

Chapter 7. Social Network Analysis in Human Geography

발제자께서는 소셜 미디어 데이터를 활용한 사회 연결망 분석(Social Network Analysis, SNA)이 인문지리학에서 어떻게 수행되는지 네 가지 범주를 예로 설명해주었고 소셜 미디어 데이터가 가지는 한계와 지향해야 할 방향, 연구 윤리적 책임을 강조했다. 또한 소셜 미디어 데이터가 '현실'을 얼마나 반영하고 있는지, 데이터가 가지는 한계를 보완하기 위해 지향해야 할 점을 제시해줬다. 필자도 소셜 미디어 데이터 기반 분석은 전통적 인문지리학 방법을 대체하는 것은 어렵다 생각하고 보완하는 도구로 사용되어야 하고 그가 가지는 문제점을 해결해야 의미있고 강력한 방법론이 될 것이라고 생각한다.

토론거리

Q1) 소셜 미디어 데이터는 인간의 '실제 경험된 공간'을 얼마나 정확하게 반영한다고 볼 수 있는가? → 디지털 공간과 현실 공간 사이의 괴리를 어떻게 해석할 수 있을까?

A) 인문지리학에서 공간은 단순한 물리 좌표가 아니다. 몸으로 체험된 공간이기도 하고, 감정, 기억 등과 같은 의미가 부여된 공간이기도 하며 권력이나 관계 속에서 구성되는 공간이기도 하다. 소셜 미디어 데이터는 한편으로 그 공간을 경험하는 사람들이 남긴 발자취라고 할 수 있다. 특히 사람들의 동선이라든지, 공간에서 생산된 감정 등은 사람들이 자발적으로 생산한 공간 데이터라고 할 수 있다. 이러한 점을 봤을 때 어느 정도 실제 경험된 공간을 반영하고 있다고 할 수 있지만 정확하다고 할 수는 없다. 앞서 발제자께서 언급한 대표성 문제, 플랫폼 간의 차이 등으로 인해 사회적으로 편향되었다고 할 수 있다. 하지만 역설적으로 정확하게 사회집단의 모습을 반영하는 곳이라고도 할 수 있다. 본문에 언급된 편향된 인구 분포(ex. 20대 남성 위주의 커뮤니티 or 20대 여성 위주의 커뮤니티)를 가진 소셜 미디어의 데이터를 확인하면 그 사회집단이 경험하고 있는 감정이나 분위기를 정확히 파악할 수 있다고 생각한다.

Q2) 감성 분석을 통한 공중 인식 연구는 실제 사회적 태도를 얼마나 대표할 수 있는가? → 표본 편향과 알고리즘 편향 문제는 어떻게 해결할 수 있을까?

A) 앞서도 언급했듯이 데이터의 편향 문제로 인해 사회적 태도를 대표하기는 불가능에 가깝다. 그러나 이러한 문제는 AI를 활용해 해결 가능성을 제시할 수 있다고 생각한다. 예를 들어 감성 분석에서의 표본 편향의 문제는 앞선 장에서 확인했던 감정 측정 지표인 SAD(Sentiment Adjusted by Demographics)지수를 활용해 조정이 어느 정도 가능하다고 생각한다. 또한 각 회사 인사팀에서 활용하는 채용/인사 AI는 "Demographic Reweighting AI"를 활용한다. Demographic Reweighting AI는 인사 부서 학습 데이터에서 오는 성별 편향의 완화를 위해 가중치를 재조정한다. 이 기술은 표본 편향의 문제를 해결할 수 있는 한 가지 방법이 될 수 있다.

Q3) GeoAI 및 소셜 미디어 분석이 인문지리학을 '정량화된 학문'으로 변화시키는 것은 바람직한가? → 질적 연구와의 균형은 어떻게 유지할 수 있을까?

A) 바람직하지 않다고 생각한다. 질적 연구는 분명 양적 연구에서 해결할 수 없는 부분을 채워준다. 예를 들어 필자가 진행했었던 성매매 집결지 해체와 관련된 분석에서도 양적 데이터만으로 해당 건물의 성매매 업소 운영 여부를 전부 파악하지 못했다. 여기서 인터뷰 등과 같은 질적 연구 방법이 해결할 수 있는 실마리가 되어준다고 생각한다.

Chapter 8. Converging GeoAI and CyberGIS for Human Geography

발제자께서는 인문지리학이 맞이하고 있는 데이터의 범람과 연구 방법론적 변화, 사이버 인프라의 결합 등의 주요 안건을 사례를 들어 언급해주었고, 마지막으로 GeoAI와 CyberGIS의 융합이 데이터를 통해 인간과 환경의 본질적인 관계를 더 정밀하고 책임감 있게 성찰하는 연구로 나아가는 토대를 마련해준다고 강조했다. 필자도 특히 인프라의 중요성을 여실히 느끼고 있어 본문의 내용에 공감가는 부분이 많았다. 또한 글 말미에 언급한 수용자의 관점에서 비판적으로 생각해야 할 부분이 전제되어야 한다는 것은 필자도 부족한 부분이라 생각이 들어 의견을 나눠봤으면 한다.

토론거리

Q1) 지리적 데이터 주권과 알 권리: 내가 작성한 소셜 미디어 포스트나 스마트폰의 이동 경로 데이터가 공공의 이익을 위한 연구에 활용될 때, 개인의 알 권리와 이후에도 자신의 데이터가 활용되는 것을 거부할 수 있는 '지리적 자기결정권'을 어떻게 보장 받을 수 있는지?

A) 지리적 데이터, 특히 위치 데이터의 경우 익명화가 매우 어렵다. 개인 특정화가 쉽고 포함된 개인 정보가 많아 엄격히 관리되어야 하는 데이터라고 할 수 있다. 물론 제도적 수준에서 데이터 사용 고지의 의무 등이 있어 거부의 기회가 있지만 그렇지 않은 경우도 빈번하다. 따라서 데이터 수집의 단계와 사용자 입장의 단계에서 한 가지씩 보장받을 방법을 토론했고자 한다. 먼저 데이터 수집의 단계에서 "차등 프라이버시" 기술이 필요하다. 수집하는 위치 데이터에 노이즈를 추가해 개인 경로는 암호화하고 집단적 패턴의 유사성만 보이는 경로만 수집하는 등의 차등 정보 수집 기술이 우선되어야 할 것이다. 다음으로 사용자 입장에서 실시간 동의 시스템이 도입되어야 할 것이다. 사용자의 위치 데이터가 연구에 활용된다는 UI를 추가해 사용자가 실시간으로 동의/거부 등을 선택할 수 있도록 한다면 지리적 자기결정권을 조금이나마 보장받을 수 있을 것이라 생각한다.

Q2) 디지털 소외 계층의 가시성 확보: 소셜 미디어나 스마트 기기 사용이 활발하지 않은 고령층, 저소득층 등 디지털 사각지대에 놓인 소외 계층의 목소리는 AI 기반 빅데이터 분석 과정에서 배제되어 누락될 가능성이 매우 큼니다. 이러한 불평등을 보완할 수 있는 방안은 무엇인지?

A) 여러 연구에서 제시되는 문제점인 과소대표 현상은 중대한 사안이다. 우리가 접했던 문제점들 예를 들어, 도시/농촌 간의 격차, 지역 인프라 불평등, 교육과 소득의 격차 등 공간적 불평등 데이터의 대표적인 문제점들이다. 이런 문제를 해결할 수 있는 것이 질적 방법론이라고 생각한다. 심층 인터뷰, 참여관찰이나 오프라인 설문 등과 같은 방법론이 디지털 소외 계층의 의견을 어느 정도 반영해줄 것이라 생각한다.

Q3) AI 분석 결과의 수용성과 신뢰: GeoAI 가 도출한 고도의 분석 결과가 나의 거주 환경이나 건강권에 직접적인 영향을 미치는 정책으로 전환될 때, 그 결정과 선택을 일반 시민에게 설명하는 기술인 'XGeoAI(Geospatial Explainable AI)'의 구현은 어느 수준까지 이루어져야 하며, 이를 위해 전문가와 대중 사이의 지식 격차를 어떻게 해소할 수 있는지??

A) XAI는 현재 특정 분야에서 수요가 크게 늘고 있다. 예를 들어 금융권에서 대출 심사를 할 때 대출이 거부되는 이유를 AI가 설명해주기도 하고 투자 자문을 받을 때도 시장 상황과 고객의 위험 선호도를 판단해 설명해주기도 한다. 이와 마찬가지로 XGeoAI 역시 시민들이 납

득할 수준(지식의 격차를 고려한)까지 구현이 되어야 한다고 생각한다. 위험도를 시각화해서 보여주거나 비용과 효율의 상관관계 그래프를 제시하거나 하는 등 시민들이 납득할 수 있는 정보를 시각화해서 가시성 높게 도출되어야 한다고 생각한다. 추가로 시민들의 불안을 고려해 AI 예측의 신뢰 구간, 오차 분포와 개략적인 시나리오를 함께 제시한다면 조금 더 투명하고 신뢰할 수 있을 것이다.