

Chapter 22. Towards Ethical Spatial Decision-Making in Geospatial Artificial Intelligence

본문과 발제자께서 글 말미에 언급하신 토론거리에 대해 역사지리학적인 관점에서 나름의 생각을 적어봤습니다.

토론거리

Q1. GeoAI 모델, 특히 편향된 데이터, 알고리즘 또는 인간의 입력에서 발생하는 편향을 어떻게 식별하고 해결하여 더 정확한 해석과 의사결정 지원을 생성할 수 있는가?

A. 어느 학문이든 마찬가지겠지만, 특히 역사지리학 내에서 데이터 편향 문제는 민감하게 받아들여야 합니다. 역사 지리 정보는 자료의 특성상 오랜 기간 축적된 공간적 불균형과 권력의 시선을 내포하고 있습니다. 예를 들어 일제강점기 시대의 지도나 개발 중심의 도시계획 주제도 등은 특정 계층에 편향된 정보를 포함하고 있어 GeoAI의 활용에 주의를 요합니다. 따라서 역사지리학에서 GeoAI를 윤리적으로 활용하기 위해서는 역사적 맥락을 읽을 수 있도록 데이터를 구축해줘야 한다고 생각합니다. 데이터의 출처와 배경을 명시해 연구자가 윤리적 판단을 할 수 있도록 돕거나, 온톨로지 형식의 자료 구축과 시계열적 교차 검증 등으로 편향을 제거한다기보다 그 배경과 구조를 드러내는 것이 바람직하다고 생각합니다.

Q2. 사용자가 의사결정 과정을 이해하고 잠재적인 편향의 근원을 감지할 수 있도록 GeoAI 모델의 투명성을 개선하기 위해 어떤 전략을 사용할 수 있는가?

A. "토론거리1"에서 다룬 내용과 같이 편향의 배경과 구조를 드러내서 사용자가 역사적 맥락을 이해할 수 있게 하는 것이 바람직한 전략이라고 생각합니다.

Q3. 인간의 판단과 윤리적 고려 사항이 AI 기반 공간 결정에 통합되도록 보장하면서, GeoAI가 인간의 의사결정을 대체하는 것이 아니라 증강하기 위해 어떤 방식으로 사용될 수 있는가?

A. 현재 AI의 수준을 연구자가 판단하고 적용할 수 있는 범위까지만 활용한 후, 최종 의사결정을 인간이 내리는 방식이 옳다고 생각합니다.

Q4. 공간 의사결정에서 GeoAI의 윤리적 사용을 규제하기 위해 기존의 법적 프레임워크를 어떻게 조정하거나 확장할 수 있는가?

A. 현재 공간 의사결정의 범위 내에서 법은 개인정보나 알고리즘의 공정성 등에 초점을 맞추고 있습니다. 물론 현재 다루는 것들도 중요하지만, 역사지리학에서의 GeoAI는 추가적인 접근이 필요합니다. 역사 자료는 앞서 언급한 것처럼, 역사적으로 공간적 불평등을 내포하고 있었습니다. 그렇기 때문에 연구자들이 연구를 계획할 때 차별 데이터를 무지성, 무비판적으로 사용하는 것에 대한 제한이 추가적으로 필요하다고 생각합니다.

Q5. 취약 계층에 미치는 영향이나 환경 지속 가능성과 같은 윤리적 차원을 설명에 반영할 수 있도록 GeoAI 시스템에 윤리적 추론을 통합하는 것이 가능한가?

A. 알고리즘에 가중치를 부여하는 등과 같은 방식으로 윤리적 판단을 일부 맡길 수 있다고 생각합니다. 하지만 윤리적 판단 역시 역사적 맥락에 따라 어느 한 쪽에서는 긍정적, 다른 한 쪽에서는 부정적 결과로 이어질 수 있기 때문에 단순히 인간이 판단할 수 있는 시나리오를 제시해주는 정도의 통합을 이루는 것이 옳을 것 같습니다.

Chapter 23. Sustainable GeoAI in Human Geography: Reproducible, Replicable, and Expandable

인문지리학에서 GeoAI의 활용에 있어 RRE(재현성, 재현가능성, 확장성) 개념은 필수적인 조건이 되어가고 있습니다. 그러나 발제문에서 언급했다시피 인문지리학의 지리공간 데이터는 고차원적이고, 특수성이 짙어 이를 실현하기 어려운 상황이라고 생각합니다. 그에 대한 해결책으로 본문과 발제자께서는 표준 데이터셋과 오픈 소스 도구의 사용, 더 나아가 XAI의 도입을 제시하셨는데, 역사 지리 자료를 다루는 입장에서 봤을 때 와닿는 해결책이라고 생각되지 않습니다.

표준 데이터셋과 오픈 소스의 사용으로 연구의 재현과 확장에 이바지할 수 있음은 자명하다고 할 수 있습니다. 하지만 역사 지리 자료의 경우 복잡한 역사적 맥락이 얽혀있고, 인과 관계를 쉽게 설명하기 어려운 경우가 많고 앞서 언급된 해결책만으로는 문제가 해결되지 않습니다. 예를 들어 XAI의 경우 '설명 가능한' AI라고 불립니다. 이것은 모델 '내부'의 설명이지 문제의 근본을 다루는 설명이 아닙니다. XAI가 특정 지역을 낙후하다고 판단할 때, 소득과 인프라 등의 부족을 다루고 어떤 변수에 얼마만큼의 가중치가 설정이 되었는지 AI 내부의 모델에 관한 설명만을 해줍니다. 그러나 AI가 낙후하다고 판단한 지역의 역사적 맥락을 이해해보면 정치적 탄압지역이라든지, 식민지 개발 구조상 예외적인 지역이었다든지 하는 맥락을 이해해야 할 이유가 존재합니다. 더 나아가 이러한 AI의 허점이 악용될 경우 정당화의 도구로 전락할 수 있다고 생각합니다.

이러한 문제점에도 불구하고 연구에 있어 RRE는 실현되어야 하는 것은 변함없습니다. 따라서 AI는 문제를 해결하는 역할이 아니라 문제를 드러내는 도구로 바라보고 연구의 재현이 역사적, 공간적 불평등을 되풀이되지 않게 윤리적으로 활용되어야 할 것입니다.

토론거리

Q1) RRE(재현성·재현 가능성·확장성) 가운데 현재 GeoAI 연구·실무 현장에서 가장 취약하다고 느끼는 부분은 어느 지점인가?

A) 재현 가능성 부분이 어렵다고 느껴집니다. 단편적으로 봤을 때 각기 다른 현장에서 통일된 데이터 포맷을 유지하고 있지 않는 경우가 많고, 본인이 사용하기 좋은 데이터로 전처리, 가공하는 작업이 동반될 것이라 취약부분이 있다고 생각합니다.

Q2) 우리가 사용하는 GeoAI 도구와 데이터 환경에서, 재현 가능한 워크플로를 구축하는 데 가장 큰 실질적 장애물은 무엇인가?

아직 GeoAI 도구를 직접 경험해보지 못해 실질적 장애물을 판단하기 어렵지만, 개인적으로 생각해봤을 때 연구 환경의 차이가 장애물이 될 것 같다고 생각합니다. 물론 인프라, 편향, 프라이버시 등 제한이 많지만 연구 환경의 차이는 현장에서 발견할 수 있는 가장 많은 오류일 것이라고 생각합니다. 예를 들어 운영체계의 상이함으로 인해 파일시스템이나 경로가 꼬일 수 있고, 버전의 차이로 인한 충돌 등이 생길 수 있습니다. 이를 해결하기 위해 연계된 기관끼리 Docker¹⁾를 활용하기도 하지만 모든 연구자들이 친절한 것은 아니기 때문에 실질적으로 장애물이 될 수 있다고 생각합니다.

1) Docker: 코드, 라이브러리, 환경설정 등을 하나로 패키징해서 어느 환경에서 작업하더라도 동일하게 작동할 수 있도록 보장하는 컨테이너 기반의 오픈소스 도구.

Q3) 지역사회 참여와 윤리적 고려를 GeoAI 연구에 실제로 반영하기 위해, 데이터 처리와 모델 설계에서 어떤 원칙과 절차가 선행되어야 할까?

'공간 해석을 누가 하느냐'에 따른 주체가 정해져야 합니다. 많은 장에서 다뤘던 AI의 문제점들이 해결되지 않기 때문에 AI를 적용할 수 있는 부분은 활용하고 해석과 판단은 연구자가 할 수 있도록 원칙이 정해져야 합니다.

Chapter 24. The Knowledge Politics of GeoAI

발제자께서는 GeoAI 기반 디지털 인도주의가 내포한 데이터 식민주의와 지식 권력의 비대칭성을 비판적으로 바라보았습니다. 발제문에 제시된 '글로벌 북부'의 행위자가 제시한 윤리적 가이드 라인이 오히려 '글로벌 남부'의 행위자를 능동적 '데이터 시민(data citizen)'이 아닌 '보호의 대상'임을 보여주는 것도 공감되는 부분이 있습니다. 하지만 이 부분에서 조금 더 미시적으로 바라봐야 할 필요가 있다고 생각합니다.

발제문에서 제시된 지도의 대상이 되는 지역 주민이 데이터 생산 과정의 주권을 잡는 것은 글로벌 남부 사회 내에서도 주의해야 할 점이 존재합니다. 발제자께서도 언급하셨지만, 글로벌 남부가 데이터 생산 과정의 주권을 잡는다는 것은 '주민들의 체화된 공간 지식과 목소리가 알고리즘의 설계 기준 자체를 그 기초부터 재구성하고, 궁극적으로 생산된 공간 데이터에 대한 소유권과 통제권, 즉 데이터 주권(data sovereignty)을 커뮤니티가 온전히 회복하는 것'이 궁극적인 결과가 되어야 합니다. 여기서 주민들은 동일한 지역에서도 계급, 성별, 권력 등에 따라 공간 인식이 전혀 다르기 때문에 대표성이 흐려질 수밖에 없습니다. 그렇게 된다면 주민을 대표한다는 것이 모호해져, "주민 참여"라는 대의 아래 또다시 데이터 권력의 독점이 이루어질 가능성이 충분합니다. 이는 주민의 자발적 참여가 실제로는 선택된 참여가 될 수도 있고, 결과적으로 다른 형태의 데이터 식민주의가 야기될 수 있습니다.

따라서 이를 해결하기 위해 주민을 하나의 주체로 보지 않고 다층적인 존재로 구분한 다음 그들 내에서도 서로 다른 행위자들이 각각의 의견을 낼 수 있는 방향으로 주민들의 참여가 이루어져야 할 것입니다.