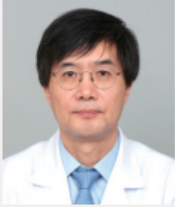


빅데이터에 대한 단상(斷想)



김종재 아산생명과학연구원장 / 교수·병리과

최근 주위에서 가장 많이 듣게 되는 단어가 '빅데이터'다. 이태원 거리에서 쉽게 볼 수 있는 빅사이즈 옷이나 신발들은 시각적 느낌이 확실하고 정서적으로도 익숙해서 그것들을 입고 신는 사람들까지 쉽게 연상이 되는 반면 빅데이터는 아직도 어떤 구체적인 느낌으로 다가오지 않는다. 더 당혹스러운 것은 빅데이터, 인공지능이 주요 국가연구사업의 중요한 화두가 되어감에 따라 빅데이터에 대해서 잘 아는 사람인 것처럼 행동하고 말을 해야 하는 경우가 늘고 있다는 사실이다.

빅데이터의 정의는 여기저기 찾아보아도 양과 속도, 다양성과 같은 속성에 대한 정성적인 기술이 주라서 결국 나 나름대로는 30년 전의 내 생활과 지금의 내 생활의 차이로 정의하기로 하였다. 80년대 학생, 전공의 시절에는 내가 어디에 가서 누구와 무엇을 했는지 이야기해주기 전에는 알 수 있는 사람들이 없었고 내가 따로 적기 전에는 데이터화할 수도 없었지만, 지금은 휴대폰 하나로 인해 내 생활의 모든 부분이 잠자는 동안마저도 데이터화되기 때문이다. 내 경우를 우리 병원 상황에 대입해서 내원하는 환자들, 직원의 수와 결부시키면 실로 상상하기 어려운 양의 다양한 데이터들이 지금 이 시간에도 생성 중이다. 앞으로 데이터의 양이 더 증가하면 '비거(bigger) 데이터' '비기스트(biggest) 데이터'라 할 것은 아니기 때문에 빅데이터라는 것은 오히려 경험하지 못했던 수준으로 데이터의 양과 종류가 폭발적으로 늘어나는 사회적 현상으로 보는 것이 더 타당하겠다.

그렇다면 왜 많은 사람들이 빅데이터의 중요성을 역설하면서 빅데이터와 인공지능이 보여주기 시작한 가능성에 열광할까? 데이터의 본질은 기록이다. 우리 모두가 역사를 공부하는 이유는 과거를 통해 현재를 개선하고 미래를 예측하기 위한 것이다. 결국 그것이 문자화된 것이든 아니든 역사라는 것은 과거 데이터의 결정체이기 때문이다. 그런데 사실은 하나일 텐데 이에 대한 해석은 다양해서 여러 사료(史料)들을 맞추어보면 특히 과거 영토에

관한 부분에서 많은 차이를 보인다. 이것은 데이터를 기록하거나 해석한 사람들의 오류와 편견이 그 주된 이유다. 같은 시대에 일정 지역을 지배했었다는 나라들이 복수인 이유이기도 하며, 이는 데이터의 정확성이 얼마나 중요한지를 일깨워주는 부분이다.

서울아산병원은 자타가 공인하는 우리나라 최고의 의료기관이다. 따라서 의료 빅데이터 분야에서도 자부심과 책임감을 가지고 양질의 데이터를 확보하고, 그로부터 중요한 의미와 새로운 사실들을 찾아내는 기술들을 적용해 건강한 사회를 이루는 데 기여해야 한다.

이러한 역할을 위해 우리가 쌓아가는 데이터는 '빅(Big)' 뿐만 아니라 '굿(Good)'이어야 하는 것이다. 최근 대형 국가연구개발사업인 인공지능의료 영상사업단, 한국데이터중심의료컨소시엄 등도 원내에 자리잡기 시작하여 빅데이터·인공지능 분야에서 우리 병원의 역량을 증명하고 있는 가운데, 시의적절하게 오는 6월 15일 열리는 아산국제의학심포지엄(2018 ASAN International Medical Symposium: Future Medicine in Digital Healthcare Innovation)은 현재와 미래 의료 분야에서 빅데이터와 인공지능이 어떠한 역할을 할 것인지, 또 해결해야 할 문제들은 무엇인지를 각계 의 전문가들을 초빙하여 알기 쉬우면서도 깊이 있게 다룬다. 해당 주제들의 기술적 측면뿐 아니라 제도적인 문제점들도 살펴볼게 될 것이다. 여러 직원들이 관심 있는 세션에 참석해 유익하고 새로운 정보들을 얻는 계기가 되었으면 한다.

의료 빅데이터에 관심 있는 직원들이 많아질수록 스피버그 감독의 영화 '마이너리티 리포트'에서 미래의 살인범을 예측한다는 예언자들의 틀로 시민들을 옥죄었던 프리크라임(PreCrime) 시스템이 아니라, 양질의 데이터와 과학에 기반하여 국민 건강을 지킬 수 있는 서울아산병원발 질병 예측 프리디지즈(PreDisease) 시스템의 적용이 현실이 되는 날이 가까워질 것이다.