

팔란티어 온톨로지 vs. 전통 온톨로지

- 기획: 페블러스 데이터 커뮤니케이션팀
- 작성: 2025-09-12
- AI 도구: Gemini
- 인터랙티브 콘텐츠: <https://blog.pebblous.ai/>

목차

- I. 서론: 온톨로지 개념의 이중주
- II. 온톨로지 기술의 기초: 지식의 디지털화
 - 2.1. 컴퓨터 과학 속 온톨로지의 탄생과 목적
 - 2.2. 핵심 구성 요소 및 구조
 - 2.3. 시맨틱 웹의 언어와 기술
 - 2.4. 전통적 온톨로지의 한계와 도전 과제
- III. 팔란티어 온톨로지: 현실 비즈니스의 디지털 트윈
 - 3.1. 기존 데이터 플랫폼의 문제점과 팔란티어의 접근 방식
 - 3.2. 팔란티어 온톨로지의 3계층 아키텍처
 - 3.3. 온톨로지 구축 및 운영의 실제
- IV. 심층 비교 분석 및 적용 사례
 - 4.1. 전통적 온톨로지 vs. 팔란티어 온톨로지 비교
 - 4.2. 팔란티어 온톨로지의 실질적 가치 창출 사례
- V. 결론 및 미래 전망
 - 5.1. 온톨로지 패러다임의 전환
 - 5.2. 데이터 기반 조직을 향한 전략적 시사점
 - 5.3. 미래 전망

I. 서론: 온톨로지 개념의 이중주

본 보고서는 현대 기술 담론에서 중요한 위치를 차지하는 두 가지 온톨로지 개념, 즉 '전통적 온톨로지'와 '팔란티어 온톨로지'에 대해 심층적으로 분석하고자 한다. 온톨로지는 본래 철학의 한 분야인 존재론(Ontology)에서 유래하여 존재의 근본 원리를 탐구하는 학문이었으나, 컴퓨터 과학 분야로 넘어오며 지식과 개념을 명확하게 정의하고 체계화하는 기술로 재탄생했다 [1, 2, 3]. 이 용어는 인공지능(AI), 시맨틱 웹(Semantic Web), 자연어 처리 등 여러 분야에서 지식의 처리, 공유, 재사용을 위한 핵심 도구로 활용되어 왔다 [1].

그러나 최근 몇 년간 빅데이터와 AI 기술의 발전은 기존의 온톨로지 개념을 재해석하고 확장하는 새로운 흐름을 만들어냈다. 특히 데이터 분석 및 운영 솔루션 기업 팔란티어(Palantir)는 온톨로지를 단순한 지식 표현 수단을 넘어, 기업의 복잡한 현실을 디지털로 모델링하고 실시간 의사결정을 지원하는 독자적인 플랫폼의 핵심으로 제시하였다.

본 보고서는 이러한 온톨로지 패러다임의 변화를 이해하기 위해, 전통적 온톨로지의 역사, 구조, 한계를 먼저 명확히 한 후, 팔란티어가 제시하는 온톨로지의 독특한 아키텍처와 기능을 상세히 설명한다. 이어서 두 개념을 다각도로 비교 분석하고, 실제 적용 사례를 통해 각각의 근본적인 차이점과 전략적 함의를 도출할 것이다. 이 분석은 데이터 기반 의사결정 시스템 구축을 고민하는 기업의 의사결정권자 및 기술 전문가들에게 귀중한 지침을 제공할 것이다.

II. 온톨로지 기술의 기초: 지식의 디지털화

2.1. 컴퓨터 과학 속 온톨로지의 탄생과 목적

온톨로지라는 용어가 컴퓨터 과학 분야에 도입된 것은 1980년대 인공지능(AI) 연구자들이 지식 공학(knowledge engineering)을 연구하면서부터였다 [1, 2]. 이들은 인간이 가진 지식을 컴퓨터가 이해하고 처리할 수 있는 형태로 표현하는 방법을 모색했고, 이 과정에서 철학적 개념인 온톨로지를 차용했다.

스탠퍼드 대학의 토마스 A. 그루버(Thomas A. Gruber) 교수는 온톨로지를 "공유된 개념화를 형식적으로 명백하게 기술하는 명세(an explicit formal specification of a shared conceptualization)"라고 정의하며 이 분야의 핵심 개념을 정립하였다 [1, 4]. 이 정의의 핵심은 '공유(shared)'와 '재사용(reusability)'에 있다 [5]. 이는 특정 분야의 전문가들이 합의한 지식 체계를 컴퓨터가 처리할 수 있는 형식으로 명시하여, 해당 지식을 여러 시스템과 인간이 공통적으로 사용하고 재활용할 수 있는 틀을 제공하는 것을 의미한다 [2, 5]. 온톨로지의 궁극적인 목적은 프로그램과 인간이 지식을 공유하도록 돋고, 정보 시스템 내의 개념을 명확히 정의함으로써 보다 정확한 정보 탐색을 가능하게 하는 데 있다 [6, 7].

2.2. 핵심 구성 요소 및 구조

전통적인 온톨로지는 일반적으로 네 가지 기본 구성을 통해 지식을 구조화한다 [1, 3, 5].

- **클래스(Class):** 개념의 범주 또는 집합을 의미한다. "사람", "자동차", "사랑"과 같은 일반적인 사물이나 개념에 붙이는 이름이다 [1, 5].
- **인스턴스(Instance):** 특정 클래스에 속하는 개별적인 실체나 사건을 의미한다. (예: '평창'이라는 인스턴스는 '올림픽' 클래스와 관계를 맺는다) [1]. 이는 클래스에 대한 구체적인 실체로 나타난 그 자체를 뜻한다 [5].
- **속성(Property):** 클래스나 인스턴스가 갖는 특징이나 성질을 나타낸다. (예: 환자의 '나이', 질병의 '증상') [3].
- **관계(Relation):** 클래스와 인스턴스 간에 존재하는 연결성을 표현한다 [1, 3, 5]. 관계는 주로 두 가지로 분류된다.
 - **계층적 관계(Taxonomic Relation):** 개념 간의 포함 관계를 나타내는 계층적 구조를 표현한다. 가장 대표적인 예는 "isA" 관계(예: "사람은 동물이다")이다 [3, 5].
 - **비계층적 관계(Non-taxonomic Relation):** 계층적 관계를 제외한 모든 관계를 의미한다. (예: "운동으로 인해 건강해진다"는 "cause"(인과관계)를 사용하여 표현한다) [5].

2.3. 시맨틱 웹의 언어와 기술

시맨틱 웹 시대의 온톨로지는 컴퓨터가 이해할 수 있는 언어를 통해 표현된다.

- **RDF(Resource Description Framework):** 가장 기초적인 온톨로지 언어이며, '주어-술어-목적어(Subject-Predicate-Object)'로 구성된 '트리플(Triple)' 구조를 통해 지식을 표현한다 [8, 9]. 이 트리플은 자원 간의 관계를 그래프 구조로 기술하며, 각 요소에 정형화된 어휘(Vocabulary)를 사용하여 상호운용성을 확보한다 [8]. 이는 데이터 간의 관계를 노드(주어, 목적어)와 에지(술어)로 시각화하여 지식 그래프를 구성하는 데 활용된다 [8, 10].
- **OWL(Web Ontology Language):** RDF 위에 구축된 언어로, 클래스 간의 계층 관계나 인스턴스 내의 논리적 제약 조건 등 더 정밀한 지식 표현을 가능하게 한다 [2, 5, 11]. 복잡하고 논리적인 추론이 필요할 때 주로 사용된다 [2].
- **SPARQL(SPARQL Protocol and RDF Query Language):** RDF로 표현된 지식 그래프를 질의하고 검색하는 표준 언어이다 [12, 13]. 데이터베이스의 SQL SELECT문과 유사하며, 특정 조건에 맞는 데이터를 찾고(WHERE), 합집합(UNION)을 수행하거나, 결과의 개수를 세는(COUNT) 등 다양한 질의 기능을 제공한다 [13, 14].

2.4. 전통적 온톨로지의 한계와 도전 과제

시맨틱 웹 기반의 온톨로지는 강력한 개념적 프레임워크를 제공했음에도 불구하고, 현실적인 적용 과정에서 여러 한계에 직면했다.

- **복잡성과 구축 난이도:** OWL과 같은 복잡한 언어로 온톨로지를 설계하고 표현하는 것은 인간과 기계 모두에게 어려운 과제이다 [15]. 이는 숙련된 지식 공학 전문가를 요구하며, 온톨로지 구축 비용을 높이는 요인이 된다.
- **확장성 및 유지보수의 어려움:** 대규모의 데이터와 끊임없이 변화하는 환경에 온톨로지를 적용할 때 확장성(scalability) 문제가 발생한다. 또한, 현실 세계의 변화에 맞춰 온톨로지를 지속적으로 진화시키는 '온톨로지 노화(ontology aging)'는 해결하기 어려운 과제이다 [15].
- **의미적 모호성 및 문화적 차이:** 언어와 문화에 따라 동일한 개념에 대한 이해가 달라질 수 있으며, 감정적 표현과 같은 비정형 데이터의 미묘한 의미를 온톨로지로 완벽하게 표현하는 데 한계가 있다 [15].
- **지식 표현과 문제 해결 간의 간극:** 전통적 온톨로지는 '지식 표현'과 '추론'에 뛰어난 장점을 가졌지만, 이것이 현실의 복잡한 문제를 직접적으로 해결하는 '운영적 연결고리'로 이어지는 데는 어려움이 있었다. 이는 IBM 왓슨의 사례에서 간접적으로 확인된다. 왓슨은 온톨로지 기반으로 방대한 의료 지식을 학습했음에도 불구하고, 실제 의료 현장의 역동적인 문제를 해결하는 데 큰 성공을 거두지 못하고 결국 사업을 접게 되었다 [16]. 이는 단순히 지식을 표현하는 것만으로는 충분하지 않으며, 그 지식이 실시간으로 변화하는 현실과 상호작용하고, 어떤 '행동'을 유발하며, 그 행동이 다시 지식으로 환류되는 '운영적 순환'이 필요하다는 점을 시사한다.

전통적 온톨로지의 '공유된 개념화(shared conceptualization)'는 주로 기술적 시스템 간의 상호운용성을 위해 동일한 개념 체계를 사용하는 것을 의미했다 [5]. 이는 기술적 차원의 문제 해결에 국한되었으며, 조직의 모든 구성원이 데이터를 기반으로 동일하게 현실을 해석하고 협업하는 '공통의 언어(common language)'라는 개념으로까지 확장되지 못했다. 이러한 한계는 온톨로지 개념이 다음 단계로 진화해야 할 필요성을 제기했다.

III. 팔란티어 온톨로지: 현실 비즈니스의 디지털 트윈

3.1. 기존 데이터 플랫폼의 문제점과 팔란티어의 접근 방식

오늘날 대부분의 기업은 ERP, MES, PLM 등 각 부서 및 기능별로 분리된 시스템(데이터 사일로)으로 인해 데이터 통합과 실시간 협업에 어려움을 겪고 있다 [17, 18]. 이로 인해 데이터 분석 및 의사결정이 파편화되고 비효율적으로 이루어진다.

팔란티어는 이러한 문제를 해결하기 위해 온톨로지를 핵심 개념으로 제시한다. 팔란티어 온톨로지는 기존 시스템을 대체하는 것이 아니라, 그 위에 올라타 분리된 데이터를 연결하는 '추상화 계층'이자 '시맨틱 레이어'로서의 역할을 수행한다 [18, 19]. 이는 단순히 데이터를 저장하는 데이터 베이스와 달리, 비즈니스의 복잡한 현실을 '객체(Object)'로 정의하고 그들 간의 관계를 모형화하여 데이터에 의미와 기능을 부여하는 구조이다 [20]. 팔란티어 온톨로지는 데이터 웨어하우징의

스타 스키마(star schema) 모델링과 유사한 점이 있지만, 단순한 데이터 모델링을 넘어 현실 비즈니스를 디지털로 표현하는 '디지털 트윈(Digital Twin)'의 역할을 수행한다는 점에서 차별점을 갖는다 [18, 21, 22, 23].

3.2. 팔란티어 온톨로지의 3계층 아키텍처

팔란티어 온톨로지는 **의미(Semantic), 동적(Kinetic), 역동적(Dynamic)**이라는 세 가지 상호 연결된 계층으로 구성되어, 비즈니스의 모든 요소를 포괄하는 포괄적인 프레임워크를 제공한다 [24, 25, 26].

- **의미 계층 (Semantic Layer):**

- 온톨로지의 가장 근간을 이루는 계층으로, 비즈니스 도메인의 개념 모델을 정의한다 [26]. 이 계층은 실세계의 개체(오브젝트, Object)와 이들의 관계(링크, Link), 그리고 속성(Property)을 디지털로 표현한다 [4, 23, 24].
- 오브젝트는 인스턴스라고도 불리며, 실세계의 단일 개체나 이벤트를 의미한다 [4]. (예: "직원"이라는 개념에서 "홍길동"은 하나의 오브젝트이다) [4].
- 링크는 두 오브젝트 유형 간의 관계를 정의한다 [4].
- 이 계층은 파운드리 플랫폼에 통합된 데이터셋을 객체 유형, 속성, 링크에 매핑함으로써 조직의 세상을 디지털로 나타낸다 [23, 27].

- **동적 계층 (Kinetic Layer):**

- 의미 계층이 정의한 개념에 현실의 '행동'을 연결하는 계층이다 [26].
- 비즈니스에서 발생하는 다양한 '액션(Action)'과 '함수(Function)'를 온톨로지 객체에 연결하여 실시간으로 비즈니스의 행동을 그래프로 표현한다 [24, 25].
- 이 계층은 AI 기반 운영의 기초를 형성하며, 프로세스 마이닝, 자동화, 그리고 실시간 모니터링 기능을 가능하게 한다 [24, 25].
- 액션은 사용자의 워크플로우를 제어하고 온톨로지 데이터를 수정하는 주요 방법이며, 사용자 변경 사항은 온톨로지에 반영되어 모든 애플리케이션에 실시간으로 적용된다 [4, 28].

- **역동적 계층 (Dynamic Layer):**

- 통합된 데이터와 액션을 기반으로 시뮬레이션 및 의사결정을 지원하는 최상위 계층이다 [24, 25].
- 'AI 기반 의사결정(AI-Guided Decisions)'을 통해 모델을 객체와 액션에 연결하여 최적의 추천을 계산한다 [24, 25].
- '멀티-스텝 시뮬레이션(Multi-step Simulations)'을 통해 다양한 시나리오를 탐색하고, 전략과 운영 간의 연결고리를 만든다 [25, 29].
- '결정 캡처 및 학습(Decision Capture & Learning)' 기능을 통해 사용자가 내린 결정과 그 결과를 새로운 데이터로 기록하고, 이를 AI/ML 모델에 다시 학습시켜 예측 능력을

지속적으로 향상시킨다 [24, 30]. 이 과정은 의사결정의 인과관계(causality)를 증명하는 중요한 역할을 한다 [30].

3.3. 온톨로지 구축 및 운영의 실제

팔란티어 온톨로지는 '온톨로지 관리자(Ontology Manager)'를 통해 구축된다. 이 과정은 기존 데이터 소스를 객체 유형, 속성, 링크에 매핑하는 작업으로, 대부분의 경우 테이블의 행이 하나의 오브젝트(인스턴스)를 이루고, 열이 속성이나 다른 오브젝트와의 링크를 이루는 관계로 매핑된다 [4, 31]. 이렇게 구축된 온톨로지는 다양한 애플리케이션(Object Explorer, Workshop, Quiver 등)을 통해 사용자의 워크플로우에 통합되어 활용된다 [4, 32, 33].

팔란티어 온톨로지는 전통적인 지식 그래프의 개념을 넘어선 '실시간 운영 시스템'이다. 전통적 온톨로지가 지식의 관계를 정적인 '지식 그래프' 형태로 표현하고 SPARQL과 같은 언어로 질의하는 Read-중심 모델에 가깝다면, 팔란티어 온톨로지는 '액션'과 '결정 캡처'라는 동적 요소를 핵심에 포함한다 [24, 25]. 이는 단순히 데이터를 읽는 것을 넘어, 데이터를 쓰고(Write), 현실 세계에 변화를 가하며, 그 결과를 추적하고 학습하는 '운영 레이어(Operational Layer)'로서의 역할을 수행한다. 이 패러다임은 단순히 정보를 찾는 것을 넘어, 실시간으로 문제를 해결하고, 의사결정을 실행하며, 그 결과를 추적하고 학습하는 '실시간 운영 시스템'으로의 진화를 의미한다.

또한 팔란티어는 데이터 사일로 문제 해결을 위해 '사전적 통제'라는 독특한 접근법을 취한다. 기존의 데이터 통합 방식은 종종 '사후적'으로 이루어져, 데이터 파이프라인(예: Airflow)과 데이터 모델(예: DBT)이 각기 다른 스키마와 메타데이터를 개별적으로 관리하며 컨텍스트 단절을 초래한다 [19]. 팔란티어는 이와 달리, '온톨로지를 거치지 않고서는 시스템 운영이 불가능하도록' 플랫폼을 설계함으로써 데이터 활동 전체를 '사전적으로' 통제한다 [19]. 이는 데이터의 생성 시점부터 의미적 통일성을 강제하여, AI 모델이 맥락을 고려한 올바른 판단을 내릴 수 있는 환경을 조성한다 [17, 19].

IV. 심층 비교 분석 및 적용 사례

4.1. 전통적 온톨로지 vs. 팔란티어 온톨로지 비교

전통적 온톨로지와 팔란티어 온톨로지는 공통의 철학적 뿌리에서 출발했지만, 그 목적, 기술 스택, 비즈니스 가치 창출 방식에서 근본적인 차이를 보인다. 팔란티어는 전통적 온톨로지의 '지식 표현' 장점('공유', '재사용', '추론')을 내재화하면서, '운영적'인 요소를 추가하여 비즈니스 가치를 극대화 했다 [4]. 아래 표는 이 두 개념의 주요 차이점을 요약한다.

분류	전통적 온톨로지 (시맨틱 웹)	팔란티어 온톨로지 (파운드리)
----	------------------	------------------

목적	지식의 형식적 표현, 공유 및 재사용 [1, 5]	현실 비즈니스의 디지털 트윈 구축 및 운영 자동화 [20, 34]
핵심 구성 요소	클래스, 인스턴스, 관계, 속성 [5]	오브젝트, 속성, 링크, 액션, 함수 [4]
기술 스택	RDF, OWL, SPARQL [2, 35]	자체 개발 마이크로서비스 아키텍처 (OMS, OSS), 파이썬, 타입스크립트 등 [4, 23]
데이터 모델	트리플 기반의 정적 지식 그래프 [8, 9]	오브젝트 기반의 동적 운영 그래프 [24]
핵심 가치	지식 추론, 의미적 상호운용성 [2, 3]	연결성, 해석 가능성, 운영 AI [36]
주요 제약	확장성, 구축 및 유지보수 난이도 [15]	초기 구축 비용, 폐쇄적 생태계 [37]

4.2. 팔란티어 온톨로지의 실질적 가치 창출 사례

팔란티어 온톨로지는 다양한 산업에서 복잡한 문제를 해결하며 실질적인 가치를 창출하고 있다. 특히 '락인 효과(Lock-in Effect)'를 통한 독점적 생태계 구축은 팔란티어의 중요한 비즈니스 전략이다 [37]. 온톨로지를 통해 고객의 핵심 운영 데이터를 단일 프레임워크로 통합함으로써, 다른 경쟁사 솔루션으로의 전환이 거의 불가능한 수준의 진입 장벽을 구축한다 [37]. 이처럼 팔란티어는 기술적 우위를 넘어, 고객의 비즈니스와 운영 방식을 깊이 이해하는 '전방 배치 엔지니어(Forward-deployed Engineer)' 모델을 통해 온톨로지를 직접 구축하고 유지보수함으로써, 기술적 문제를 인적 자원으로 해결하는 독특한 전략을 구사하고 있다 [22].

아래 표는 주요 산업에서의 구체적인 활용 사례를 요약한다.

산업	해결한 문제	주요 기술 요소	비즈니스 성과
제조/항공	복잡한 생산 공정의 비효율성 [38]	생산 계획, 재고, 인력, 부품 데이터 통합 [38]	에어버스 A350 생산 30% 이상 가속화 [38, 39]
제조/소	공급망 데이터 사일로와 원가 분석의 어려움	7개 이상의 ERP 데이터 통합, 디지털 트윈	SKU(재고 단위)별 원가 및 수익성 파악, 원자재 구매 최적화

비 재	[34, 39]	구축 [34]	[34, 39]
공 공/ 국 방	대규모 범죄 네트워크 분석 및 추적 [37]	고담 플랫폼의 객체 및 관계 시각화 [37]	금융 사기 탐지, 군사 작전 효율 성 증대 [37]
금 융	증권 시장 데이터의 복 잡성 [40]	AI 온톨로지 및 분석 워 크플로 연구 [40]	시장 데이터 기반의 AI 분석 역 량 강화 [40]

V. 결론 및 미래 전망

5.1. 온톨로지 패러다임의 전환

본 보고서의 분석에 따르면, 온톨로지 개념은 '지식의 표현과 추론'이라는 정적이고 이론적인 차원에서 '현실 비즈니스의 운영과 의사결정'이라는 동적이고 실용적인 차원으로 진화했다. 전통적 온톨로지가 지식의 '명세화'에 집중했다면, 팔란티어 온톨로지는 '운영적 요소(액션, 결정)'를 핵심에 포함하여 지식과 행동을 직접적으로 연결하는 '디지털 트윈'으로 구현되었다 [20, 24]. 이는 온톨로지가 단순히 정보를 검색하는 도구를 넘어, 실시간으로 문제를 해결하고, 의사결정을 실행하며, 그 결과를 학습하여 시스템을 지속적으로 개선하는 핵심 인프라가 될 수 있음을 보여준다.

5.2. 데이터 기반 조직을 향한 전략적 시사점

데이터 기반 조직으로의 전환을 모색하는 기업은 단순히 데이터를 모으고 시각화하는 것을 넘어, 데이터에 의미와 맥락을 부여하고, 이를 행동으로 연결하는 시스템을 구축해야 한다. 팔란티어 온톨로지는 이러한 비전을 실현하는 구체적인 청사진을 제시하며, '지속적으로 진화하는 조직의 의사 결정 체계'를 구축하는 핵심적인 도구가 될 수 있다 [17, 36]. 특히 팔란티어의 온톨로지는 비즈니스 프로세스 전반에 걸쳐 데이터의 연결성(Connectivity), 해석 가능성(Interpretability), 규모의 경제(Economies of Scale)를 제공함으로써, 조직이 데이터 자산의 가치를 극대화하고 복잡성을 제어하며 디지털 전환을 가속화할 수 있도록 돋는다 [36].

5.3. 미래 전망

데이터와 AI의 미래에서 온톨로지의 중요성은 더욱 커질 것이다. 인공지능 에이전트와 같은 새로운 데이터 소비자가 등장함에 따라, 온톨로지는 그들이 복잡한 현실 세계를 이해하고 맥락에 맞는 올바른 판단을 내릴 수 있도록 '지식의 표준'을 제시하는 필수적인 인프라로 자리매김할 것이다 [19]. 결국, 온톨로지는 데이터를 단순히 저장하는 공간이 아닌, 조직의 집단 지성이 지속적으로 개선되

고 진화하는 살아있는 플랫폼으로서의 역할을 수행하게 될 것이다.