성.

### (3) 의료데이터의 활용 현황

의료데이터는 그 자체가 상품이 되어 거래되기도 하고 데이터를 분석하여 개인 맞춤형으로 질병의 예방, 진단, 치료, 관리까지 의료 서비스를 제공하는 형태로 활용되기도 한다.

의료데이터 거래 플랫폼으로 가장 대표적인 기업은 미국의 ‘페이션즈라이크미(PatientsLikeMe)’다. 페이션즈라이크미는 플랫폼에 가입한 환자들의 투약, 질병 치료 후기를 공유하는 서비스 플랫폼을 운영하면서 수집한 투약 정보, 부작용, 증세, 가족력, 신체 정보 등을 익명화하여 제약기업이나 연구기관에게 판매하는 비즈니스 모델을 가지고 있다. 국내에서도 의료데이터를 거래하는 비즈니스 모델을 가진 기업들이 등장하고 있다. ‘휴먼스케이프(Humanscape)’는 희귀질환 환자들이 입력한 데이터를 기반으로 환자들에게 치료제 개발 현황, 질환 정보, 임상 시험 정보 등을 제공하는 ‘레어노트(rarenote)’ 서비스를 제공하면서, 이렇게 축적된 데이터를 신약 개발, 연구, 치료 등 데이터를 필요로 하는 곳에 제공하는 사업을 하고 있다.<sup>6)</sup>

최근 데이터의 활용이 두드러지고 있는 분야는 치료 영역으로 디지털 치료제를 소개하고자 한다. 디지털 치료제(디지털 치료 기기, DTx)는 환자의 데이터를 이용하여 생활 습관을 개선하도록 인지행동치료를 제공하여 질병에 대한 기존 치료(의약품 등)에 대한 순응도를 높임으로써 환자가 가지고 있는 질병을 치료하는 소프트웨어 기기이

다. 국내에서는 2023년 2월 처음으로 ‘(주)에임메드’가 개발한 불면증 인지치료 소프트웨어 ‘솜즈(Somzz)’가 식약처의 허가를 받았으며, 4월에 두 번째로 ‘웰트(주)’가 개발한 불면증 인지치료 소프트웨어 ‘WELT-I’가 디지털 치료 기기로 허가 받았다. 불면증 인지치료 소프트웨어는 불면증을 지속시키거나 악화시키는 심리·행동·인지적 요인들에 대한 중재를 목표로 하는 인지행동치료법을 모바일 앱으로 구현<sup>7)</sup>한 것으로 수면일기를 작성하고 그 데이터를 바탕으로 수면제한요법, 수면습관 교육, 이완요법 등을 일정 기간 동안 제공하여 불면증 증상을 개선할 수 있도록 돕는다. 디지털 치료 기기로 허가받기 위해서는 ‘작용원리에 관한 자료’, 즉, 치료 작용기전에 대한 과학적(임상적) 근거를 제출해야 하고, 허가 이후에도 실사용 데이터(RWD, Real World Data)를 활용한 실사용 증거(RWE, Real World Evidence) 자료를 마련하여 지속적인 유효성 근거를 마련해야 하므로 지속적인 데이터의 수집과 분석이 요구되는 분야이다.

개인이 가진 고유한 특성인 유전 정보의 분석도 서비스로 제공되고 있다. 소비자 직접 의뢰(DTC, Direct to consumer) 유전자 분석 서비스란 특정 항목의 유전자에 대해 의료기관을 방문하지 않고 소비자가 직접 유전자 검사 전문 기관을 통해 받을 수 있는 유전자 검사를 말한다.<sup>8)</sup> 소비자가 직접 유전자 검사를 할 수 있는 시료인 타액이

6) 휴먼스케이프 회사소개서.

7) 식품의약품안전처(2023), “소프트웨어(앱)으로 질병 치료하는 디지털 치료 기기 환자 선택권 넓히고 편의성 높인다”, 보도자료, 2월 15일.

8) 김민영(2021), “DTC 유전자를 활용한 개인 맞춤 서비스의 최신 동향”, 『BRIC View 동향리포트』.

〈표 1〉 보건의료데이터 활용 국내 주요 기업 사례

| 데이터 활용 형태            | 기업명    | 주요 사업 모델  |
|----------------------|--------|---|
| 환자 제공 의료 정보          | 휴먼스케이프 | - 환자가 입력한 희귀질환 데이터를 기반으로 치료제 및 임상시험 정보 등을 제공하고 적절한 임상시험 참여를 제안하는 서비스를 제공<br>- 커뮤니티와 서비스 제공을 통해 축적한 난치질환 환자 정보를 데이터로 가공하여 제약사, 병원, 연구기관 등 데이터가 필요한 곳에 제공하는 사업                                      |
| 생체 및 증상 정보-디지털 치료    | 에임메드   | - 국내 첫 디지털 치료 기기로 허가된 불면증 증상 개선을 목적으로 한 인지치료 소프트웨어 '솜즈(Somzz)'<br>- 모바일용 애플리케이션에 알고리즘을 순차적으로 적용하여 불면증 환자들에게 교육, 실시간 피드백, 행동 중재 및 푸시 알림 메시지 등을 6~9주간 제공  |
|                      | 웰트     | - 국내 2호 디지털 치료 기기로 허가된 불면증 인지행동치료 소프트웨어 'WELT-I'<br>- 환자가 입력하는 '수면 일기' 데이터를 기반으로 적정 취침 시간, 행동 중재, 수면 방해 습관 분석, 이완요법 등을 6주간 수행   |
| 의료 영상 정보 인공지능 솔루션 개발 | 뷰노     | - 딥러닝 엔진인 VunoNet을 기반으로 뼈(VUNO Med-Bon Age), 뇌(Deep Brain), 흉부(Chest X-ray, LungCT AI), 안저(Fundus AI)의 영상을 분석해 질환의 진단을 보조하는 인공지능 솔루션 개발   |
| 유전 정보 분석             | 마크로젠   | - 개인 유전체 분석 서비스 '마이저놈스토리'는 의료기관을 통해 받을 수 있는 질병 예측 유전자 검사와 소비자가 의료기관을 방문하지 않고 직접 개인의 특성이나 웰니스 항목을 검사할 수 있는 헬스&뷰티 유전자 검사를 제공<br>- 영양소, 운동, 피부/모발, 식습관, 니코틴 의존성, 카페인 대사, 비만, 중성지방 농도 등의 항목에 대한 분석 제공 |

자료: 저자 작성.

나 상피 세포를 채취하면 서비스 분석 기관은 채취된 시료를 분석하여 유전자 검사 결과를 소비자에게 전달하는 과정으로 서비스가 이루어진다. 가장 대표적인 DTC 분석 서비스 기업인 미국의 '23andMe'는 유전자 검사를 통한 조상 찾기, 즉 혈통에 대한 정보를 알려주는 서비스를 통해 유전자 검사 서비스에 대한 소비자들의 관심을 끌었으며, 2018년 DTC 서비스 기업 중 가장 처음으로 암과 관련된 유전자(BRCA1/BRCA2) 변이에 대한 검사를 제공할 수 있는 FDA 검증을 받았고, 이후 대장암(MUTYH), 2022년 1월에는 전립선 암(HOXB13)과 관련된 유전자 분석도 처방전 없이 서비스를 이용할 수 있도록 FDA 승인을 받았다.<sup>9)</sup> 이에 반해 현재 우리나라에서는 운동 능력, 영양

소, 피부, 모발 등 웰니스 영역에 국한되어 서비스가 이루어지고 있다. 2016년 “의료기관이 아닌 유전자 검사기관이 직접 실시할 수 있는 유전자 검사 항목에 관한 규정”과 관련한 보건복지부 고시에 따라 DTC 유전자 검사 서비스가 허용되면서 12가지 검사 항목, 유전자 46개가 규정되었고 그 이후 점차적으로 검사 항목이 늘어나 현재 70개로 확대되었지만 여전히 웰니스 분야에 대한 유전자 검사 항목만 허용될 뿐 질병과 관련된 유전자 검사는 의료기관을 통해서만 받을 수 있다.

9) 23andMe Press Release, “23andMe Receive FDA Clearance for Direct-to-Consumer Genetic Test on a Hereditary Prostate Cancer Marker”, <https://investors.23andme.com/news-releases/news-release-details/23andme-receives-fda-clearance-direct-consumer-genetic-test/>(검색일: 2022. 11. 15).

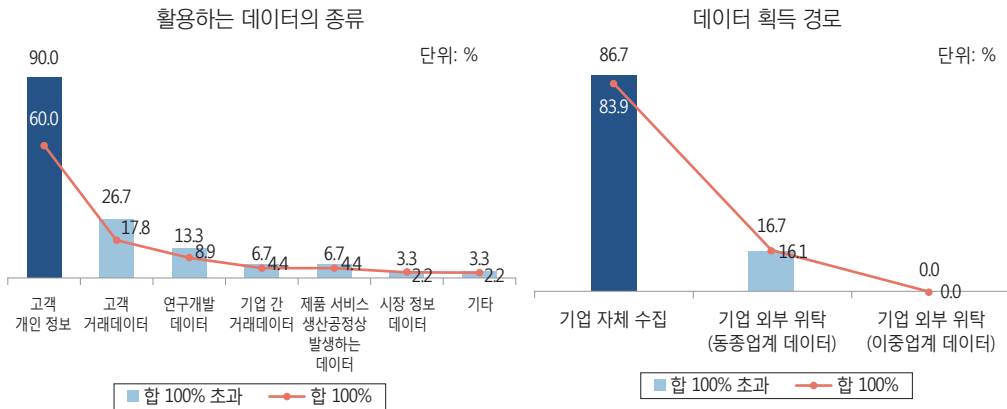
### 3. 국내 보건의료산업에서의 데이터 활용 실태

앞 장에서 의료데이터의 수집-유통-활용이 어떻게 이루어지고 있는지 전반적인 데이터의 가치 사슬에 대해 살펴보았다면, 이 장에서는 실질적으로 기업들의 데이터 활용 실태를 조사한 결과를 소개하고자 한다. 의료데이터의 수집-유통-활용 가치사슬 전 주기에 대한 현황을 살펴보기 위하여 보건 및 의료업에 대한 조사 결과를 제시하고자 한다.

보건 및 의료업에서는 제품 및 서비스의 개발 및 개선(38.5%)을 위하여 고객의 개인 정보를 기

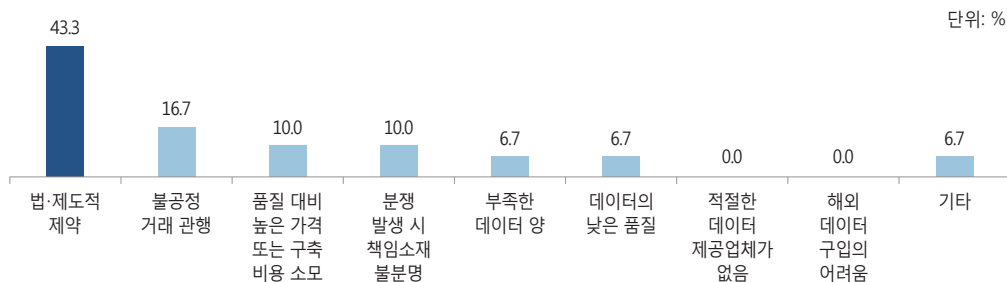
업 자체적으로 수집하여 활용하는 것으로 나타났다. 데이터를 수집하는 데 법·제도적 제약(43.3%)이 가장 큰 장애요인으로 조사되었으며, 동일한 이유인 법·제도적 제약(37.5%)과 분쟁 발생 시 책임소재 불분명(25.0%)의 이유로 외부 데이터를 수집하거나 활용을 위탁하지 않는 것으로 조사되었다. 의료데이터가 민감한 개인 정보이기 때문에 정보 보호를 위한 규제가 타 데이터에 비해 강하게 작용하고 있고 데이터의 소유권에 대한 논의가 여전히 지속되고 있어 유출 등과 같은 문제가 발

〈그림 2〉 활용하는 데이터 종류와 획득 경로



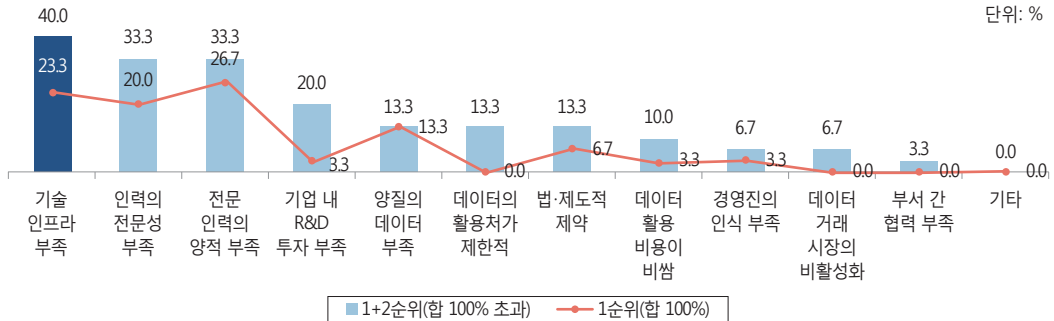
자료: 산업연구원 실태조사.

〈그림 3〉 데이터 수집 및 획득 시 장애요인



자료: 산업연구원 실태조사.

〈그림 4〉 데이터 활용 장애요인



자료: 산업연구원 실태조사.

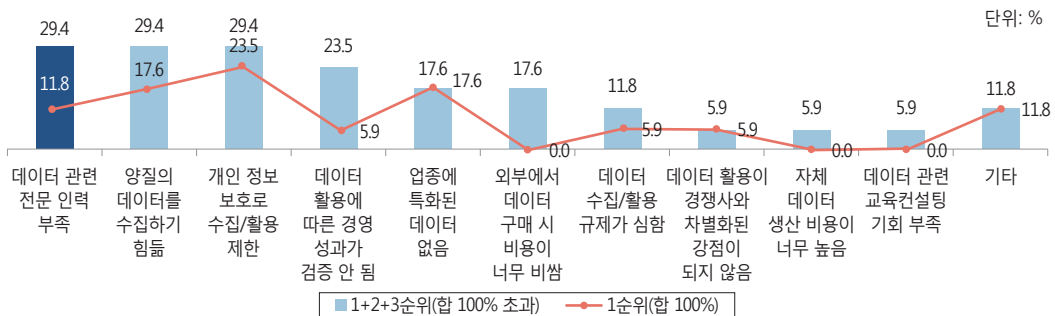
생했을 때의 책임 문제 등에서의 리스크를 최소화 하기 위하여 기업 자체적으로 수집하는 비중이 높은 것으로 보인다. 기업들은 외부 데이터를 구입할 의향이 있지만 기업 자체적으로 수집한 데이터를 다른 기업과 거래할 의향은 크지 않은 것으로 조사되었다. 이러한 이유에서 데이터 판매 기업과의 직접 거래 또는 데이터 거래소를 이용한 데이터 거래 이용 경험(23.3%)도 낮다고 조사되어 데이터 유통 시장에서의 데이터 거래가 활발히 이루어지고 있지 않은 것으로 보인다.

데이터 활용 기업들은 활용에 필요한 기술 인프라의 부족, 인력의 전문성 부족 등을 활용의 가장 큰 장애요인으로 응답한 반면, 데이터 미활용기업

은 전문 인력의 부족뿐만 아니라 데이터의 수집 단계에서부터 개인 정보 보호에 대한 제한으로 인해 어려움을 느끼고 있으며, 양질의 데이터를 수집하는 것에도 어려움을 느껴 활용하지 않는 것으로 응답하였다. 정부의 지원에 대한 조사에서는 데이터를 적극적으로 활용하기 위한 정부의 제도적 지원으로 데이터의 표준화(40.4%)와 데이터 품질 인증(23.4%)에 대한 요구가 높은 것으로 조사되었다.

요컨대 보건 및 의료서비스업에서는 제품 및 서비스를 개발하기 위하여 데이터를 수집하고 활용하고자 하지만 의료데이터가 가진 정보의 특성상 개인 정보 보호를 위한 비식별화 조치 등의 법·제도적 제약과 책임 소재의 문제, 양질의 데이터 확보

〈그림 5〉 미활용기업이 데이터를 활용하지 않는 이유



자료: 산업연구원 실태조사.



어려움 등의 이유로 수집·유통·활용에 어려움을 느끼고 있는 것으로 조사되었다. 따라서 데이터의 활용을 활성화하기 위해서는 양질의 데이터를 수

집하고 안전하게 유통하여 결합을 통해 의미있는 가치를 창출할 수 있도록 데이터 가치사슬 전(全) 단계에 대한 정책 지원 제도의 마련이 필요하다.

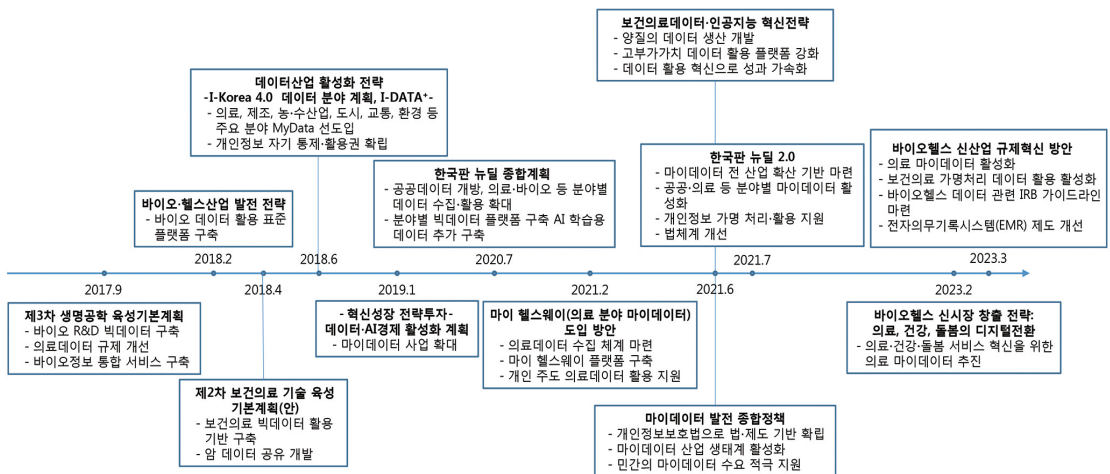
## 4. 보건의료데이터 관련 국내 정책 현황

삶의 질 향상과 더불어 산업 성장의 핵심 자원으로 데이터의 중요성이 부각되면서 정부는 보건의료데이터의 수집과 활용을 위해 다양한 정책들을 추진해 오고 있다. 보건의료데이터 관련 초기 정부 정책들은 생명공학, 과학기술 분야의 기본계획 및 바이오헬스산업 육성 전략 등에 의료 기기 개발 및 기술 개발을 위한 사업의 일환으로 필요한 데이터를 구축하는 것에 초점이 맞추어져 있었다. 그러나 전 산업에 데이터가 활용되고 데이터를 단순하게 수단이나 보조 자원으로 인식하는 것이 아니라 그 자체에 높은 가치를 부여함에 따라

‘데이터 산업’, ‘마이 데이터’, ‘보건의료데이터’ 처럼 데이터가 중심인 데이터 전략 정책이 추진되기 시작하였다. 또한 국가 단위의 디지털 대전환 전략인 한국판 뉴딜 종합계획의 한 분야로 추진되고 있는 디지털 뉴딜의 D.N.A 생태계 강화를 위한 과제인 일명 디지털담 구축에 의료 분야가 포함되면서 국가 단위의 전략에도 보건의료데이터 관련 정책이 포함되어 추진되고 있다.

정부는 의료데이터 활용 확대를 위하여 데이터 3법 개정과 더불어 보건의료데이터 활용 가이드 라인을 발표하였으며, 의료데이터 중심병원사업,

〈그림 6〉 의료데이터 관련 정부 정책 타임라인



자료: 저자 작성.

보건의료빅데이터 플랫폼 구축 사업 등을 통해 민간과 공공이 보유하고 있는 의료데이터의 활용을 활성화하기 위해 노력하고 있다. 정부는 사업들을 통해 활용성이 높은 데이터를 구축하고 폐쇄적인

데이터 활용 환경을 개선함으로써 원활한 데이터 흐름을 만들기 위하여 중장기적 관점에서 데이터의 구축부터 통합·연계, 개방을 통한 활용까지의 방안을 모색하고 있다.

## 5. 결론 및 시사점

의료데이터는 민감 정보라는 데이터의 특성상 안전성과 윤리성을 고려하는 것이 매우 중요하기 때문에 데이터 활용의 필요성에 대해서는 인식하고 있음에도 불구하고 다른 데이터들에 비해 활용도가 낮은 것이 현실이다. 의료데이터의 효율적인 활용을 위하여 의료데이터의 수집부터 분석을 위한 의료데이터의 거래(이동성), 마지막으로 의료데이터의 활용 단계 각각에 대해 정책 방안을 생각해 보고자 한다.


첫째, 데이터 생태계를 활성화하기 위해서는 데이터 구축이 필요하다. 데이터의 구축을 위해서는 무엇보다 먼저 다양한 이해관계자 사이의 합의가 필요하다. 의료데이터를 생성하는 정보의 주체, 의료기관, 활용기업(데이터 이용자), 규제 당국 등이 함께 데이터 구축 시 이득과 손해가 무엇인지 논의하고 의료데이터 활용에 선행하여 적절한 보호가 이루어짐을 명확히 할 필요가 있다. 또한 사회적 합의를 기반으로 한 자발적인 데이터 구축과 선순환적 데이터 활용생태계 초석을 마련하기 위해서는 선제적으로 정부가 주도하여 산재되어 있는 데이터를 연계하고, 대량으로 확보되어야 할 필요성이 있는 정보를 수집할 필요가 있다.

데이터를 구축할 때에는 직관적이며 이해하기 쉬운 형태의 데이터 등록 절차 및 제공 방식을 설계하고 제공한 데이터가 어떤 방식으로 안전하게 보관되고 있고 어떻게 활용되고 있는지를 데이터 제공자가 원하는 경우, 언제든지 확인 가능한 시스템을 만들어 모든 과정을 투명하게 공개하여 데이터 운영에 대한 신뢰도를 높여야 한다.

둘째, 데이터를 분석했을 때 유의미한 결과를 도출할 수 있을 만큼의 활용 가치가 높은 데이터로 만들기 위해서는 데이터 간 자유로운 이동성이 요구된다. 이를 위해서는 상호 운용이 가능하도록 데이터의 표준화를 통한 호환성 확보가 무엇보다 중요하다. 의료기관에 따라 상이한 데이터의 표준화·디지털화 수준을 고려하여 표준화에 대한 필요성과 중요성에 대한 공감대를 형성하고 연구와 분석에 대한 수요가 높은 분야부터 단계적으로 표준화 대상에 대한 범위를 확장하여 데이터 결합이 확산될 수 있도록 지원해야 한다. 지속적으로 데이터가 거래, 유통되기 위해서는 데이터 품질 기준을 마련하여 고품질의 데이터가 꾸준히 제공될 수 있도록 해야 하며 데이터 거래 또는 전송을 위한 인프라 구축 등에 대한 적절한 인센티브

의 마련도 요구된다. 이러한 노력을 기반으로 기관의 데이터뿐 아니라 개인이 자신의 데이터에 대해 주도권을 가지고 데이터를 필요로 하는 기업이나 기관에 자유롭게 공유, 거래할 수 있는 구조를 만든다면 소비자 친화적인 혁신 서비스의 출현이 가능할 것이다.

셋째, 데이터의 구축과 자유로운 이동이 가능해지면 다양한 활용을 위해 효과적인 법·제도의 운용이 필요하다. 데이터3법이 개정되고 안전한 활용을 위한 가이드라인을 만들고 있지만 여전히 데이터 활용 환경 변화에 대한 대응은 미흡한 것으로 보인다. 산업의 경계가 무너지고 다양한 데이터가 전 산업에 활용되고 있는 시점에서 의료데이터와 이종 데이터를 결합하여 활용하고자 했을 때에는 어떠한 가이드라인을 따라야 하는지 규정이

모호한 경우가 발생할 수 있다. 보건의료데이터 활용 가이드라인에는 의료데이터 간 결합 방식 및 반출 절차를 명시하고 있으나, 의료데이터와 비의료데이터(금융데이터 등)를 결합하여 사용하고자 할 때 어떠한 법령에 따라 어떤 기관에 결합을 신청해야 하는지 관련 절차가 명확하지 않다는 점이 문제가 될 수 있다. 따라서 모든 데이터 영역의 활용을 아우를 수 있는 거버넌스 체계의 구축이 요구된다. 이종 데이터 결합에 대한 운영 체계의 유연성, 데이터 종류별 전담 인력의 구성과 협업 등 이종 데이터 간 결합에 대한 명확한 규정이 마련된다면 데이터 결합에 대한 접근성이 보다 용이해지고 이전보다 효율적으로 결합 과정이 이루어지기 때문에 다양한 분야에서 혁신적인 서비스의 개발이 이루어질 것으로 생각된다. 



최은희

서비스산업연구본부 서비스미래전략실 전문연구원  
ehchoi@kiet.re.kr / 044-287-3845

「새로운 직업/직무 출현에 따른 서비스산업의 인력정책 방향」(공저, 2021)  
「서비스산업 활성화 전략 수립 연구」(공저, 2021)