

“AI에 데이터 읽는 법 학습시켜 … 세상에 없던 통계 서비스 선보일 것”

통계청서 ‘국가데이터처’로 승격 … 안형준 초대 처장

“인공지능(AI)에 통계 읽는 법을 학습시켜 모든 국민이 손쉽게 통계를 활용할 수 있는 서비스를 2027년께 내놓을 계획입니다.”

안형준 초대 국가데이터처장은 13일 한국경제신문과의 인터뷰에서 “생성형 AI를 활용해 통계 문턱을 낮추겠다”며 이같이 말했다. AI가 통계 숫자는 물론 맥락까지 이해하도록 학습시켜 예컨대 ‘지역 인구 감소와 여성 취업률의 상관 지수’ 같은 복합적인 질문에 대답할 수 있도록 하겠다는 설명이다. 지금까지는 경제학자 등 전문가가 아니면 통계를 전문적으로 해석하고 가공하는 것이 불가능했지만, 이런 서비스를 이용하면 일반 국민도 통계를 다양한 분야에 활용할 수 있다.

안 처장은 1990년 출범한 통계청의 첫 내부 출신 청장으로 지난 8월 취임했다. 기획재정부 외청이던 통계청이 지난 1일 국가데이터처로 승격되면서 첫 국가데이터처장에 올랐다. 그는 “국가데이터처는 통계뿐 아니라 통계의 원재료인 정부 부처 데이터까지 폭넓게 활용하자는 취지에서 출범했다”며 “부처 간 ‘데이터 칸막이’를 없애고 민간 부문 데이터와 연계해 지금까지 없었던 새로운 통계 서비스를 제공할 것”이라고 말했다.

▶통계청이 국가데이터처로 승격됐습니다. 가장 달라지는 점은 뭔가요.

“청과 달리 처는 독자적인 입법권을 갖게 됩니다. 통계도 사회와 산업구조 변화에 맞춰 바꿀 필요가 있는데, 기존 통계법은 기재부 소속법이다 보니 원하는 시점에, 원하는 방향으로 개정하는 게 쉽지 않았습니니다. 자체 입법권을 가지면 통계 분류나 신설, 폐지, 개편 등을 독립적으로 할 수 있다는 차이가 있습니다.”

▶부처 이름에서 ‘통계’를 빼고 ‘데이터’를 추가한 이유가 있습니까?

“국가데이터처의 핵심은 부처 간 데이터 칸막이를 없애는 겁니다. 지금도 행정안전부, 과학기술정보통신부, 보건복지부 등 각 부처가 자기 부처 소관 데이터를 관리할 수 있는 권한이 있습니다. 다만 해당 데이터를 연계, 조정하는 기능을 가진 곳은 아무 데도 없습니니다. 이게 국가데이터처에 주어진 미션입니다. 예컨대 최근 내놓은 ‘포괄적 연금통계’도 기존에 저희가 보유한 데이터에 각 부처의 모든 연금데이터(주택연금, 공무원연금, 군인연금, 사학연금, 기초연금 등)를 더해 만든 겁니다.”

▶임기 동안 이것만큼은 이루고 싶은 게 있다면요.

“AI 메타데이터베이스를 반드시 성공시키고 싶습니다. 지금의 생성형 AI는 대규모언어모델(LLM) 기반이기 때문에 문장을 읽는 것은 잘하지만 통계와 숫자, 표를 해석하는 능력은 떨어집니다. 통계청은 이 한계를 2년 전부터 포착하고 통계 숫자를 읽는 것은 물론 흐름까지 분석할 수 있는 별도의 메타데이터를 구축하고 있습니다. 통계의 맥락도 파악할 수 있어 ‘물가와 국민소득의 관계’ 혹은 ‘외국인 직접투자가 지방 건설 경기에 미치는 영향’ 같은 전혀 다른 분야를 조합한 질문에도 답할 수 있게 됩니다.”

▶메타데이터는 구체적으로 어떻게 만드는 건가요?

“우리나라 국가 통계가 1380종 정도 됩니다. 우선 통계마다 모두 메타데이터를 구축한 다음 이걸 통합하는 작업이 필요합니다. 내년 초에는

쏜부처 데이터 총괄 ‘처’로 격상

독자 입법권 갖고 통계 분류·개편 가능
부처간 데이터 칸막이 없애는 게 핵심

생성형 AI 활용해 통계 문턱 낮춘다

통계 흐름 분석해내는 ‘AI 메타데이터’
2027년 3월 구현 목표 … 국제표준 선도
저출생 대책 근간될 새 인구지표도 공개

인구주택총조사에 변화된 사회 반영

지역 균형발전 기준 될 지수 개발 중

통계청 3대 지표 중 하나인 ‘고용동향’ 통계부터 메타데이터를 둘러볼 생각입니다. 필요 예산은 이미 편성돼 있어요. 고용동향에서 잘 작동하는 게 확인되면 그다음 통계로 확산시킬 겁니다.”

▶완성 시점은 언제로 보시는지요.

“2027년 3월 실제로 구현해 돌려보는 것을 목표로 하고 있습니다. 이런 계획을 지난 6월 경제협력개발기구(OECD) 통계청장 회의에서 발표했더니 각국 참석자들이 모두 ‘환영한다’며 박수를 보내더군요. 메타데이터가 제대로 만들어진 국제 표준이 될 수 있습니다. 지금까지 전 세계를 통틀어 통계 숫자와 맥락, 그 흐름까지 제대로 분석해내는 AI는 없었으니까요. 세계 모든 통계 기관이 한국의 메타데이터 알고리즘을 활용해 통계를 읽어내는 세상이 올 수 있는 겁니다.”

▶국가데이터처의 AI가 세계 통계 기관들의 기준이 되는 거군요.

“그렇습니다. 다만 6월 저희와 똑같은 발표를 한 나라가 한 곳 있었습니니다. 스위스인데요. 누가 세계 최초의 ‘AI 메타 기반 통계 기관’이 될지 선두 경쟁이 시작된 셈입니다. ‘우리가 제대로 만들어서 표준화를 선도해보자’라는 생각으로 열심히



안형준 국가데이터처장이 13일 정부대전청사에서 한국경제신문과 인터뷰하고 있다. 안 처장은 “AI에 통계 읽는 법을 학습시켜 일반 국민도 통계를 손쉽게 다양한 분야에 활용할 수 있는 서비스를 2027년께 제공할 것”이라고 말했다. 최혁 기자

히 구축하고 있습니다.”

▶통계청 내부 출신의 첫 청장이자 초대 국가데이터처장이 됐습니니다.

“제가 2016년 ‘통계데이터허브국’ 초대 국장을 맡았습니다. 통계데이터허브국은 통계청 내부 통계부터라도 연결하는 작업이 필요하다는 생각에서 출발했습니다. 당시 빅데이터 열풍이 불 때였죠. 국의 역할을 정립하고 필요 인력을 설계하는 작업을 맡았습니다. 그로부터 약 10년 뒤, 이제는 각 부처 간 데이터 융합과 활용을 총괄하고 거기에 AI까지 더하는 국가데이터처장이 됐습니니다. 한마디로 짜릿합니다.”

▶올해 눈여겨봐야 하는 통계가 있다면요.

“대한민국의 모든 가구, 인구, 주택의 규모 및 특성을 파악하는 ‘인구주택총조사’가 올해로 100년을 맞았습니다. 일제강점기 때부터 시작돼 해방, 6·25전쟁, 민주화 등 우리나라의 현대사가 고스란히 투영돼 있는 통계인데요. 오는 11월 말까지 조사가 이뤄집니다. 5년마다 진행되는 조사인 만큼 변화된 사회 분위기가 항목에 반영되기도 하는데요. 올해는 ‘결혼 의향’ 그리고 ‘비혼 동거’ 여부를 묻는 항목을 신설했습니다. 과거에

는 묻고 싶어도 민감한 주제라 묻기 어려웠던 항목도 시대가 변하면서 함께 조사하게 된 거죠.”

▶지난해 출생아 100명 중 6명은 혼외자라는 통계도 있었습니다.

“그렇죠. 출생뿐 아니라 미혼, 동거, 그리고 ‘가족 돌봄 시간’ 통계도 새롭게 나올 예정입니다. 다문화 가구가 늘어난 만큼 가구 내 사용 언어, 한국어 실력 등의 항목도 추가해 조사합니다. 사회 변화에 따라 통계도 진화하는 셈이죠. 올해는 10년에 한 번씩 이뤄지는 종교 통계도 조사합니다. 우리나라 국민이 가장 많이 믿는 종교는 무엇인지, 무교는 얼마나 되는지 등을 확인할 수 있습니다.”

▶연말에 또 새롭게 공개되는 통계가 있습니까.

“12월에 ‘인구 동태 패널’을 발표할 예정입니다. 1983년생부터 1995년생까지 우리나라 가입 기 인구 850만 명의 삶을 추적한 패널인데요. 저출생 대책의 근간이 되는 자료가 될 겁니다. 이 사람들이 몇 세에 결혼했는지, 맞벌이인지 아닌지, 결혼하고 아이를 낳기까지 얼마나 걸렸는지 등을 조사해 소득, 주택 소유 여부 같은 요인이 최종적으로 결혼과 출산에 어떤 영향을 미쳤는

지 보여주는 통계가 될 겁니다.”

▶새로 개발 중인 경제 통계도 있나요?

“연말에 ‘포괄적 민간소비’ 통계를 발표합니다. 기존 소매판매 통계에는 재화만 잡혔는데 여기에 서비스까지 더하는 개념입니다. 최근에 소비가 늘어나고 있는데, 기존 통계로는 소비를 종합적으로 보지 못한다는 문제점이 있어서 개발하게 됐습니니다. 그리고 지역내총생산(GRDP) 같은 경우에는 작년까지만 해도 1년에 한 번만 발표했는데, 이제부터는 매 분기 발표하는 쪽으로 개선했습니니다. 또 사업체를 조사하는 경제총조사(경총)도 내년 진행하는데 항목에 AI 활용 여부와 스마트공장 등을 추가했습니다.”

▶모두 정책 수립에 도움이 되는 통계군요.

“맞습니다. 이재명 대통령이 여러 번 언급한 ‘지역균형지수’도 지방행정연구원과 업무협약(MOU)을 맺고 개발 중입니다. 기존에 지방시대 위원회에서 만든 지역균등화지수와 달리 하나의 종합적이고 포괄적인 지수를 개발하려고 노력 중이예요. 수도권과의 거리, 인구수 등을 반영한 지수인데 지방자치단체 지원 정책의 기준점으로 삼을 수 있도록 개발할 계획입니다.”

남정민/이광식 기자