

김 씨는 "의료 데이터에는 숫자뿐 아니라 CT나 엑스레이 같은 이미지 데이터도 많다"며 "요즘은 이런 데이터를 어떻게 처리해야 할지 공부하고 있다"고 말했다. (사진=신정빈 기자)

데이터에서 생명으로
숫자 너머의 질문을 찾는 시간

신정빈 기자 wjdqls2002@khu.ac.kr

공부하는 사람들 ⑧
정밀의료학 전공

학부 4학년을 마친 학생들은 취업과 대학원 진학의 갈래 앞에 선다. 이중 대학원 진학은 '연구자의 삶'을 택하는 것이다. 이들이 한 명의 연구자로 성장하는 과정에서 만들어내는 '앎'은 개인의 성취인 동시에 사회 공공재가 된다. 이번 학기 대학주보는 '공부하는 사람들'의 이야기에 주목한다. 여덟 번째 주인공은 의료 빅데이터로 전 세계의 질병과 죽음의 흐름을 연구하는 정밀의료학과 김소은(석박통합 6기) 씨다.

끝을 보는 성격 연구자의 길로 들어서다

김소은 씨는 한번 무언가에 꽂히면 끝을 봐야 하는 성격이다. 어릴 적부터 하고 싶은 일이 생기면 부모님이 만류해도 끝내 해내곤 했다. 대학원 진학도 그렇게 마음먹은 일 중 하나였다.

어릴 적 막연히 의사를 멋있다고 여겼던 김 씨는 타 대학에서 데이터사이언스를 전공하며 생명과학 분야와는 멀어졌다. 그러던 중 당시 학과장이 의료 데이터를 연구하는 모습을 보며 두 분야를 접목할 수 있다는 사실을 알게 됐다. 의료 데이터는 생명과학에 대한 관심과 전공 지식을 함께 활용할 수 있는 분야였다.

의료 데이터는 개인정보와 밀접해 접근이 쉽지 않고, 연구자 수도 적은 분야다. 김 씨에게는 낯선 영역을 탐구한다는 점이 매력으로 다가왔다. 김 씨는 "아직 많은 사람이 하지 않는 분야를 공부하는 것이 새로운 걸 하고 있다는 생각에 더 재미있었다"고 말했다.

김 씨는 학과장의 소개로 학부 3학년 때 경희디지털헬스센터에서 학부연구생으로 연구를 시작했다. 예상과 다른 결과가 나오기도 하고, 처음 구상한 분석 방법을 선행 연구에 맞춰 수정해야 할 때도 물론 있었다. 하지만 김 씨는 오히려 새로운 방법을 찾아가는 일이 자신의 성격과 잘 맞는다고 느꼈다. 박사과정까지 공부를 이어가겠다고 마음먹은 계기였다. 김 씨는 "연구직으로 진로를 이어가고 싶다는 생각도 있었지만, 무엇보다 연구 자체가 재미있어 지금까지 계속하고 있다"고 말했다.

스스로 답을 만들어가는 대학원에서의 공부

연구를 시작하며 공부하는 방식도 달라졌다. 학부 시절에는 교과서에 정리된 지식과 분석 방법론을 배웠다면, 대학원에서는 실제 데이터의 특성에 맞춰 분석 방법부터 스스로 찾아야 했다. 새로운 사실을 발견하는 것만큼이나, 의문을 품는 태도가 중요하다는 설명이다. 김 씨는 "남들은 그냥 넘어가는 것도 '왜 이렇게 될까'를 생각해야 연구를 할 수 있는 것 같다"고 말했다. 이어 그 차이를 "글로 보는 것과 현장에서 직접 배우는 것"이라고 표현했다.

학부 연구생 시절 첫 프로젝트는 세계보건기구(WHO)의 사망 통계를 활용한 연구였다. 질병별 사망 추세를 분석하고 2050년까지 사망 규모가 어떻게 변화할지 예측하는 연구였다. 처음이었기에 선행 연구를 찾고 분석 방법을 정하는 것부터 쉽지 않았다. 김 씨는 관련 논문과 자료를 찾아가며 자신만의 결과를 도출했다. 그는 "새롭게 내 것을 만들어간다는 게 재미있었다"며 "당시 익힌 것들이 지금까지 새로운 연구를 하는 데 좋은 기반이 된다"고 말했다.

낯선 의학 지식은 김 씨가 넘어야 할 과제였다. 연구 초반 제2형 당뇨병 관련 선행 연구를 읽을 때는 'T2DM'이라는 약어가 반복해서 등장해 당황하기도 했다. 김 씨는 "이게 도대체 뭐길래 계속 나오지 싶었는데 알고 보니 제2형 당뇨병(Type 2 Diabetes Mellitus)을 의학 용어로 짧게 줄인 말이었다"고 당시를 떠올리며 웃음을 띠었다.

부족한 의학 지식은 동료들과 역할을 나누며 채워갔다. 김 씨가 데이터 분석과 AI 모델링을 맡으면 의학을 공부하는 동료들은 분석 결과가 인체의 작동 원리, 즉 '기전'에 부합하는지 함께 검토한다. 데이터상 연관성이 나타나더라도 사람의 몸에서 같은 현상이 발생할 수 있는지 의학적으로 살펴야 하기 때문이다. 김 씨는 "서로 다른 전공의 연구자들이 의견을 나누는 과정이 즐겁다"며 "혼자였으면 연구를 포기하고 싶을 수도 있었을 것 같다"고 말했다.

김 씨는 새롭게 찾아낸 결과가 의학적 기전으로도 설명될 수 있다는 사실을 확인할 때 특히 보람을 느낀다고 했다. 김 씨는 "혹시 저희의 해석이 오해나 잘못된 정보를 낳지는 않을지 항상 고민한다"며 "의학적·임상적으로 타당한지 수없이 검증하고, 확신이 들 때 결과를 제출한다"고 말했다.

데이터로 세상을 보는 법 원인을 좇는 질문 필요

첫 WHO 연구에서 김 씨는 국가별 소득 수준에 따라 사망 원인에도 차이가 나타난다는 사실을 발견했다. 고소득 국가에서는 대사질환이나 자살로 인한 사망이 두드러진 반면, 저소득 국가에서는 영양 부족이나 감염병으로 인한 사망이 많았다. 하지만 숫자가 모든 현실을 반영하는 것은 아니라는 설명이다. WHO 통계는 각국이 보고한 사망 기록을 모은 자료이기 때문에, 사망 보고 체계가 충분하지 않은 국가에서 발생한 죽음은 통계에서 반영되지 않기 쉽다. 김 씨는 "보고되지 않았을 뿐, 그런 나라에서 오히려 더 많은 사람이 숨지고 있을 수 있다"고 말했다.

김 씨는 계속해서 질병과 사망에 관한 데이터를 분석하고 있다. 현재는 전 세계 204개국에서 380여 개 질병에 관한 사망 현황을 담은 '세계질병부담(GBD)' 데이터가 주 연구 자료다. 김 씨는 "자살 관련 연구를 할 때마다 우리나라가 다른 나라에 비해 유독 높다는 게 계속 보였다"며 "안타깝기도 하고, 사회적으로 함께 풀어가야 할 문제라는 생각이 들었다"고 전했다.

김 씨를 이끄는 원동력은 '누군가의 건강과 삶 개선에 보탬이 될 수 있다는 가능성'이라고 했다. 그는 "데이터 하나하나가 누군가의 실제 임상 경과이자 삶의 기록이라는 점을 떠올릴 때 분석에 임하는 마음가짐을 다시금 다잡게 된다"고 말했다.

앞으로의 숙제는 연관성을 넘어 원인을 찾는 것이다. 김 씨는 병원 전자의무기록(EMR)을 활용해 질병이 왜 생기는지까지 밝혀보고 싶다고 말했다. 스스로를 "새로운 걸 찾기 위해 일단 열심히 해보는 사람"으로 정의한 김 씨는 데이터 속에서 다음 질문을 찾고 있다.



김 씨는 올해 '세계 상위 2% 연구자'로 선정됐다. (사진=신정빈 기자)