

소셜 네트워크 데이터란 소셜 네트워크로부터 수집할 수 있는 정보로, 그 방대한 양 때문에 다양한 사회 현상을 예측할 수 있다. 흔히 ‘빅데이터’라고도 하며, 이는 다양한 방면에서 유용하게 쓰이고 있다. 휴대전화의 기록 데이터를 기반으로 사용자의 위치를 예측하는 알고리즘을 고안할 수 있고, 집단에서 어떻게 사고방식이 바뀌어 가는가와 같은 사람들의 행동 양식도 이 소셜 네트워크 데이터를 통해 예상할 수 있다.¹⁾

소셜 네트워크 데이터를 이용하면 사회 구성원들의 대인관계를 도식화하여 정리할 수도 있을 것이다. 스마트폰의 각종 앱들은 종종 연락처를 요구한다. 연락처 관련 앱이 아니더라도 ‘네이버 지도’ 앱의 권한에는 ‘연락처 수정’이 있으며, ‘Google 드라이브’ 앱에도 연락처를 확인할 수 있는 권한이 있다. 연락처뿐만 아니라 연락하는 내용도 수집할 수 있는 채널도 있다. 바로 소셜 네트워크 서비스, 이른바 SNS이다. 특히 페이스북 같은 경우, 얼마나 자주 친구들과 메시지를 주고받았는지, 서로 태그를 했는지에 대한 정보에 쉽게 접근할 수 있다. 즉, 서로 아는 사람인지 확인하는 것에 그치지 않고 친밀한 정도까지 파악할 수 있다. 또한 2013년 3월 기준으로 4천만 명 가량이 사용하고 있는 카카오톡은 거의 모든 스마트폰 사용자들이 연락을 주고받을 때 쓰는 수단으로써 페이스북과 같은 기능을 할 수 있다.²⁾

이런 소셜 네트워크를 기반으로 얻은 대인관계 데이터를 이용하면 학교폭력이나 따돌림을 예측할 수 있다. 10대들의 학교폭력은 학교 안에서만 일어나는 것이 아니라 학교 밖으로 연장되는 경우가 많다. 실제로 2014년 설문조사 자료를 보면, 학교폭력 피해 경험이 있는 청소년들 중 약 10%가 ‘사이버 괴롭힘’을 경험해본 적이 있다고 답했으며, 여학생의 경우 이 수치는 17%로 상승한다.³⁾ 이것이 전수조사도 아닐뿐더러 사실대로 답하지 않을 가능성이 있는 설문조사라는 점, 그리고 금품갈취 등 다른 유형의 학교폭력과 사이버 괴롭힘이 결합될 가능성이 있다는 점 등으로 미루어보았을 때, 실질 수치는 이보다 높을 수 있다.

그렇다면 소셜 네트워크를 이용하여 어떻게 학교폭력을 예측할 수 있을까? 먼저 소셜 네트워크 데이터를 수집하여 모욕적인 단어를 필터링하여 미리 청소년들의 잘못된 행동을 예측할 수 있다. ‘카톡감옥’이란 용어가 있다. 카카오톡 대화방에 계속 초대하여 대화방에서 나가지 못하게 하는 것이다. 초대를 거절하는 기능이 없다는 점을 악용한 것이다. 따돌림 당하는 아이들은 이 ‘카톡감옥’에 갇혀 수많은 또래에게 인신공격에 가까운 욕을 강제로 보아야만 한다.⁴⁾ 대화방을 나가더라도 다시 초대되기 때문이다. 이 경우에는 단어 필터링으로만 손쉽게 학교폭력을 예측할 수 있다.

그러나 종종 모욕적인 단어가 누구를 향한 것인지 알 수 없어 예측하기 어려울 때도 있다. 페이스북에서는 자신의 타임라인(일종의 전용게시판)에 상대방에 대한 모욕적인 비난을 누구나 손쉽게 올릴 수 있다. 자신의 ‘친구’ 혹은 ‘친구가 아닌 사람’도 볼 수 있는 공개적인 공간이기 때문에 비난받는 사람은 변명의 여지도 없이 많은 사람들의 비난을 견뎌내야 한다. 상대

1) 「Dr Seldon, I presume」, The Economist, 2013. 02. 23., 2015. 05. 30. <http://www.economist.com/news/science-and-technology/21572159-data-social-networks-a-re-making-social-science-more-scientific-dr-seldon-i>

2) 이윤희, 「국내 SNS 이용 현황과 주요 이슈 분석」, 『INTERNET & SECURITY FOCUS』, 2014년 8월 호, 한국인터넷진흥원, 60p.

3) 「2014년 2차 학교폭력 실태조사 분석결과 발표」, 한국교육개발원, 3p.

4) 한솔, 「출구없는 ‘카톡 감옥’ 이제 없어지나?」, 뉴스원, 2014. 12. 08, 2015, 05, 30. <http://news1.kr/articles/?1991462>

를 명시하기도 하지만 소위 '저격글'은 상대를 암시할 뿐 직접적으로 언급하지 않는 방식을 이용하여 비난을 받는 상대의 반박을 피해간다. 그러나 소셜 네트워크는 작성자의 대인관계와 다른 친구들과의 대화내용을 모두 파악할 수 있기 때문에 이렇게 대상이 명시되지 않은 비난 글의 대상도 쉽게 파악 가능하다.

그러나 소셜 네트워크를 기반으로 하는 데이터 수집에는 다양한 위험성이 존재한다. 첫째, 다양한 문맥 속에서 이 데이터의 의미가 잘못 수용되어 배치될 수 있다.⁵⁾ 예를 들어 모욕적인 언사를 기준으로 필터링하기 때문에 스스로를 비관한 글이 데이터 수집 과정에서는 타인에 대한 비방글로 변질될 수 있다. 둘째, 잘못된 정보가 수집됐을 때 사용자들은 이를 수정할 권리도 없을뿐더러 잘못된 정보가 수집되었는가의 여부를 확인하기도 쉽지 않다. 넘쳐나는 정보 속에서 자신과 관련된 정보를 찾는 것은 사막 위에서 바늘을 찾는 것과 다르지 않기 때문에 사용자 대부분은 이 사실을 모르고 있는 경우가 많다. 즉, 자신이 비방글을 써서 학교폭력 의심 대상자로서 정보가 수집되었다는 사실을 모른다. 정보가 수집되었다는 사실 자체를 인식하지 못하기 때문에 잘못된 정보를 수정할 권리도 제대로 행사할 수 없다. 세 번째, 설령 가해여부 사실이라 하더라도 단 하나의 데이터가 청소년 한 명에게 영원히 '가해자'라는 낙인을 찍을 수 있다. 학교폭력의 가해자와 피해자는 고정적인 게 아니라 끊임없이 변화 가능한 요인임에도 불구하고⁶⁾, 이렇게 한 번의 데이터로 한 청소년을 가해자 후보에 올려놓는 것은 지나치게 가혹할 수 있다. 이는 모욕적인 단어라는 데이터가 깊이 있는 질적 내용이 아니라 단순 사실 여부만을 전달하기에 일어나는 위험이다.

위험성이 다양한 만큼 이에 대한 해결방안도 다양하다. 첫 번째 위험성 문제는 기본적으로 소셜 네트워크 데이터의 수집이 컴퓨터로 이루어지기 때문에 발생하는 문제다. 컴퓨터는 문맥을 읽을 수 없기 때문에 특정 글이 특정 단어를 포함하는가의 여부만을 가려낸다. 따라서 필터링 시스템을 더 엄밀하게 설계하여 비난의 대상을 정확히 파악하게 할 수 있다. 그러나 더 확실한 방법은 '인간계산 기술'을 이용하는 것이다. 인간만이 엄밀한 단어의 뉘앙스까지 알 수 있기 때문에, 인간과 컴퓨터를 결합하는 작업을 통해 이 문제를 해결할 수 있다. 둘째, 사용자가 언제든지 데이터의 출처를 트래킹할 수 있도록 법제화하는 것이다. 출처 트래킹을 통해 자신의 데이터가 누구에 의해 어떻게 수집되는지를 분명히 알면, 사람들은 자신의 잘못된 정보를 발견했을 때 손쉽게 수정가능하다. 또한 이러한 법 규제로 인해 컴퓨터 기술이 발전하면 데이터를 자동으로 업데이트하여 수정할 수 있는 소프트웨어까지도 가능해질 것이다.⁷⁾ 그러나 자신의 데이터가 수집되었는가를 자동으로 알 수 있는 방법은 조금 더 생각해보아야 할 문제다. 세 번째 문제는 어떤 질적 특성을 양적 데이터로 파악하려는 과정에서 발생하는 문제다. 이를 해결하려면 다층적인 정보를 활용해야 한다. 즉, 단순히 비방글 관련 데이터만으로 학교폭력 여부를 예측하면 안 된다. 이전에 비난의 대상 또한 비방글을 쓴 적이 있는지, 대상 간에 얼마나 연락을 자주 하는지, 두 대상이 동시에 아는 친구들도 비슷한 비방글을 썼는지 등도 함께 수집하여 이를 총체적으로 살펴보아야 한다. 소셜 네트워크 데이터의 힘은 엄청난 양의 정보를 정렬시킬 수 있다는 데에서 나오며, 소셜 네트워크는 이런 다층적인 정보를 수집하는 데 최적의 환경을 제공해준다.

5) Frank Pasquale, 『The Black Box Society』, Harvard University Press, 2015, 148p.

6) 「2014년 2차 학교폭력 실태조사 분석결과 발표」, 한국교육개발원, 4p.

7) Frank Pasquale, op. cit., 146p.