

02

issue

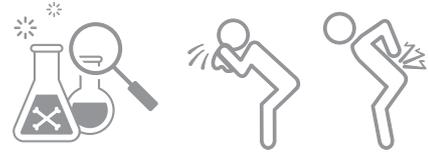


디지털 뉴딜: 빅데이터를 활용한 직업병 위험도 예측

직업건강연구실 역학조사부
이상길 부장
서회경 차장



들어가며

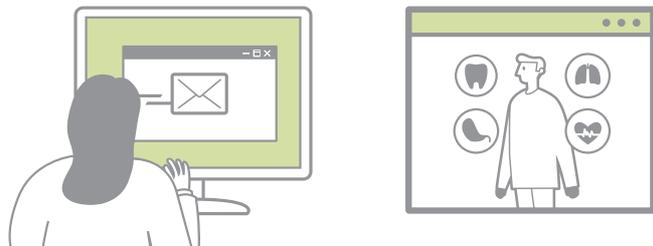


산업이 발달함에 따라 새로운 유해물질의 사용, 생산공정의 변화가 가속화되고 있다. 또한 제4차 산업혁명 등 산업구조 변화에 따라 새로운 직종이나 다양한 방식의 일자리가 생겨나고 있다. 이러한 노동방식, 노동환경의 변화는 기존에 없었던 새로운 직업병을 초래한다.

1988년 원진레이온 사건, 90년대 석면으로 인한 악성중피종, 2000년대 반도체 노동자들의 급성 백혈병까지 직업병은 시대에 따라 변화해왔다. 급변하는 산업구조 및 고용환경에 따른 직업병 이슈에 선제적으로 대응하고, 새로운 유해물질에 신속히 대응할 수 있는 직업병 예방·관리 시스템 구축이 필요한 이유다. 새로운 직업병을 미리 예측하고, 건강에 취약한 직종을 발굴해 다양한 건강유해요인을 찾아내어 맞춤형 대책을 제시할 수 있다면 노동자의 생명과 건강을 지키는데 이보다 좋은 예방책은 없을 것이다.

한편, 정부는 코로나19로 큰 타격을 입은 경기 회복을 위해 구조적 대전환을 이루고자 ‘한국판 뉴딜’ 사업을 추진하고 있다. 디지털 뉴딜, 그린 뉴딜, 안전망 강화를 큰 축으로 분야별 투자와 일자리 창출을 위한 사업을 추진할 계획이다. 이 중 디지털 뉴딜은 코로나19로 디지털 경제로의 전환이 가속화됨에 따라 우리나라의 강점인 ICT 경쟁력을 바탕으로 4차 산업혁명을 선도하는 산업 데이터의 수집과 활용에 주목하고 있다.

한국판 뉴딜을 뒷받침하기 위해서는 공공기관들이 뉴딜의 성과창출을 적극적으로 뒷받침할 수 있는 환경을 조성할 필요가 있다. 공공기관 스스로 공공기관 고유업무에 뉴딜을 접목하는 것으로, 공공기관 빅데이터 활용 촉진 및 전 산업 디지털화, 공공자원의 공유 자원화, 디지털·비대면 활용 공공서비스 혁신 등이다. 산업안전보건연구원은 ‘빅데이터를 활용한 직업병 위험도 예측 사업’을 한국판 뉴딜과제로 제안하여 지난 8월 말 최종 선정되었다. 보고는 빅데이터를 활용해 국민 안전을 위한 다양한 예측 시스템 개발 사례들을 살펴보고, 연구원에서 제안한 직업병 위험도 예측 사업의 특징과 향후 계획을 소개하고자 한다.



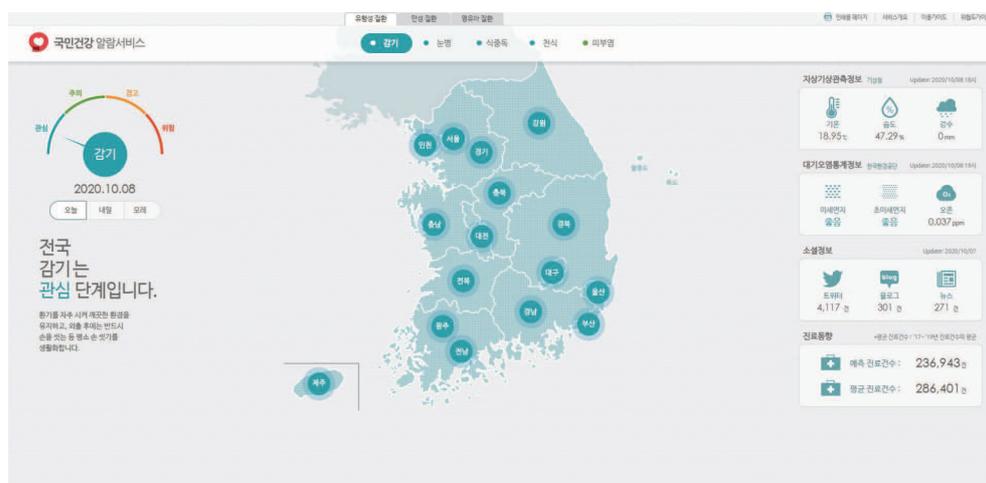
공공 빅데이터를 활용한 예측 시스템 사례

1 국민건강보험 국민건강알람서비스

최근 국민의 생명과 안전을 위해 빅데이터를 활용한 다양한 예측 시스템들이 활용되고 있다. 국민건강보험공단은 2014년에 식품의약품안전처, 기상청, 국립환경과학원과 협업하여 ‘국민건강알람서비스’를 개발했다. 국민건강알람서비스는 식품의약품안전처에서 제공하는 식중독 확진건수, 기상청에서 제공하는 지상기상관측정보, 한국환경공단에서 제공하는 대기오염통계정보, 소셜미디어 정보 및 언론기사 등 데이터를 활용한 것으로, 주요 질병의 위험도 동향과 알람을 제공하는 서비스이다.

세부적으로는 질병에 영향을 끼치는 기상 환경 요인을 분석해 감기, 눈병, 천식, 식중독, 피부염과 같은 대표적인 질병을 지역별, 계절별로 파악해 유행 전 미리 대응할 수 있도록 하였다. 기상청의 지상기상관측정보, 한국환경공단의 대기오염통계정보를 통해 해당 지역의 기상상황 및 대기오염 정도를 파악할 수 있으며, 해당 질병에 대한 최신의 소셜미디어 피드수 및 언론 기사도 살펴볼 수 있다. 각 질병의 과거 3년간 지역별 평균 진료건수도 확인할 수 있다. 영유아 폐렴, 만성질환 고위험자의 COPD(만성폐쇄성폐질환)에 대한 급성악화 가능성 및 유행정도 등의 예측정보도 제공하고 있다.

국민건강보험공단의 국민건강알람서비스는 2018년 국제사회보장협회(International Social Security Association, ISSA) 아태지역 사회보장 포럼에서 유행병 확산 방지 및 국민 건강증진에 기여한 바를 인정받아 특별공로상을 수상하기도 하였다.



[그림 1] 국민건강 알람서비스 화면(<http://forecast.nhis.or.kr/menu.do>)

2 질병관리청 심혈관질환자 임상역학자료 기반 미세먼지 노출 피해 위험예측모델¹⁾

질병관리청은 2020년 ‘심혈관질환자 임상역학자료 기반 미세먼지 노출 피해 위험예측모델’ 연구 용역을 추진하고 있다. 심혈관질환자 고위험군 자료와 기상청의 미세먼지 노출자료를 연계해 위험도 예측모델을 구축하는 것을 목표로 삼고 있다.

미세먼지 노출 피해 위험예측모델 구축에 활용하는 자료는 인체영향평가자료(의료기관 전자 의무기록 등록자료(EMR), 한국인 심근경색증 또는 심부전 등 등록관찰 코호트 자료) 등 임상 등록 또는 추적조사 기반 자료와 한국환경공단, 국립환경과학원 및 기상청 대기 기상 자료와 같은 대기질 평가 자료다. 임상등록 또는 추적조사 기반 자료와 미세먼지 노출 자료를 연계해 지역별, 대기오염 물질별 노출 자료를 제공하고, 이를 데이터베이스화 할 계획이다. 자료가 연계 되면 월별, 연간 노출량 자료 등 지표를 생산하는 것도 목표로 하고 있다. 지역정보를 반영한 개인별 맞춤형 미세먼지 노출 유효 기준치 모델은 머신러닝 알고리즘을 활용하여 위험도 예측 모델을 개발할 계획이다.



3 건강보험심사평가원 보건의료 빅데이터 개방시스템

건강보험심사평가원은 1977년부터 우리나라 건강보험제도와 함께하면서 보건의료 데이터의 표준(분류, 코딩, 적용기준 등)을 만들어 일찍부터 ICT와 접목시키면서 이를 정제·축적해왔으며, 2015년 공공기관 최초 수요자 중심 맞춤형 보건의료 빅데이터 개방시스템을 구축하였다.

국내 보건의료 분야 빅데이터 분석·활용 기반을 마련하였으며, 공공과 민간의 폭넓은 활용을 위해 보건의료 빅데이터센터 및 원격서비스를 운영하고 있다. 건강보험심사평가원에서 보유·개방하는 의료데이터는 전국민의 진료정보와 의료기관, 제약회사, 유관기관 등 다양한 경로에서 수집한 의약품정보, 치료재료정보, 의료자원정보, 비급여정보, 의료질평가정보 등을 분석·정제한 데이터로, 공공데이터(데이터셋, 오픈API), 의료빅데이터분석(빅데이터센터, 원격가상화분석), 의료통계분석(질병, 의약품, 의료기관) 등 다양한 서비스를 제공함으로써 국민과 보건의료 산업 분야, 의료연구기관 등 이용자에게 유익한 의료정보와 서비스를 지속적으로 발굴·개방하고 있다.

1) 질병관리청 홈페이지 연구용역 공고문 참조(2019-10-18) “2020년 학술연구개발용역(심혈관질환자 임상역학자료 기반 미세먼지 노출 피해 위험예측모델 개발) 입찰 공고 안내”

직업병 예측 시스템 개발 주요내용

1 사업개요

산업안전보건연구원에서 추진하고자 하는 직업병 예측 시스템의 목표는 산업보건 데이터 웨어하우스(Data Warehouse)를 구축해 직업병 예측 모델을 개발하고 직업병 발생 위험도 정보를 제공하는 것이다. 즉, 다양한 직업병의 발생 원인을 규명하기 위해 노출과 반응관계 증명 등 과학적 분석을 실시하고, 이를 통해 직업병과의 연관관계를 파악하기 위한 것이다. 소요 기간은 2021년부터 최대 2027년까지 약 7년이며, 올해는 최종 사업계획을 수립하고, 사업 운영을 위한 TF 팀을 구성해 사업추진의 발판을 마련하고자 한다.

시스템에 활용할 산업보건 공공데이터는 산업재해통계, 작업환경측정자료, 특수건강진단자료 등 고용노동부 자료를 중심으로 타 부처 정보를 포함할 예정이다. 산업안전보건연구원은 2014년 국립암센터, 2018년 국민건강보험공단, 2020년 건설근로자공제회와 공공기관 업무협약(MOU) 체결을 통해 공공데이터 수집 및 연계 기반을 조성해왔다.

2 직업병 예측 시스템 운영 체계

1 데이터 수집 및 환경 분석

다양한 산업보건 데이터를 수집하고 이를 연계하기 위해서는 데이터의 속성을 파악해야 한다. 연계 대상으로 살펴보고 있는 산업보건 관련 데이터는 작업환경측정자료와 작업환경실태조사 자료, 특수건강진단 자료와 근로환경조사 자료 등의 고용부 자료와 국민건강보험 자료, 암등록 자료, 통계청 사망자료 등이다. 각 데이터를 연계하기 전 데이터별 범주와 형태를 살피고 특성을 파악하는 과정과 데이터의 신뢰도 조사가 필요하다. 데이터 통합관리 시스템 구축을 위해 정보화 전략계획(ISP)을 수립할 예정이며, ISP 수립은 2021년을 목표로 삼고 있다.

2 데이터 연계 체계 구축

연계할 데이터의 환경 분석이 완료되면, 각 데이터의 수집 경로와 연계할 체계를 설계한다. 데이터 연계를 위해서는 표준코드 구축 등의 데이터 표준화 작업이 선행되어야 하는데, 2021년부터 작업환경측정 자료와 특수건강검진 자료 표준화 작업을 시작할 계획이다. 데이터 정제 작업과 함께 개인정보 보호에 따라 대체 식별화를 처리할 제반 기술도 구현해야 하며, 국민건강보험 자료 등 타 부처 자료 연계를 위한 법제도적 장치에 대한 검토와 지원도 필요하다. 데이터 연계를 위해 필요한 과정들을 거치고 나면, 본격적으로 산업보건 통합 DB를 구축할 수 있다. 산업보건 통합 DB를 효율적으로 활용할 수 있는 데이터 웨어하우스(DW)를 구축하면 직업병 예측 시스템 운영 체계의 뼈대가 완성될 것으로 보인다.

3 데이터 보안 및 관리 체계 수립

데이터 활용에 있어 가장 중요한 부분은 보안체계를 마련하는 것이다. 통합에 활용하는 데이터는 근로자 개인의 민감한 정보가 많이 포함되어 있으므로 보안지침을 마련하고, 데이터 반출이나 재식별의 위험이 없도록 보안 사항을 살펴야 한다. 향후 데이터를 개방할 때 보안이나 위험성의 문제는 없는지도 꼼꼼히 살펴야 한다. 또한 각 데이터들의 품질관리 및 정보 업데이트 체계도 필요하다. 데이터 오류 검출과 수정 등 데이터 품질관리와 각 데이터의 지속적 업데이트 방안도 데이터 관리 방안 수립을 통해 검토할 것이다.

3 데이터 개방 관련 사항

통합 데이터를 개방해 다양한 연구결과를 창출하기 위해서는 데이터를 활용하고자 하는 사람 누구나 이해할 수 있는 표준 DB를 제공해야 한다. 예를 들어 특수건강진단 코호트 자료를 구축해 이를 표준 DB로 제공할 수 있다. 또한 데이터를 활용하고자 하는 연구자, 정책입안자 등 다양한 고객 수요 맞춤형 DB를 지속적으로 제공해 데이터 접근성을 높이고 관련 연구를 활성화 할 계획이다. 산업보건 통합 DB 분석 결과는 일반인들에게도 쉽게 제공할 수 있도록 시각화 자료 및 직업병 발생 현황 연보 등을 통해 제공할 것이다.



[그림 2] 직업병 예측 시스템 운영 체계(안)

3 직업병 위험도 예측 시스템 활용

데이터 수집과 연계, 보안 및 개방관련 사항을 기본적으로 정립한 후, 이를 토대로 직업병 위험도 예측 시스템을 개발하고자 한다. 예측 시스템의 기본 방향은 일반국민, 근로자, 사업주 등 다양한 수요를 충족할 수 있도록 예측식을 활용한 위험도 지수 개발, 직업병 유형별 영향요인 분석자료 제공 등이다. 예를 들면 업종별, 지역별로 자주 발생하는 질병이나, 사업장의 특정 위험 지표 등의 정보 제공 같은 것이다.

예측 시스템은 궁극적으로 직업병 유형별 고위험군 도출에 활용하여 역학조사를 수행하게 하고, 이러한 역학조사 평가 결과를 활용해 직업병 예방사업을 수행할 수 있게 할 것이다. 고위험 사업장을 대상으로 맞춤형 예방 서비스를 수행할 수 있으며, 예방사업의 효과에 대한 피드백을 통해 전반적인 산재예방 서비스의 질을 높이는 데 기여할 것이다. 직업병 위험도 예측 시스템은 산업보건시스템의 공공성을 확장하게 하는 도구로 활용할 수 있다. 즉, 산업보건 공공데이터 허브로 작동하며 직업병 예측 모델의 개방을 통해 일하는 사람 누구나 자신의 건강위험도를 예측할 수 있게 될 것이다.



[그림 3] 직업병 위험도 예측 사업 절차(안)

빅데이터를 활용한 직업병 예측 시스템에 거는 기대

코로나19 이후 국민의 기본생활 유지와 안전과 관련한 필수노동자인 배달노동자, 요양·돌봄 노동자, 의료기관 종사자들에 대한 관심과 우려가 높아지고 있다. 배달 물량이 많아지면서 발생한 택배 노동자의 과로사 문제는 코로나19 이후 사회 변화가 우리에게 던진 새로운 숙제이다. 이렇듯 급변하는 산업구조와 고용환경의 변화에 발 빠르게 대응하기 위해서는 다양한 직업병 이슈에 선제적으로 대응할 수 있는 직업병 예측 시스템이 필요하다.

산업보건 DB를 통합, 연계한 직업병 예측 시스템을 통해 직업병 고위험군을 발굴하고, 이 결과를 직업병 예방 사업에 선제적으로 활용한다면 더 많은 노동자의 생명과 건강보호에 기여할 수 있을 것이다. 데이터 분석 결과는 직업병 예방사업의 우선순위 설정에 활용하고, 이를 토대로 다양한 역학조사를 실시해 일하는 사람의 건강에 유해한 요인들을 찾아내고 대책을 제시할 수 있을 것이다. 또한 산업보건 통합, 연계 데이터의 개방은 민간시장 활성화를 유도하는데도 기여할 것이다.

참고문헌

- » 국민건강보험공단 국민건강 알람서비스 <http://forecast.nhis.or.kr/menu.do>
- » 질병관리청 2020년 학술연구개발용역 심혈관질환자 임상역학자료기반 미세먼지 노출 피해 위험예측모델 개발