디지털 사회
DIGITAL SOCIETY
제8호

연세대학교 사회과학데이터혁신연구센터
SOCIAL SCIENCE DATA INNOVATION CENTER

발 행 인: 조화순
발 행 일: 2018년 1월 1일
홈페이지: http://democracy3.0.yonsei.ac.kr
ISSN 2586-3525(Online)

데이터 큐레이션과 사회과학의 미래

김연숙
서울대학교 한국정치연구소 연구원

뉴 밀레니엄을 지나고도 20년 가까이 흘렀다. 20세기 말에 새로운 천 년을 준비하며 막연한 두려움과 희망을 가졌던 기억이 난다. 2018년 새해가 밝아오고 우리는 또 새로운 한 해에 대한 희망과 걱정 속에 미래를 설계한다. 새로운 세기를 맞이한 후 거의 20년이 흐르는 동안, 우리는 사회 각 부분에서 엄청난 변화와 급격한 혁신의 물결을 경험하였다. 뉴 밀레니엄의 변화와 혁신의 물결은 사회과학자들에게 좌절과 회의감을 주는 생소한 변화였을 수 있다. 과거 인기 있던 학문분야라 할 수 있는 정치학, 사회학, 행정학 등의 사회과학이 자연과학이나 공학에 비해 실용성이 떨어지는 학문으로 치부되고, 사회과학의 존립까지 위협받는 위기에 직면하게 되었기 때문이다. 사회과학이 현실과 괴리된 사변적 학문분야로 실용성이 떨어지고, 비생산적인 논쟁만 일삼는 무기력한 학문으로 전락한 느낌이다. 우리가 살아가는 삶 속에 호흡하고, 우리의 현실을 잘 설명하고 예측할 수 있는 사회과학은 불가능한 것인가?

우리는 매일 셀 수 없이 많은 양의 데이터가 생산되고 있는 환경 속에 살아가고 있다. 엄청난 양의 새로운 데이터는 우리의 삶, 우리의 사회, 경제뿐 아니라 정치적 환경을 설명하는 지표로서 사용된다. 또한 그것에 따라 우리의 삶과 정치, 사회의 작동 원리가 정의되고, 데이터가 알려주는 방향성에 따라 우리가 나아갈 경로가 새롭게 정해진다고 볼 수 있다. 개인적 차원에서도 데이터의 생산, 관리, 공유의 과정은 매우 중요해지고 있다. 사회 속에 살아가는 정치적 동물인 우리는 데이터의 홍수 속에 자신에게 필요한 데이터를 스스로 정의·수집하고, 집적된 데이터를 기준으로 가치관을 정립해 간다. 담론적 수준에서의 사회과학은 이제 주위의 데이터를 어떻게 모으고 효율적으로 관리·서비스하는가에 관한 '데이터 큐레이션(data curation)'의 단계로 진화하고 있다.

'데이터 큐레이션'이란 다양한 형태로 존재하는 데이터를 특정한 관심과 목적에 따라 새롭게 구성하여 전시(서비스)하는 행위를 뜻한다. 이를 통해 정보 탐색시간을 줄일 수 있고, 필요에 따라 선택적인 정보만을 공유할 수 있다. 이러한 시스템을 구축하고 설계하는 전문가를 '데이터 큐레이터'라 한다. 일상생활 속에서 소셜 네트워크를 통해 공유되고 타인과 소통하는 도구로 사용되는 데이터는 학술적 영역에서도 그 중요성이 높아지고 있다. 다양한 형태의 데이터가 어떠한 방식으로 가공되어야 사용자의 접근성을 높일 수 있는지, 연구자가 필요로 하는 정보형태로 제공될 수 있을지 고민하는 것 이 '큐레이터'의 주요 역할이다. '데이터 큐레이션'의

개념은 미래 사회과학이 다루어야 할 새롭고 혁신적인 분야의 핵심요소라 할 수 있다. 과거, 현재 그리고 미래의 사회, 정치, 경제현상을 설명하는 데이터 과학은 '아카이브(data archive)'의 대중화로써 사회과학 분야에서도 확대되고 있다.

컴퓨터 엔지니어만이 슈퍼 컴퓨터가 있는 랩(lab)실에서 데이터를 조작하고, 특정한 소프트웨어 지식이 있는 전문가만이 데이터베이스를 관리하고 공유할 수 있던 시절이 있었다. 하지만 이제는 개인들이 가진 다양한 형태의 디바이스를 통해 데이터를 직접 생산하고 서로 공유할 수 있는 시대이다. 특히, 정치학이나 사회학 등을 전공한 학술연구자들은 이제 간단한 소프트웨어 지식만으로도 우리 사회의 특성과 문화, 집단의 가치, 정치적 정향성, 투표결과의 예측 등이 가능해 졌다.

현재의 사회과학은 더이상 저명한 학자의 이론이나 거창한 담론적 논의로만 설명되지 않는다. 미래에는 나만의 '프레임'으로 다양한 형태의 데이터를 재구성하여 공유하는 패러다임을 수용하고, 그것을 어떻게 적용할 수 있을지에 달려있다고 해도 과언이 아니다. 또한 다양한 데이터를 통해 정치·사회현상의 인과관계를 탐구하고, 실용적 목적에 부응할 수 있는 융합적 시각으로의 전환이 필요할 때이다. '데이터 큐레이션'을 통해, 우리의 현실을 잘 설명할 수 있는 사회과학 연구가 가능해질 것이라 예상해 볼 수 있다. 미래 공상과학 소설에서나 나올법한 2020년이 먼 미래가 아닌 바로 2년 후이다. 우리가 속한 정치, 사회, 경제 환경을 과학적으로 해석하고 예측할 수 있는 사회과학적 해법을 고민해야 할 것이다. 이것은 데이터에 대한 지식과 분석기법의 이해, 데이터 공유에 대한 긍정적인 인식, 그리고 혁신적 큐레이션 기법에 대한 관심을 통해 가능해지리라 기대해 본다.