

## 데이터 저널리즘:

# 자료 더미에서 원석을 캐 빛나는 보석으로 가공

전통적으로 기자는 주로 글과 말로 먹고사는 사람들이었다. 여기에 사진기자는 사진 이미지, 방송기자라면 영상 감각이 중요하지만, 여전히 텍스트와 이미지는 두 축을 크게 벗어나진 않는다. 그들이 전달하는 정보 대부분은 사람을 취재해 나온다. 하지만 몇 년 전부터 기자들의 뉴스제작 문법과 관행을 흔드는 수상한 개념이 출현했다. 바로 데이터 저널리즘이 그것이다.

데이터 저널리즘의 가장 큰 특징은 지도나 그래프 같은 시각화 방식과 외형적인 측면보다는 뉴스에 접근하고 취재해 가공하는 전체적인 과정에서 찾을 수 있다. 한마디로 뉴스메이커, 즉 사람들을 직접 취재하는 대신 방대하고 복잡한 데이터를 우선하여 다룬다. 이를 가리켜 "데이터를 인터뷰한다"고 표현하기도 한다. 당연히 뉴스제작의 처음부터 마무리까지 데이터가 큰 비중을 차지한다. 기존의 보도 방식에서 통계와 데이터가 보조적인 위치에서 머물렀다면, 데이터 저널리즘은 데이터에서 뉴스가 나온다고 해도 과언은 아니다.

구체적인 사례를 들어보자. 자전거로 출퇴근하거나 여가에 운동 삼아 자전거를 즐겨 타는 사람들이 늘어나면서, 관련 교통사고도 증가했다. 한두 건의 피해 사례를 먼저 취재한 다음 경찰청 집계 수치를 섞어 제도적인 미비점을 지적하는 전문가 인터뷰를 함께 편집하는 게 전형적인 리포트 제작법일 것이다. 반면에 데이터저널리즘팀은 데이터부터 따져봤다. 교통안전공단의 자전거 사고 다발 지점 데이터를 웹크롤링 기법으로 수집하면, 지난 3년 동안 2번 이상 사고가 난 지점 8,500곳을 추릴 수 있었다. 이 지점을 지도로 시각화한 뒤, 전국의 교차로 위치와 중첩해보니, 60%를 넘는 자전거 사고가 자전거 도로와 차도가 만나는 교차로 부근에서 발생하는 사실을 확인할 수 있었다. 또한, 매년 6월과 9월 중순에 자전거 사고가 급증하며, 평일 출퇴근 시간에 사고가 몰렸다. 공간과 시간의 관점에서 추이를 분석해보니, 시내에서 자전거로 출퇴근하는 사람들이 겪는 위험이 고스란히 드러난 것이다. 결과는 별도의 생생한 현장 취재물과 함께 가상 스튜디오 리포트로 방송됐다.

정부는 전국의 공공시설의 장애인 편의시설 설치 현황을 5년마다 조사해 발표한다. 휠체어를 타고 입구에 들어가기에 불편함은 없는지, 점자블록은 제대로 갖춰져 있는지 등 다양한 사항을 점검한다. 2013년 자료를 보면, 공공시설의 장애인 접근성은 62.3%가 적절한 수준인 것으로 평가됐다. 데이터저널리즘팀이 최근 자료로 분석해보니 실상은 달랐다. 서울시 의뢰로 지체장애인 서울지원센터가 서울 소재 1,800여 개 공공건물의 접근성을 조사한 데이터에서는 전체 시설의 68%가 중대한 문제를 안고 있었다. 열쇠는 자료보다는 분석 방법의 차이에 있었다. 정부 보고서가 각 조사 항목의 점수를 평균을 내



어 집계했다면, 취재진은 중요 항목 중에서 '한 개라도' 문제가 있는 건물을 모두 골라냈다. 현관 입구가 좁건, 경사로가 없건, 한 가지라도 미비하면, 장애인이 휠체어를 타고 들어가는 데 문제가 발생할 수 있다는 점을 주목한 것이다. 보도자료를 넘어서 시민의 입장에서 직접 데이터를 따져 보면 보다 차별화된 기사를 쓸 수 있다.

데이터를 활용하면 기존의 취재 방식보다 사회 이슈에 더 입체적으로 다가갈 수 있다. 마치 하늘에서 큰 그림을 조망한 뒤, 개별 지점으로 들어가 확대경을 들이대듯 정밀하게 다시 들여다볼 수 있는 것이다. 특히 인터랙티브 차트와 지도를 통해서 시청자나 독자가 직접 자세한 내용을 탐색할 수 있도록 돕는다.

물론 데이터저널리즘팀을 괴롭히는 현실적인 제약과 오해도 존재한다. 무엇보다 언론이 활용할 수 있는 양질의 공공데이터가 턱없이 부족하다. 데이터 저널리즘을 정보산업계의 빅데이터 분석과 동일시하는 것도 흔한 착각이다. 실제로 필자가 다루는 대부분의 데이터는 엄밀한 의미의 빅데이터와는 한참 거리가 멀다. 오히려 스몰데이터라고 표현하는 것이 맞다. 데이터베이스이건, 문자를 담은 비정형 데이터이건, 컴퓨터를 이용해 자료 더미에서 원석을 캐내고, 그것을 다시 빛나는 보석으로 가공해 공익적인 뉴스로 전달하는 것이 데이터 저널리즘의 영역이다. 아직은 극소수 언론인이 외로운 실험을 진행 중이지만, 앞으로 어떤 모습으로 진화해갈지 언론계의 더 많은 관심이 절실하다.