

인간 대상 연구를 위한 데이터에 대한 관계적 관점과 지각의 인지적 침투 개념의 확장*

오 경 준[†]

이 글은 인간 대상 연구에 의해 생성되는 인간 유래 데이터의 특징을 명료하게 만들고 과학적 실천에서 연구 주체로서 인간 대상자의 의의를 재검토한다. 이를 위해 먼저 나는 과학적 관찰에 대한 전통적 의견을 비판하며 제시된 데이터에 대한 관계적 관점을 소개하고, 그 의의와 한계를 설명한다. 데이터에 대한 관계적 관점은 데이터 중심 과학으로서 데이터와 메타데이터의 적극적인 활용을 통해 연구를 위한 데이터베이스를 구축, 사용하는 데이터 중심 생물학의 특징을 규명하고 이를 발전시키는 데에는 유용하지만, 데이터의 전파와 공유를 위한 탈맥락화 과정에서 인간 대상자와 상호 작용에서 생성되는 인간 유래 데이터의 고유한 특징을 반영하지 못 한다. 나는 인간 유래 데이터의 고유한 특징을 반영하여 과학적 실천에서 새롭게 고려해야 할 연구 주체로서 인간 대상자에 대한 새로운 관점을 위해 데이터에 대한 기존의 관계적 관점과 지각의 인지적 침투 개념의 적용 대상을 확장하는 것을 시도한다.

주요어: 데이터에 대한 관계적 관점, 지각의 인지적 침투, 탈맥락화, 정보 유기체, 데이터 중심 과학

* 1. 이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.RS-2025-02217259, 뇌인지 다학제 국제 거버넌스 기반 인공지능 편향성 검출-교정-설명가능 지능적 자율진화 플랫폼 개발, 50%)과 2025년도 위즈진 주식회사의 재원 및 지원(50%)을 받아 수행된 연구임.

2. 논문의 전반적인 구조와 내용을 고도화하여 논문의 완성도를 높이는 데에 도움을 주신 심사자들에게 깊은 감사를 표한다.

[†] 한양대학교 과학기술윤리법정책센터, 연구원 (silvercord22@gmail.com)

1. 들어가는 글

과학철학의 오랜 역사에서 과학적 지식이 구축되고 정당화되는 방법을 명료하게 이해하기 위해 실험과 관찰을 통해 만들어진 결과물로서 데이터의 본성을 해석하는 일은 많은 논쟁을 일으켜왔다. 특히, 데이터와 현상, 데이터와 데이터 모형, 현상과 이론의 관계를 논의하기 위한 전제를 제공한다는 측면에서, 과학자가 가져야 할 데이터에 대한 관점을 제시하는 것은 과학적 실천에서 데이터를 이해하는 데에 있어 가장 중요한 작업이라 평가할 수 있을 것이다(Suppes 1962; Bogen and Woodward 1988).

제임스 보겐과 제임스 우드워드는 이론과 관찰을 이분법적으로 나누는 관찰에 대한 고전적인 관점으로부터 이론과 현상 사이의 관계에 대해 다음과 같은 설명을 제공했다. 그들에 따르면, 데이터는 과학적 실천에서 직접 관찰한 것에 대한 기록으로서, 측정과 실험을 통해 생산된다. 그들은 반면 현상은 일반적으로 관찰 불가능하며, 다양한 측정과 실험을 통해 관찰 가능한 데이터로부터 추론되는 것이라 설명한다. 더불어 이론은 이러한 현상을 예측하는 것으로서, 데이터는 현상에 대한 주장의 증거를 제공하는 대상으로 여겨진다(Bogen and Woodward 1988). 사비나 레오넬리는 제임스 보겐과 제임스 우드워드가 제시한 이론, 현상, 데이터에 관한 관점을 데이터를 세계의 표상으로 해석한다는 측면을 강조하여 ‘데이터에 대한 표상적 관점’이라 부른다. 그는 현상에 대한 주장의 증거로서 객관적인 정체성을 갖는 데이터 개념 대신, 데이터의 정체성과 중요성이 연구 과정에서 사용되는 데이터의 기능과 맥락으로부터 파생된다는 측면에서 그 증거적 지위가 연구 과정에서 과학자와 비과학자의 능동적인 사용과 해석에 의해 형성되는 관계에 따라 달라지는 데이터 개념, 즉 ‘데이터에 대한 관계적 관점’을 제안한다.¹⁾

1) Leonelli (2016), pp. 65-6.

이 글의 목표는 우리로 하여금 과학적 관찰의 의미를 새롭게 해석하고 과학 활동에서 과학자와 비과학자의 새로운 역할에 주목하게 만드는 ‘데이터에 대한 관계적 관점’을 인간 대상 연구에 적용했을 때 발생하는 문제점을 비판적으로 검토하고, 이러한 문제점을 해소하는 한편, 데이터 중심 과학의 안전하고 합리적인 발전을 위한 ‘데이터에 대한 확장된 관계적 관점’을 제시하는 것이다.

이를 위해 본 논문의 논의는 다음과 같이 구성될 것이다. 2절에서는 사비나 레오넬리가 제안한 ‘데이터에 대한 관계적 관점’의 내용과 그 의의를 간략히 살펴볼 것이다. 3절에서는 ‘데이터에 대한 관계적 관점’을 과학 활동으로서 인간 대상 연구에 적용했을 때 발생하는 문제점을 중심으로 그것의 한계에 대해 논의할 것이다. 4절에서는 ‘데이터에 대한 관계적 관점’을 인간 대상 연구에 적용했을 때 발생할 문제를 해소하고, 데이터가 과학자, 연구 대상자, 비과학자 등의 다양한 주체들과 갖는 상호작용을 다룰 수 있는 ‘데이터에 대한 확장된 관계적 관점’을 제시할 것이다. 이 과정에서 과학자와 관찰 대상 사이에서 발생하는 현상, 혹은 특징으로 이해되었던 지각의 인지적 침투가 관찰 대상인 연구 대상자와 연구 대상자가 상호 작용 중인 세계 사이에서도 일어난다는 측면에서 확장된다는 것 또한 설명할 것이다. 5절에서는 제시한 ‘데이터에 대한 확장된 관계적 관점’의 의의를 설명하고 그 한계를 파악하여 더 합리적인 후속 연구를 위해 필요한 요소를 확인할 것이다.

2. 과학적 관찰의 재정의로서 데이터에 대한 관계적 관점의 의의

사비나 레오넬리가 ‘데이터에 대한 표상적 관점’으로 규정한 제임스 보겐과 제임스 우드워드의 설명에 따르면, 과학자가 과학 활동의 일환으로서 관찰을 통해 “획득한” 원시 데이터는 “주어진 것”으로서 객관적이고 고정된 인식론적 지위를 갖는다. 반면, 그가 대안으로 제안한 ‘데이터에 대한 관계적 관점’은 과학자가 자신의 연구 목적, 채택한 연

구 방법, 연구 수행에 수반된 다양한 실천적 요소들의 영향을 받아 데이터의 내용을 결정한다는 측면에서 마음 의존적이며²⁾, 연구 수행 과정의 직접적인 주체가 아닌 비과학자 역시 과학자가 생산한 데이터를 사회 혹은 다른 과학자에게 보급하는 관행을 규제, 관리하는 집단을 형성한다는 측면에서 데이터를 공동 생산하고 데이터의 내용을 결정하는 주체일 수 있다는 점을 설명한다.³⁾ 이에 따르면, “주어진 것”으로서 객관적이고 고정된 원시 데이터의 인식론적 지위는 사라지고, 과학자, 측정 혹은 실험에 사용되는 도구, 활용되는 이론적 틀, 데이터와 데이터의 보급 과정을 형성하는 사회적 맥락 사이의 복잡한 상호 작용의 인식론적 지위가 강조된다.

다양한 요소들 사이의 복잡한 상호 작용에 의해 형성된 데이터는 모형화를 통해 증거적 가치를 부여 받는다. 주어진 데이터 집합에 대한 가장 적합한 해석을 선택한다는 측면의 데이터 모형화는 분석에 사용될 통계적 원리, 데이터의 특색, 데이터 정렬 방식과 시각화 방식, 데이터 모형화를 수행하는 과학자의 관심사와 전문성에 따라 달라진다. 더불어, 그는 과학자들이 데이터의 증거적 가치를 높이기 위해 데이터에 대한 표상적 관점에 입각하여 데이터를 객관적이고 맥락 독립적 단위로 개념화할 필요 없으며, 과학자의 주관이 개입된 데이터의 맥락 의존적 특징을 명시적으로 밝히는 데이터 패키징을 통해 오히려 데이터의 이동성과 증거적 가치를 높일 수 있다고 주장한다.⁴⁾

김진영은 데이터에 대한 표상적 관점에 대한 사비나 레오넬리의 비판에도 불구하고, 데이터에 대한 그의 관계적 관점이 데이터에 대한 표상적 관점과 상호 배타적인 것이 아니라 상호 보완적으로 이해되어야 한다고 주장한다(Gim 2023). 그는 데이터의 증거적 가치가 레오넬리의 설명과 달리, 데이터가 패키징 되고 이동되기 전에도 여전히 세계의 표상적 원천에 의존한다고 주장한다. 그는 사비나 레오넬리가 데이터 패키징, 데이터 큐레이션, 데이터의 순환 및 재사용과 같은 사회적 절차

2) *Ibid.*, p. 74.

3) *Ibid.*, p. 54.

4) *Ibid.*, pp. 91-2.

의 중요함을 강조하지만, 데이터가 갖는 인식적 가치 자체를 부정하지 않았다는 측면에서 여전히 데이터에 대한 표상적 관점과 양립 가능하다고 주장한다. 그에 따르면, 현상의 구조적 특징들에 대한 데이터의 내용은 자연에 의해 결정되며, 현상의 물리적 형태에 대한 데이터의 내용은 연구자의 관심사에 의해 결정된다.

데이터에 대한 표상적 관점과의 공존 가능성과 같은 비판에도 불구하고, 사비나 레오넬리가 제시하는 데이터에 대한 관계적 관점은 여전히 과학적 실천, 특히 과학적 관찰이란 개념을 새롭게 정의하는 데에 크게 기여한다. 전통적으로 여러 과학 분야에서 관찰은 이론을 평가하는 순수하고 이론 중립적인 객관적 기반으로서 여겨지고 있으며, 과학 철학 분야의 연구를 통해 관찰의 이론적재성과 같은 개념이 제시된 이후에도 이를 극복하거나 연구자의 주관성과 관계된 요소를 줄여야 한다는 규범적 생각이 자리 잡고 있다. 인지적으로 침투된 연구자의 관찰 경험이 그 경험의 내용에 해당하는 지각 믿음을 정당화할 수 없지만, 관찰 믿음에 관한 이론적재성을 수용하면서도 경험의 인식적 독립성을 수용할 수 있도록 만드는 경험의 현상적 성질에 대한 관점을 제시하는 윤보석의 주장은 관찰이나 현상의 표상으로서 데이터의 객관성을 포기하지 못 하는 과학철학 연구의 일반적인 경향성을 보여주는 좋은 사례다(윤보석 2021).

과학적 관찰의 인식론적 지위를 확립하고 과학적 실천의 합리성을 구원하려는 이러한 시도는 그 자체로 높은 연구로서 가치를 지니고 존중되어야 한다. 하지만, 데이터를 생성, 가공, 공유, 확산하는 복잡한 과정 또한 적절히 해석해야 하는 데이터 중심 과학 활동을 위한 전제로서 이 같은 인식론적 지위를 포기하는 것처럼 보이는 사비나 레오넬리의 시도 역시 과학적 실천과 관련된 새로운 합리성을 제시한다는 측면에서 중요한 가치를 갖는다. 그는 ‘데이터 여정(data journey)’이라 부르는 과학적 데이터의 생성, 모형화, 패키징, 탈맥락화, 공유, 확산을 통해 과학자 사회가 추구할 수 있는 새로운 생산성에 입각한 합리성과 비과학자의 개인적, 사회적 요구에 응답하는 과학적 실천의 실용적 가치 추구 방법을 제안한다. 이를 위해, 그는 관찰이 본질적으로 이론 의존적,

맥락 의존적이라는 것을 받아들이고, 데이터나 관찰 자체의 객관성을 추구하는 것 대신, 특정한 맥락 안에서 엄격하고 투명하며 책임감 있는 데이터 생산 관행의 산물로서 객관성을 이해해야 한다고 주장한다.

그에게 있어 데이터가 생산되고 확산되는 적절한 조건을 밝히는 것은 과학적 지식으로 간주되는 것과 그 인식론적 가치를 평가하기 위해 필수적이다.⁵⁾ 이를 위해 그는 데이터와 관련된 전 과정을 표현하는 개념으로서 ‘데이터 여정’을 다룬다. ‘데이터 여정’은 데이터가 그 유래로부터 다양한 연구 맥락으로 이동하는 방식을 의미한다.⁶⁾ 그가 다루는 ‘데이터 여정’은 인간 유래 데이터의 익명화나 메타데이터의 공유를 통해 데이터가 확산되는 데이터베이스 개념을 활용하는 데이터 중심 생물학의 특징을 명료하게 만들고, 그 바람직한 발전 방향을 제시하기 위해 제안되었지만, 하나의 관찰이나 실험을 함께 수행하는 공동 연구자들의 상호 작용이나 연구 종료 후 해당 연구의 절차로부터 생성된 데이터를 재사용하여 새로운 연구를 수행하는 경우처럼 여러 과학적 실천 상황에도 충분히 적용 가능하다. 동료 연구자들 사이에서 원활한 협력을 위해 정의되고 공유되는 데이터 모형화의 내용과 절차, 이를 거친 데이터 패키징, 하나의 연구에서 생성한 데이터를 또다른 연구에 활용하는 과정에서 기존의 데이터 모형화의 내용과 절차를 재검토 하는 한 명의 연구자는 모두 사비나 레오넬리의 관점에서 충실한 ‘데이터 여정’의 과정을 경험하게 될 것이다. ‘데이터 여정’이 설명할 수 있는 이런 복잡성은 과학적 관찰의 현재를 구체적으로 드러내고, 연구자가 과학적 관찰의 생산성과 다산성을 증가시키도록 돕는 적절한 설명 도구라 평가할 수 있다.

데이터에 대한 관계적 관점은 데이터 혹은 관찰의 이론 중립적인 특징을 옹호하는 관찰에 대한 고전적인 설명과 데이터 혹은 관찰의 이론적 재성을 인정하고 다양한 종류의 이론적재성을 비판적으로 고려하는 관찰에 대한 현대적인 설명이 모두 다루지 못 하는 관찰의 또다른 특징을 드러내기도 한다. 관찰의 이론적재성에 대한 논의들이나 그로부터 파생된 지각의 인지적 침투에 대한 기존의 논의들은 대부분 관찰 자체를 그

5) *Ibid.*, p. 2.

6) *Ibid.*, pp. 39-40.

분석 대상으로 삼는다는 측면에서 관찰과 이론, 혹은 데이터와 이론 사이의 더 구체적인 상호 작용을 명시적으로 드러내기 어렵다. 반면, 사비나 레오넬리는 ‘데이터에 대한 관계적 관점’을 통해 데이터 중심 생물학에서 이론의 독특한 지위를 다룬다. 그에 따르면, 데이터 중심 생물학에서 이론은 연구자에 의해 숙지된다는 측면이 다음과 같이 강조된다.

철학자 케네스 워터스가 제시한 유용한 구분을 바탕으로, 나는 데이터 중심 생물학이 이론 **주도적**(theory-driven)이 아니라, 이론 **숙지적**(theory-informed)이란 생각을 옹호한다. 데이터 중심 생물학은 이론에 기반을 두지만, 이론이 궁극적인 결과를 미리 결정하도록 내버려두지 않는다. 데이터 중심 과학에 대해 물어야 할 질문은 어떤 형태의 개념적 틀을 포함하는지 여부가 아니라, 다른 형태의 탐구와 구별되는 방식으로 이론을 활용하는지 여부이다(Leonelli 2016, p. 136, 강조는 원저자).

여기서 그는 이론을 연구자에 의해 알려지는 피동적인 것으로 설명하는 시도를 통해 데이터 중심 생물학에서 연구자의 주관성이 갖는 의미를 강조한다. 이 때, 연구자는 관찰의 이론적재성에 의해 영향을 받는 수동적인 존재도, 지각의 인지적 침투의 대상도 아니며, 적극적으로 이론을 탐색, 선택, 이해, 적용하는 주체다(Waters 2008). 그리고 자신이 수행하는 관찰의 인식론적 가치를 높이기 위해 관찰 가능한 세계나 데이터의 객관성을 확보하려고 노력하는 대신, 이론을 숙지하는 것과 같은 자신의 주관적 개입을 투명하게 드러내고 이를 기초로 더 생산적인 과학적 실천을 추구하기 위해 노력하여 ‘데이터 여정’의 인식론적 가치를 높일 수 있다.

이 때, 데이터에 대한 관계적 관점을 받아들이며 기대할 수 있는 관찰의 마음 의존적 특징의 독창적인 합리성이 드러난다. 데이터에 대한 관계적 관점에 입각한 관찰의 마음 의존적 특징, 즉 ‘데이터 여정’의 주관적 특징은 더 이상 가치 중립적이고 순수한 객관적 증거를 훼손하는 부정적인 요소가 아니라, 새로운 맥락에서 잠재적 세계나 현상에 대한 새로운 과학적 지식을 구축할 수 있는 데이터 패키징을 위한 적극적인 고려 대상이자 이러한 고려를 가능하게 만드는 ‘데이터 여정’의 본성으로 여겨진다.

3. 탈맥락화 과정에서 드러나는 ‘데이터에 대한 관계적 관점’의 문제점

그런데, 사비나 레오넬리가 새로운 맥락에서 잠재적 세계나 현상에 대한 새로운 과학적 지식을 구축할 수 있는 ‘데이터 여정’을 위해 필요하다고 주장하는 ‘데이터 큐레이션’의 내용은 아쉽게도 그가 고려하지 못한 의외의 문제를 일으킨다. 그에 따르면, 적절한 ‘데이터 큐레이션’은 단순한 데이터 저장 행위를 넘어선 적극적인 인식론적 과정이다.

그는 ‘데이터 큐레이션’을 다음과 같이 상호 연결된 두 단계의 과정, ‘탈맥락화(decontextualization)’와 ‘재맥락화(recontextualization)’로 나누어 설명한다. ‘탈맥락화’는 데이터를 원래의 맥락에서 추상화하고, 데이터의 ‘국소적(local) 특이성’을 제거하며, 표준화된 형식으로 데이터를 패키징 하여 더 넓은 사용을 위한 이동성과 상호 운용성을 갖추도록 만드는 과정이다. 사비나 레오넬리는 이 과정을 데이터가 이동하고 다른 연구 절차와 효과적으로 상호 작용하는 데에 필수적인 것으로 간주한다. ‘재맥락화’는 ‘탈맥락화’와 동등하게 중요한 후속 단계로서, 데이터 생산 맥락과 그 이후의 ‘데이터 여정’에 대한 정보로서 ‘메타데이터’가 데이터의 증거로서 잠재 능력을 이해하고 새로운 연구 맥락에서 적절하게 재목적화하는 데 필수적인 역할을 수행한다. 그는 이 때, 과학자들의 ‘체화된 지식(“노하우”)’이 데이터를 비판적으로 평가하고 효과적으로 활용하는 데 필요하다고 강조한다.⁷⁾

사비나 레오넬리는 성공적인 ‘탈맥락화’와 ‘재맥락화’에 의해 이루어진 ‘데이터 큐레이션’이 목표와 전문성 측면에서 많은 공통점을 갖지 않는 과학자들 사이에서 데이터의 공유와 확산을 통해 기존 데이터에 대한 완전히 새로운 해석을 얻을 수 있다는 측면을 강조하는데,⁸⁾ 이는 명백히 그가 ‘데이터 여정’으로 추구하고자 하는 독보적인 인식론적 가치와 실천적 가치를 잘 보여준다.

7) *Ibid.*, p. 5, pp. 29-31.

8) *Ibid.*, p. 31.

그러나 사비나 레오넬리가 과학자 사회에서 데이터의 인식적 가치를 높이는 것뿐만 아니라, 비과학자 사회, 즉 시민 사회의 다양한 가치를 추구하기 위해 데이터를 다양한 응용 분야에서 효과적으로 만들기 위해 강조한 ‘탈맥락화’는 데이터 자체에 내재된 인간적 특징을 제거하는 ‘비인간화(dehumanization)’를 일으킬 수 있다. 그는 ‘탈맥락화’에서 발생할 수 있는 문제를 예방하기 위한 노력에 대해 언급하고 있으나,⁹⁾ 그가 관심을 두는 문제는 어디까지나 데이터의 ‘불량’ 여부나 데이터의 신뢰도와 같이 데이터의 배포 과정에서 일어나는 문제와 그에 대한 책임에 국한되어 있으며, 이는 ‘데이터에 대한 관계적 관점’으로 파악 가능한 과학자와 비과학자, 그들의 행위 목적과 다양한 상호 작용 사이에서 발생 가능한 도덕적 문제나 인권에 대한 고려와 같은 인간적 문제들과는 차이가 있다. 따라서, 구체적이고 명시적으로 규정 가능한 메타 데이터의 정의와 그 적용 방법이 제시되지 않는다면, 데이터는 그 여정 중 특정한 인간 활동, 의도, 맥락을 생성한 데이터 주체와의 연결을 잃고 일반적이고 추상적인 기록물로 전락할 수 있다. 이는 복잡한 인간 경험이나 본성을 추상적이고 재사용 가능한 단위로 축소하여 과학적 실천의 비인간화에 직접적으로 기여할 수 있다.

특히, 인간 스스로를 위해 자동화된 알고리즘과 인간 유래 데이터를 사용하기 위한 맥락에서 조성되고 있는 최근의 논의들은 앞서 설명한 ‘탈맥락화’에 의한 과학적 실천의 비인간화에 대해 우려하도록 만든다. 일반적으로 ‘알고리즘에 의한 비인간화(algorithmic dehumanization)’는 알고리즘과 데이터를 사용하여 개인이나 집단이 가진 자유, 평등, 존엄성 등의 가치를 침해하는 의도적, 또는 비의도적 행위로 정의된다(Schultz et al. 2024). 만약, 새로운 연구 절차를 위한 ‘재맥락화’를 위한 ‘탈맥락화’를 통해 데이터 접근권, 데이터 이동권, 데이터 삭제권, 잊힐 권리와 같은 인간 유래 데이터를 생산한 데이터 주체의 다양한 데이터 권리가 침해된다면, 이는 ‘알고리즘에 의한 비인간화’를 야기하는 데이터 중심 과학 활동의 구조적 문제라 여길 수 있다.

데이터베이스를 통해 확산, 발전하는 데이터 중심 생물학 분야에서

9) Ibid., pp. 46-7, p. 53, 59, 110, pp. 142-3, pp. 174-5.

‘데이터 여정’이 풍요롭게 지속될수록, 전체 ‘데이터 여정’에서 특정한 인간 대상자의 행동, 신체 특징, 인체 유래물로부터 최초로 데이터를 생산하는 상황이 차지하는 시간적, 물리적 비율은 필연적으로 작아진다. 이는 앞서 설명한 두 가지 문제들, 다양하고 특수한 인간적 특징을 상실한 추상적인 기록물로 전락한 데이터의 가치 축소와 과학적 실천 과정에서 인간 대상자의 존엄성에 대한 충분한 고려의 부재에 대한 취약성이 점점 증가할 수 있다는 측면에서 더 큰 우려를 낳는다.

4. ‘정보 유기체’ 개념을 적용한 ‘데이터에 대한 확장된 관계적 관점’

앞서 살펴본 데이터에 대한 기존의 관계적 관점이 야기할 수 있는 문제를 방지하고, 데이터 주체로서 인간 대상 연구에 참여하는 인간 대상자와 데이터의 관계를 더욱 명료하게 설명하기 위해, 나는 루치아노 플로리디가 제안한 ‘정보 유기체(inforg)’ 개념을 데이터 중심 과학, 더 나아가 데이터를 생산하는 모든 과학적 실천 절차와 상황에 적용하는 것을 제안하고자 한다.

루치아노 플로리디의 개념들은 정보를 존재론적 논의의 맥락에서 이해하는 그의 정보 철학에 뿌리를 두고 있다. 그는 모든 정보적 실체, 속성, 상호 작용, 과정, 관계의 총체로서 ‘정보권(infosphere)’이란 개념을 가정한다. 그리고 독립적이고 고유한 실체가 아닌, ‘정보권’ 안에서 정보적 상호 작용과 관계를 맺는 정보적 실체로 존재하는 구성 요소가 ‘정보 유기체’다. ‘정보 유기체’는 근본적으로 정보로 구성되는데, 이때 정보는 고전적인 의미의 새넌 정보로서, 발생 확률을 가진 사건들의 불확실성을 양적으로 나타내며, 특정 사건에 대한 불확실성을 줄여주는가에 따라 측정 가능한 양을 갖는다. 이는 단순히 추상적 개념으로서 정보일 뿐만 아니라, 현실의 구성 요소로서 물질 및 에너지와 같은 물리적 존재자들과 동등한 위치로 취급된다.¹⁰⁾

루치아노 플로리디는 ‘정보 유기체’의 유형에 대한 엄격한 분류 체계

를 제안하진 않지만, ‘정보권’ 내에서 정보 행위자들에 대한 분류는 시도하고 있다. 먼저 인간은 대표적인 ‘정보 유기체’로서 의미론적 정보를 처리하는 능력, 또는 의미론적 엔진으로서의 능력을 그 특징으로 가진다. 이때, 의미론적 정보를 처리하는 능력은 의미를 생성할 수 있다는 것을 의미한다. 그는 인간을 우주에서 현실에 대한 지식을 발전시킬 수 있는 유일하게 알려진 의미론적 엔진이자 의식적인 정보 유기체로 묘사한다. 그의 정보적 존재론이 보편적 설명을 추구하고 있음에도 불구하고, 그는 의미론적 정보를 처리하고 의미를 생성하는 인간의 고유한 능력에 기반하여 인간을 명시적으로 구별한다. 그에 따르면, 의미론적 정보를 처리하는 인간의 능력은 인간이 현실에 대한 지식을 발전시키는 것뿐만 아니라, 도덕적 행위자로 행동할 수 있도록 만든다. 루치아노 플로리디는 인간 ‘정보 유기체’ 외에도 자동적으로 정보를 처리하는 정보적 실체이자 구문론적 정보 처리기로서 알고리즘, 자연적이지 않지만 도덕적 영향을 미치는 행동을 수행할 수 있는 실체로서 도덕적 설명 책임을 갖는 자율적 체계로서 ‘인공적 행위자’, 인간을 포함한 생물학적 유기체로서 ‘정보권’ 내에 존재하는 ‘자연적 행위자’, 인간과 ‘인공적 행위자’ 사이를 매개할 수 있는 디지털 장치와 조합된 ‘정보 유기체’로서 하이브리드 행위자들을 ‘정보 유기체’로서 ‘정보권’ 내 행위자들로 제시한다.¹¹⁾

이에 따르면, 과학적 실천에서 과학자, 즉 연구자와 비과학자, 인간 연구 대상자는 인간 ‘정보 유기체’로, 잠재적 관찰 세계 혹은 현상 중 불확실한 사건을 일으키는 행위자로서 다양한 대상들은 ‘자연적 행위자’로, 연구 절차 중 디지털 장치에 의해, 혹은 특정한 제도나 문화에 의해 불확실한 사건을 자동적으로 일으키는 기술적 매개물들을 ‘알고리즘’으로 여길 수 있다. 이 중 데이터의 내용을 결정하는 연구자와 비과학자, 연구 절차와 데이터 모형화 및 데이터 패키징 중 정보를 자동적으로 처리, 가공하는 ‘알고리즘’은 사비나 레오넬리의 데이터에 대한 관계적 관점에 의해 충분히 다루어졌으므로, 나는 데이터의 유래, 생성의 맥락과 절차에 대한 ‘탈맥락화’ 문제의 구체적인 대상이 되는 동시

10) Floridi (2013), pp. 14-8.

11) Floridi (2014), pp. 94-6.

에, 데이터에 대한 기존의 관계적 관점에서 다루지 못 했던 인간 유래 데이터의 데이터 주체, 즉 인간 ‘정보 유기체’로서 인간 대상 연구의 인간 대상자를 주로 검토하고자 한다.

인간 대상 연구는 여러 영역에서 정의되고 있는데, 이 논의에서는 주로 두 영역의 개념을 아울러 활용하고자 한다. 하나는 연구 윤리 분야의 한 개념으로서 인간 대상 연구다. 특히 인간과 인체유래물의 연구를 다루는 대한민국 법률인 ‘생명윤리 및 안전에 관한 법률’은 인간 대상 연구를 “사람을 대상으로 물리적으로 개입하거나 의사소통, 대인 접촉 등의 상호 작용을 통하여 수행하는 연구 또는 개인을 식별할 수 있는 정보를 이용하는 연구”라 규정한다.¹²⁾ 또 하나는 주로 인간의 심리적 과정을 포함한 인간의 행위와 개인들 사이의 상호 작용을 연구하는 ‘사회 행동 과학 연구(Social and Behavioral Sciences Research)’다.¹³⁾ 전자의 경우, 주로 연구 대상자에게 침습적 행위나 연구 대상자의 환경을 조작하는 물리적 개입이 포함된 연구를 말한다는 측면에서 이에 해당하는 의학, 약학, 생리학, 생화학 분야가 포함된다. 후자의 경우, 실재하는 인간을 관찰이나 실험을 통해 실증적으로 연구한다는 측면에서 이에 해당하는 심리학, 사회과학 분야가 포함된다.

또한 양적 연구를 지향하는 사회, 행동 과학자들이 만들어내는 수치(numerical) 데이터와 질적 연구를 지향하는 사회, 행동 과학자들이 만들어내는 서술적(narrative) 데이터 모두가 이러한 인간 대상 연구의 수행 과정에서 생산되는 데이터에 포함된다.¹⁴⁾ 또한, 인간 대상자가 일상적인 생활 속에서 제공하는 생체 신호와 행동학적 특징을 실시간, 근실시간으로 기록하는 ‘생태 순간 평가(ecological momentary assessment)’ 데이터와 같은 데이터 역시 본 논의에서 다루는 인간 대상 연구 과정에서 생산되는 인간 유래 데이터에 해당한다.

인간 대상 연구 과정에서 생산되는 이 같은 맥락의 데이터는 독립적

12) 생명윤리 및 안전에 관한 법률 제2조 제1항.

13) NIH (2025). 해당 문헌은 “Behavioral and Social Sciences”란 표현을 사용하고 있으나, 더 널리 쓰이는 “Social and Behavioral Sciences”란 표현과 혼용 가능하다는 판단에 의해 ‘사회 행동 과학 연구’란 용어를 사용했다.

14) Tashakkori *et al.* (2021), p. 4.

으로 존재하는 대상으로서 인간 대상자가 홀로 생산하거나, 인간 대상자와 연구자 사이의 관계에서 생산되는 것이 아니다. 인간 대상자는 데이터를 생산하는 매순간 매우 다양한 요소들과 복잡한 인과 관계를 맺는다. 연구자에 의해 제시된 자극과 물리적 관계를 맺기도 하고, 이를 자신의 주관적 경험을 통해 의식 수준으로 떠올리기도 한다. 연구자가 미처 파악하거나 고려하지 않지만, 삶을 지속하기 위해 대기 중의 산소와 물리적인 상호 작용을 맺고, 중력의 영향에 의해 골격근의 수축을 경험한다. 연구자에 의한 관찰 이전에 경험한 세계와의 어떤 상호 작용에 대한 기억을 자각하여 신체의 변화를 겪기도 한다. 이 중 어떤 것은 이미 연구자에 의해 알려지거나 특정한 이론이나 지식에 의해 규정된 채로 연구자에 의해 다시 관찰되기도 하지만, 어떤 것은 연구자에 의해 전혀 알려지지 않거나, 연구자의 목적에 의해 의도적으로 무시되기도 한다.¹⁵⁾

사비나 레오넬리의 데이터에 대한 관계적 관점에 따르면, 관찰 가능한 데이터의 내용을 결정하는 것은 연구를 수행하는 과학자와 과학자 사회, 그리고 비과학자 사회다. 즉, 관찰 가능한 데이터의 내용, 혹은 관찰해야 할 데이터의 내용은 일단 연구를 수행하는 과학자가 수립한 구체적인 연구 목표와 그가 속한 전문 분야의 동료 과학자들이 추구하는 좀더 추상적인 연구 목표 혹은 인식론적 가치에 의해 결정된다. 또한, 데이터의 내용은 더 적절한 ‘데이터 여정’을 위해 과학자가 아닌 다양한 개인들로 구성된 사회가 추구할 수 있는 다양한 가치에 의해 결정되기도 한다.

하지만, 관찰 가능한 데이터의 원천으로서 인간 행위자, 즉 잠재적 관찰 세계에서 불확실한 사건을 일으키는 동시에 이것의 의미를 자신만의 추상화 능력을 활용하여 인과적 관계를 형성하는 다른 대상들과의 관계를 통해 파악하는 인간 ‘정보 유기체’가 인간 대상 연구에서 데이터의 원천이 되는 과정은 다음과 같이 더욱 복잡한 특징을 지니며, 나는 이를 적극적으로 고려하는 태도를 “데이터에 대한 확장된 관계적 관점”이라 부르고자 한다.

15) 본 논문이 다루는 인간 대상 연구 영역과 인간 유래 데이터의 종류, 논의를 위한 인간 대상 연구에 대한 상세한 예시적 설명을 요구하여 논문의 내용을 명료하게 만드는 데에 도움을 주신 심사자에게 감사를 표한다.

- 1) 연구자와 다른 동료 과학자들, 혹은 과학자가 아닌 다양한 개인들로 구성된 사회가 추구하는 인식적, 비인식적 가치에 의해 연구 절차 중 생성될 데이터의 내용이 결정된다.
- 2) 결정된 데이터의 내용에 의해 인간 대상자는 ‘정보권’ 내에 존재하는 다양한 ‘자연적 행위자들’, ‘알고리즘’, ‘인공적 행위자’, 혹은 또다른 인간 ‘정보 유기체’와 함께 상호 작용을 한다.
- 3) 인간 대상자는 자기 자신의 주관적 믿음을 형성하는 다양한 동기와 지식들, 자신이 지금까지 상호 작용해왔던 다른 사회적 관계를 통해 상호 작용의 결과로서 추구할 데이터의 내용을 결정한다.
- 4) 인간 대상자는 수행한 상호 작용과 관련된 의미론적 정보를 지식으로 만든다.
- 5) 연구자는 4)의 지식을 현재 고려 중인 맥락의 ‘데이터 여정’에 적합한 데이터 모형을 통해 데이터 패키징을 수행한다.
- 6) 연구자는 데이터 패키징을 마친 데이터를 공유, 확산하는 과정에 필요한 추상화를 통해 형성된 ‘정보권’ 내에 존재하는 다른 과학자들, 비과학자들, 데이터를 자동적으로 처리하는 무형의 알고리즘들, 각각의 행위자들 사이를 매개하는 ‘자연적 행위자들’과 상호 작용을 한다.
- 7) 데이터 큐레이션의 주체는 데이터 큐레이션을 위한 탈맥락화 과정에서 재맥락화한 목적과 상호 작용에 필요한 추상화를 거쳐 새로운 ‘정보권’을 형성한다.
- 8) 데이터 큐레이션의 주체는 인간 ‘정보 유기체’로서 데이터 큐레이션을 위한 ‘정보권’ 안에서 인간 대상자, 다른 과학자들, 비과학자들, 데이터 처리자들, 무형의 알고리즘들, 각각의 행위자들 사이를 매개하는 ‘자연적 행위자들’과 새로운 상호 작용을 한다.
- 9) 데이터는 공유를 통해 확산된다.
- 10) 1), 2), 3), 4), 5), 6), 7) 과정에서 인간 ‘정보 유기체’들은 도덕적 책임을, 2), 6) 과정에서 ‘인공적 행위자’들은 설명 책임을 진다.

인간 대상 연구에서 위와 같이 새롭게 해석한 데이터에 대한 확장된 관계적 관점에 따르면, 이전의 데이터에 대한 관계적 관점에 의해 형성된 과정에 비해 새로운 차이가 발생한다. 먼저, 데이터의 원천인 현상

으로 해석되던 인간 대상자는 이전에 포착되지 않았던 ‘정보 유기체’로서 처야 할 도덕적 책임을 새롭게 지게 된다. 이것은 우리가 기존의 인간 대상 연구 윤리에 의해 도덕적 고려 대상으로 여겨졌던 인간 대상자에게 부여해야 할 새로운 역할과 책임, 권한을 고려해야 한다는 것을 의미한다. 또한, 우리가 비인간화의 문제를 야기할 수 있었던 기존의 탈맥락화를 기반한 데이터 큐레이션에 대응할 수 없었던 것과 달리, 8)의 과정을 통해 데이터 큐레이션의 주체를 데이터 큐레이션을 위한 새로운 ‘정보권’에서 새로운 상호 작용을 통해 도덕적 책임을 다하는 주체로 고려할 수 있다.

또한, 데이터에 대한 확장된 관계적 관점은 여러 도덕적 주체들과 고려 대상이 복잡하게 상호 작용하는 인간 대상 연구를 정보적 실체들 사이에서 형성되는 하나의 ‘정보권’으로 고려한다는 측면에서, 의식적 주체로서 갖는 의도나 개별 행위자의 궁극적 의도와 같이 현실적으로 파악하거나 고려하기 매우 어려운 요소에 주목하는 대신, 도덕적으로 중립적인 것처럼 보이거나 도덕적으로 무시할 수 있는 상호 작용의 결과로 발생하는 복잡적 주체의 도덕적 행위를 다루는 ‘분산된 도덕성’과 같은 새로운 실천적 방안을 고려할 가능성을 우리에게 제공한다.¹⁶⁾

더불어, 데이터에 대한 확장된 관계적 관점은 연구 절차 내 새롭게 포착될 수 있는 발생 가능한 ‘인공적 해악’에 대응할 수 있다. ‘인공적 해악’은 루치아노 플로리디가 제안한 ‘정보권’에서 포착할 수 있는 새로운 종류의 해악이다. 그에 따르면, ‘인공적 해악’은 전통적인 윤리학 체계에서 포착할 수 없으며, 인간 ‘정보 유기체’가 선의를 가지고 있음에도 불구하고 ‘자연적 행위자’의 의도 없는 행위나 ‘알고리즘’의 자동적 상호 작용의 실패, ‘인공적 행위자’의 블랙박스 속 행위에 의해 비극을 초래한다는 측면에서 치명적이다.¹⁷⁾ ‘인공적 해악’의 발생 가능성은 ‘데이터 여

16) Floridi (2013), pp. 262-4.

17) 플로리디는 인공적 해악의 대표적인 사례로 수술 도중 모니터의 오작동으로 인해 수술을 정상적으로 진행하지 못 해 환자가 겪는 치명적인 위협과 같은 상황을 설명한다. 이러한 ‘인공적 해악’은 해악의 영향을 받는 행위자, 즉 피동자의 안녕을 부패하게 만들거나 훼손되는 것으로 규정되며, 플로리디는 이를 새로운 정보 기술에 대한 높은 이해와 정보 윤리적 관점을 통해 방지할 수 있는 새로운 종류의 문제라 여기고 있다.

정'을 통해 데이터가 지속적으로 공유, 확산되는 과정에서 새로운 연구 절차와 그와 관계된 새로운 행위자들 사이에서 증가하는 복잡성을 고려했을 때, '데이터 여정'의 지속과 함께 증가할 가능성이 매우 크다.

더불어, 데이터에 대한 확장된 관계적 관점은 이전까지 우리가 포착할 수 없었던 인간 대상자의 지각적 특징을 고려하도록 돕는다. 이전까지 고전적인 관점에서 관찰을 수행하는 연구자가 겪는 지각의 인지적 침투만을 고려했던 것과 달리, 이제 우리는 인간 대상자가 겪는 그의 지각의 인지적 침투를 고려할 수 있다. 연구 대상으로서 현실(reality)에 스스로 의식적 주체로서 인간 대상자의 주관적 현실(subjective reality)과 언어와 문화를 통해 공유되는 상호 주관적 현실(intersubjective reality)을 포함하여 다룰 수 있는 연구 방법론을 확립하는 것은 이미 사회 행동 과학 분야에서 오랜 연구 주제였다.¹⁸⁾ 하지만 과학철학 분야에서 연구 대상자가 외부 세계를 지각한다는 측면에서 세계와 맺는 복잡한 상호 작용을 고려하여 연구 대상자를 경험적으로 연구하는 것의 본성을 다루는 일은 아직 주목 받고 있지 못 하다. 일찍이 이언 해킹과 같이 타인에 의한 규정에 영향을 받거나 자기 자신에 대한 지식의 숙지 여부에 영향을 받는 독특한 종으로서 인간의 특징을 연구한 사례는 존재했지만(Hacking 1996), 이는 어디까지나 연구 대상으로서 연구 대상자와 연구자 사이의 관계나 연구 대상자와 이론이나 지식의 영향을 받은 타인과의 관계에서 드러나는 특징이라는 측면에서 앞서 설명한 정보 유기체로서 연구 대상자가 맺는 정보적 관계의 일부에 불과하다. 이언 해킹이 제안한 '고리 효과(looping)'와 같은 개념이 주로 특정한 정신 질환이나 각인 효과와 같이 어느 정도 체계적으로 구성된 이론적 개념이나 특정 분야의 지식을 중심으로 설명된다는 점은 이 같은 선행 연구가 인간에 의해 부여된 어떤 가치나 그의 주관적 요소, 즉 전통적인 의미에서 관찰의 이론적재성과 유사한 방식으로 대상자에게 영향을 미치는 요소만을 다룬다는 것을 보여준다.¹⁹⁾

Ibid., p. 187을 참고할 것.

18) Tashakkor *et al.* (2021), p. 20-1.

19) 이언 해킹의 '인간중' 관련 논의를 적용하여 빈약했던 인지적 침투의 확장에 대한 논

정보 유기체로서 인간 대상자가 스스로 시도한 추상화를 통해 의미 있는 정보를 생산하며 정보적 관계를 맺는 다른 정보적 대상들과 상호 작용하는 것을 파악하는 일은 우리가 그간 포착되지 않았던 독립적인 체계, 혹은 더 광범위한 체계로서 또 하나의 상호 주관적 믿음 체계를 연구자와 데이터의 원천 사이에서 발견 가능하다는 것을 말한다. 따라서 이는 정당화의 맥락에서 관찰이나 경험, 믿음의 인식론적, 존재론적 가치를 다뤘던 기존의 논의와 달리, 이제 우리가 은닉되어 있었던 현상의 새로운 이면을 밝히는 발견의 맥락에서 데이터의 원천이 가진 새로운 지각적 특징으로서 정보 유기체들 사이에 존재하는 광범위한 지각의 인지적 침투를 고려해야 한다는 것을 보여준다.

5. 데이터에 대한 확장된 관계적 관점의 의의와 한계

루치아노 플로리디의 ‘정보 유기체’ 개념과 사비나 레오넬리의 데이터에 대한 관계적 관점을 통합하는 “확장된 관계적 관점”은 우리가 인간 대상 연구에서 이미 발생 중이지만 드러나지 않았던 윤리적 문제들과 데이터 중심 과학의 발전과 ‘데이터 여정’의 지속에 따라 새롭게 발생할 윤리적 문제들에 대해 더 깊고 명료하게 이해할 수 있도록 돕는다. 이는 데이터 소유권, 데이터 접근권, 데이터 이동권, 데이터 전송권, 데이터 삭제권, 잊힐 권리 등의 데이터 권리와 인간의 선의나 악의와 독립적으로 비인간 행위자들과의 복잡한 상호 작용 속에서 일어나는 ‘인공적 해악’과 관계된 새로운 도전과 책임에 대해 고려하도록 우리를 이끈다.

사비나 레오넬리가 제안하는 관계적 데이터 관점은 이미 실용 단계에서 많은 연구를 수행하고 있는 인간의 질병, 유전 정보, 검체와 시료 등의 의료적, 생물학적 데이터베이스를 기초로 만들어진 인간의 데이터

관련 행위와 연구를 기초로 제안되었다. 그러나, 이러한 의료적, 생물학적 데이터베이스는 난치병이나 불치병의 치료법과 치료제에 대한 연구에 활용할 수 있다는 실용적 가치를 이유로, 혹은 익명화를 통해 데이터의 원천인 실제 인간에 대한 추가적인 해악이 쉽게 눈에 띄지 않는다는 점을 이유로, 여러 윤리적 난점에도 불구하고 빠르게 도입, 확산되었다. 데이터에 대한 확장된 관계적 관점은, 데이터의 이동성을 높이고, 손쉽게 데이터와 데이터 모형을 재맥락화 하기 위해 경시되어왔던 의학, 생물학 분야의 인간 유래 데이터의 윤리적 활용 방안에 대한 새로운 시각을 제공할 수 있다. 제시되는 새로운 맥락과 절차, 목적에 따라 끊임 없이 변화하는 ‘데이터 여정’에 따라, 우리는 취약한 연구대상을 위한 새로운 동적 동의 과정이나 프라이버시 보호 방법을 고안할 수 있을지 모른다.

더불어, 인간 대상자가 연구자에 의해 주어진 자극을 경험하는 과정에서, 실험 환경이나 자신이 관찰되는 상황에서 다양한 외부 환경을 지각하며 겪는 인지적 침투에 대한 새로운 고려는 관찰이나 데이터의 마음 의존적 특징에 관여하는 연구의 주체가 연구자 뿐만이 아니라는 사실을 우리에게 환기시킨다. 이것은 연구 대상자를 또 하나의 연구 수행 주체로 해석하도록 이끈다는 측면에서, 우리로 하여금 연구자와 연구 대상, 전문가와 비전문가 사이에서 존재하는 위계가 위협하는 연구 대상자의 자유와 자율성, 존엄성에 대해 한 번 더 숙고하도록 만든다.

이와 같은 데이터에 대한 확장된 관계적 관점의 의의에도 불구하고, 이러한 데이터에 대한 새로운 관점은 언제나 비판적으로 검토되어야 할 필요가 있다. 사비나 레오넬리가 데이터의 공유와 확산을 위한 ‘데이터 여정’의 확대를 위해 제안했던 데이터에 대한 관계적 관점이 데이터 큐레이션 과정으로 다루었던 탈맥락화가 인간 유래 데이터의 데이터 주체에 대한 우리의 인식을 방해하고 알고리즘에 의한 비인간화를 야기할 수 있었다는 사실은 데이터에 대한 관계적 관점이 가진 근본적인 문제를 지적한 것이 아니라, 데이터에 대한 새로운 관점들이 다양한 연구 절차와 목적, 문화를 가진 여러 전문 분야에 적용될 때 미처 예상하지 못 했던 문제가 발견될 수 있다는 점을 지적한 것에 불과하기 때

문이다. 이는 데이터 중심 과학으로서 전문 분야가 증가하고, 더욱 다양한 인간 유래 데이터를 연구 과정에서 다루는 전문 분야가 증가함에 따라, 이미 받아들이게 된 데이터나 관찰, 연구 윤리에 대한 우리의 관점을 늘 비판적인 시각으로 검토해야 한다는 것을 의미한다.

현상, 데이터, 이론에 대한 오래된 믿음을 비판적으로 검토하여 기존의 과학적 실천의 생산성을 높이고, 인류에게 유용한 새로운 연구의 사회, 경제적 비용을 안전하고 합리적으로 줄일 수 있도록 만들기 위해 사비나 레오넬리가 데이터 기반 생물학으로부터 제안한 데이터에 대한 관계적 관점은 그 실용적 가치에도 불구하고 데이터의 증거적 가치를 옹호하기 위한 새로운 해결책을 제시하거나(Gim 2023), 익명화를 거쳐 다시 동의를 얻기 어려운 인간 대상자로부터 유래한 데이터 집합으로 구성된 데이터베이스를 지속적으로 활용하기 위한 윤리 프레임워크를 마련해야 하는 등, 새롭게 해결해야 할 다양한 문제를 야기하기도 한다. 그러나 본 연구가 데이터에 대한 확장된 관계적 관점을 제안한 것처럼 이러한 문제들을 해결해 나가는 과정에서, 인류는 건강과 복지와 같은 실질적 안녕을 추구하는 것뿐만 아니라 관찰과 경험, 현상과 이론과 같이 발명과 함께 그 적절한 활용 방안에 대해 지속적으로 고민해 온 여러 개념들에 대한 이해를 넓혀갈 수 있을 것이다.

참고문헌

- 생명윤리 및 안전에 관한 법률[법률 제20327호, 2024.2.20., 일부개정]
 윤보석 (2021), 「인지적 침투, 결과주의, 그리고 경험적 합리성: 천현득 교수의 논의를 중심으로」, 『과학철학』 24권 2호, pp. 1-29.
 Tashakkori, A. et al. (2021), *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences*, Sage publications, Inc.
 Bogen, J. and Woodward, J. (1988), “Saving the Phenomena”, *The*

- Philosophical Review* 97(3): pp. 303-52.
- Floridi, L. (2013), *The Ethics of Information*, Oxford university press.
- Floridi, L. (2014), *The Fourth Revolution*, Oxford university press.
- Gim, J. (2023), 「Data, representation, and evidential values in biology」, 『과학철학』 26권 2호, pp. 31-58.
- Hacking, I. (1996), “The Looping Effects of Human Kinds”, in Sperber, D. et al. (eds.), *Causal Cognition: A Multidisciplinary Debate*, Oxford university press.
- Lionelli, S. (2016), *Data-centric Biology: A Philosophical Study*, The University of Chicago press.
- National Institution of Health (2025), About Behavioral and Social Sciences Research
- Schultz, M. D. et al. (2024), “Algorithms and Dehumanization: A Definition and Avoidance Model”, *AI & society* 40: pp. 2191-211
- Suppes, P. (1962), “Models of Data”, in Nagel, E. et al. (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford, CA: Stanford university press, pp. 252-61.
- Waters, C. K. (2008), “How Practical Know-how Contextualizes Theoretical Knowledge: Exporting Causal Knowledge from Laboratory to Nature”, *Philosophy of Science* 75: pp. 707-19.

논문 투고일	2025. 06. 23.
심사 완료일	2025. 07. 09.
게재 확정일	2025. 07. 21.

Expanding the Relational View on Data and Cognitive Penetration of Perception for Human Subject Research

Kyungjoon Oh

This paper clarifies the characteristics of human-oriented data generated by human subject research and reexamines the significance of human subjects in scientific practice. To this end, I first criticize the traditional view on scientific observation, introduce the relational view on the data presented by Sabina Leonelli and explain its significance and limitations. The relational view on data is useful for clarifying and advancing the characteristics of data-centric biology, which constructs and uses databases for research through active use of data and metadata as a data-centric science, but it does not reflect the unique characteristics of human-oriented data generated from interactions with human subjects in the process of decontextualization for dissemination and sharing of data. I attempt to expand the scope of application of the existing relational view on data and the concept of cognitive penetration of perception to provide a new perspective on human subjects as active subjects of research that should be newly considered in scientific practice by reflecting the unique characteristics of human-oriented data.

Keywords: the relational view on data, penetration of perception, decontextualization, inforg, data-centric science

