

# 페블러스 미국 특허 (US 12,481,720 B2) 기술 및 비즈니스 가치 분석 보고서

- 작성일: 2025년 11월 28일
- 기획: (주)페블러스 데이터 커뮤니케이션팀
- 인터랙티브 콘텐츠: <https://blog.pebblous.ai/>

## 1. 특허 개요: 권리의 확보

### 1.1. 서론

본 섹션은 페블러스(Pebblous)가 확보한 미국 특허 US 12,481,720 B2의 전략적 중요성을 분석합니다. 이 특허는 회사의 핵심 기술 자산을 법적으로 보호하는 첫 번째 방어선이자, 경쟁사가 모방할 수 없는 기술적 해자(Technological Moat)를 구축하는 법적 근간입니다. 본 분석은 특허의 법적 권리 범위를 명확히 하고, 이 권리가 어떻게 페블러스의 독점적 시장 지위를 보장하는지를 논증하는 것을 목표로 합니다.

### 1.2. 특허 기본 정보

페블러스가 확보한 핵심 지적 재산권의 주요 정보는 다음과 같습니다.

항목	내용
특허 명	COMPUTING DEVICE THAT PERFORMS A METHOD FOR DIAGNOSING PROPERTIES OF DATA AND A SYSTEM COMPRISING THE COMPUTING DEVICE
특허 번호	US 12,481,720 B2
발행 일	2025년 11월 25일
출원 번호	18/511,617
출원 일	2023년 11월 16일

발명자	Joo Haeng Lee, Jeong Won Lee
양수인	PEBBLOUS INC., Daejeon (KR)
특허 존속 기간	2042년 11월 19일까지 (82일의 기간 연장 포함)

### 1.3. 부여된 권리의 범위

미국 특허청장(Director of the United States Patent and Trademark Office)은 페블러스에 해당 발명에 대해 미국 내에서 타인의 제조(making), 사용(using), 판매 제안(offering for sale), 또는 판매(selling) 행위를 배제할 수 있는 독점적 권리를 부여했습니다.

비즈니스 관점에서 이는 다음과 같은 강력한 의미를 가집니다.

- 시장 독점권: 경쟁사는 페블러스의 허가 없이 특허에 명시된 "데이터의 속성을 진단하고 개선하는 방법"을 상업적으로 구현하거나, 해당 기술이 적용된 제품 또는 서비스를 제공할 수 없습니다.
- 경쟁 우위 확보: 이 독점권은 페블러스의 핵심 솔루션인 '데이터클리닉(DataClinic)'의 기술적 근간을 보호함으로써, 경쟁사들이 유사한 기능으로 시장에 진입하는 것을 효과적으로 차단합니다.
- 가치 창출 기반: 이 권리는 페블러스가 라이선스 계약, 기술 제휴 등 다양한 비즈니스 모델을 통해 자사의 지적 재산을 수익화할 수 있는 법적 토대가 됩니다.

### 1.4. 결론 및 전환

결론적으로, US 12,481,720 B2 특허는 단순한 기술 문서를 넘어, 2042년 11월 19일까지 페블러스의 비즈니스 활동을 보호하는 강력하고 장기적인 법적 방패입니다. 이는 페블러스가 안정적으로 시장을 선도하고 기술 혁신에 집중할 수 있는 기반을 제공합니다. 다음 섹션에서는 이 특허가 구체적으로 보호하는 핵심 기술의 원리와 작동 방식에 대해 심층적으로 분석하겠습니다.

## 2. 핵심 특허 기술 분석: '데이터 지도'를 통한 진단과 개선

### 2.1. 서론

본 섹션의 목적은 특허가 보호하는 핵심 기술의 작동 원리를 명료하게 분석하는 것입니다. 이 기술은 AI 개발의 가장 큰 병목 현상인 '데이터 품질' 문제를 해결하기 위해, 눈에 보이지 않는 데이터의 특성을 시각적으로 진단하고 과학적으로 개선하는 독창적인 방법론을 제시합니다. 복잡한 기술 원리를 비전문가도 쉽게 이해할 수 있도록 '데이터 지도'라는 비유를 사용하여 단계별로 설명하겠습니다.

## 2.2. 1단계: 진단 - 데이터 이미징(Data Imaging)

특허 기술의 첫 단계는 AI 학습 데이터를 진단 가능한 형태로 변환하는 '데이터 이미징'입니다. AI 학습 데이터를 고차원의 가상 공간인 '임베딩 공간(Embedding Space)' 위에 개별 '데이터 포인트(Data Point)'로 변환하여 배치합니다. 이 과정의 핵심은 데이터의 추상적인 '의미론적 유사성'(예: 비슷한 종류의 새 이미지들)을 임베딩 공간 내의 물리적인 '근접성'으로 매핑하는 것입니다. 결과적으로, 이 임베딩 공간은 데이터의 전체적인 구조와 관계를 한눈에 파악할 수 있는 일종의 '데이터 지도' 역할을 합니다. 이 지도 위에서 데이터 포인트의 물리적 \*\*밀도(density)\*\*와 \*\*분포(distribution)\*\*는 데이터셋의 건강 상태를 나타내는 핵심 지표가 됩니다.

## 2.3. 2단계: 개선 - 분포 조정 및 합성 데이터 생성

진단된 '데이터 지도'의 문제를 해결하는 두 번째 단계는 데이터의 분포를 과학적으로 '개선'하는 과정입니다. 진단된 데이터 지도의 분포(속성) 문제를 해결하기 위해, 특허 기술은 데이터 포인트의 위치를 조정하거나 추가하여 '수정된 데이터 포인트 세트(Modified First Data Point Set)'를 생성합니다.

진단된 '데이터 지도'에서 데이터가 부족하여 비어있는 영역(저밀도 갭)이 발견되면, 특허 기술은 바로 그 위치에 \*\*원본 데이터셋에는 존재하지 않았던 새로운 데이터 포인트\*\*를 정밀하게 생성하여 추가합니다. 이것이 바로 데이터셋의 약점을 보강하는 '합성 데이터(Synthetic Data)' 생성 기술의 핵심 원리이며, 본 특허가 보호하는 독점적 권리입니다.

## 2.4. 3단계: 보고 - 시각화 및 진단 리포트

마지막으로, 진단 및 개선 결과를 사용자에게 명확하게 전달하는 단계입니다. 특허는 원본 데이터 지도를 \*\*데이터 이미지(Image of Data, IOD)\*\*로, 개선된 데이터 지도를 \*\*수정된 데이터 이미지(Modified Image of Data, MIOD)\*\*로 시각화하여 제공하는 방법을 명시합니다. 이 시각적 자료들은 데이터의 속성, 품질, 개선 제안 등을 포함하는 최종 '진단 보고서(Diagnostic Information)' 형태로 통합되어 사용자에게 제공됩니다.

## 2.5. 결론 및 전환

요약하자면, 이 특허 기술은 데이터의 문제를 눈으로 보고(진단), 정량적으로 분석하며(밀도/분포 측정), 과학적으로 개선(분포 조정 및 합성 데이터 생성)하는 통합적인 방법론을 법적으로 보호합니다. 이는 주관적인 경험에 의존하던 데이터 품질 관리를 객관적인 공학의 영역으로 끌어올린 혁신입니다. 다음 섹션에서는 이 핵심 기술이 페블러스의 실제 제품군에 어떻게 적용되어 사업적 가치를 창출하는지 구체적으로 분석하겠습니다.

---

## 3. 페블러스 제품 및 비전과의 연계성 분석

### 3.1. 서론

본 섹션의 분석 목표는 앞서 분석한 추상적인 특허 기술이 어떻게 페블러스의 핵심 제품 및 장기 비전의

구체적인 기술 엔진으로 작동하는지 명확히 연결하여, 특허의 실질적인 사업 가치를 입증하는 것입니다. 이 특허는 단순히 분리된 기술이 아닌, 페블러스의 모든 솔루션을 관통하는 핵심 DNA 역할을 수행합니다.

3.2. 특허 기술과 핵심 솔루션 매핑

특허 기술은 페블러스의 현재와 미래 제품 포트폴리오 전반에 걸쳐 기술적, 법적 토대를 제공합니다.

페블러스 요소	특허와의 대응 관계	상세 설명
데이터클리닉 (DataClinic)	핵심 엔진 및 방법론	특허의 명칭 자체가 <b>***데이터의 속성을 진단하는 방법***</b> 을 명시하며, 데이터클리닉이 AI 학습 데이터의 품질을 종합적으로 진단하고 보정하는 핵심 기술 그 자체임을 보호합니다. 특허의 밀도 및 분포 분석 기술은 데이터클리닉의 진단 알고리즘의 근간입니다.
페블로스코프 (PebbloScope)	데이터 시각화 및 커뮤니케이션	특허가 정의하는 '데이터 이미징(IOD/MIOD)' 기술은 페블로스코프의 직접적인 기술적 토대입니다. 고차원 데이터를 3D 공간으로 변환하여 분포적 특성을 시각적으로 탐색하게 하는 페블로스코프의 기능은 특허의 시각화 및 보고 방법을 상용화한 것입니다.
데이터 다이어트 (Data Diet)	중복 및 유사 데이터 제거	이는 특허가 명시한 '분포 속성 조정' 및 '경계 밖 데이터 포인트 삭제' 기능에 직접적으로 대응됩니다. 특허 기술은 과밀집 클러스터를 식별하고 제거함으로써, 데이터 다이어트의 기술적 정당성과 독점성을 법적으로 보장합니다.
데이터 벌크업 (Data Bulk-up)	정밀 타겟팅 합성 데이터 생성	이는 특허의 핵심 청구항인 <b>***원본에 없던 새로운 데이터 포인트를 추가***</b> 하는 메커니즘을 상업적으로 구현한 것입니다. 이로써 페블러스는 경쟁사가 모방할 수 없는 정밀 타겟팅 합성 데이터 생성 능력을 독점적으로 보유하게 됩니다.
AADS (Agentic AI Data Scientist)	자율 운영 시스템의 핵심	AADS는 데이터 업무의 완전한 자율 운영 체계를 지향합니다. 이 특허 기술은 AADS라는 자율 에이전트가 데이터 품질 문제(밀도, 분포 불균형 등)를 스스로 진단하고 개선하는 데 사용하는 핵심 '도구(tool)' 또는 '함수(function)' 역할을 수행합니다.

### 3.3. 결론 및 전환

이처럼 US 12,481,720 B2 특허는 페블러스의 현재 제품 포트폴리오의 기술적 정당성을 확보할 뿐만 아니라, 미래 비전인 AADS의 근간까지 아우르는 핵심 자산입니다. 이는 페블러스의 기술 로드맵 전체를 보호하는 강력한 방패 역할을 합니다. 다음 섹션에서는 이 독점 기술이 국제 표준과 어떻게 부합하여 더 큰 신뢰성과 경쟁력을 갖게 되는지 분석하겠습니다.

## 4. 국제 표준(ISO/IEC 5259-2)과의 기술적 매핑

### 4.1. 서론

본 섹션은 페블러스 특허 기술의 또 다른 전략적 가치를 분석합니다. AI 데이터 품질에 대한 국제 표준인 ISO/IEC 5259-2는 AI 시스템이 갖춰야 할 데이터 품질 특성, 즉 '무엇을(What)' 측정해야 하는지에 대한 이론적 프레임워크를 제공합니다. 그러나 이 표준은 '어떻게(How)' 그 추상적인 품질 특성들을, 특히 비정형 데이터에 대해 정량적으로 측정할 것인지에 대한 구체적인 기술적 해답은 제시하지 못했습니다. 페블러스의 특허 기술은 바로 이 '잃어버린 연결고리(lost connection)'를 해결하는 독점적인 기술적 구현체로서, 표준의 이론을 현실의 공학으로 전환시키는 핵심적인 역할을 합니다. 즉, 페블러스의 특허 기술은 ISO 표준이 제기한 '무엇을(What)'이라는 철학적 질문에 '어떻게(How)'라는 공학적 해답을 제공하는, 현재까지 유일하게 입증된 기술적 구현체입니다.

### 4.2. ISO 5259-2 추가 품질 특성과의 상세 매핑

페블러스 특허 기술은 ISO/IEC 5259-2가 정의하는 AI/ML 특화 추가 품질 특성들을 다음과 같이 정량적으로 측정하고 개선하는 방법을 제공합니다.

ISO 품질 특성	페블러스 특허 기술의 역할 및 매핑	상세 설명
유사성 (Similarity,	임베딩 공간의 밀도	<b>**Sim-ML-1 (샘플 유사성)**</b> 은 데이터셋 내 중복/유사 샘플을 측정합니다. 특허 기술은 데이터 이미징 후 임베딩 공간의 밀도를 측정하여, 과밀집 클러스터를 시각적으로 식별함으로써

Sim-ML-1/3)	측정 및 내재 적 차원 산출	유사성 위반을 객관적으로 증명합니다. 또한 **Sim-ML-3 (샘플 독립성)**은 내재적 차원 산출을 통해 데이터의 본질적인 복잡도와 정보 중복성을 측정하여 달성합니다.
대표성 (Representativeness, Rep-ML-1)	매니 폴드 형상 분석 및 벌크 업	**Rep-ML-1 (대표성)**은 데이터셋이 실제 환경을 얼마나 잘 반영하는지 평가합니다. 특허 기술은 매니폴드 형상을 분석하여 데이터가 부족한 **저밀도 영역(gap)**을 식별함으로써 대표성 부족을 증명합니다. 나아가 '수정된 데이터 포인트 추가' (즉, 벌크업)를 통해 이 문제를 개선하는 방법을 제공합니다.
다양성 (Diversity, Div-ML-1/2/3)	매니 폴드 크기 및 거리- 밀도 측정	**Div-ML (다양성)**은 데이터가 얼마나 다양한 시나리오를 포함하는지 평가합니다. 특허 기술은 매니폴드의 크기와 거리-밀도 측정을 통해 데이터가 다양한 특징 공간을 얼마나 포함하는지를 정량적으로 판단하는 근거를 제공합니다.
효율성 (Efficiency, Eff-ML-3)	밀도 측정 기반 데이 터 다이 어트	**Eff-ML-3 (공간 낭비 위험)**은 데이터 처리의 효율성을 측정합니다. 특허의 밀도 측정 기술로 식별된 중복/유사 데이터 (과밀집 클러스터)에 대한 데이터 다이어트 처방은 불필요한 저장 공간과 컴퓨팅 자원 낭비를 줄여 GPU 효율을 향상시키는, 표준의 요구사항을 직접적으로 이행하는 활동입니다.

### 4.3. 결론 및 전환

결론적으로, 페블러스의 특허 기술은 ISO 표준의 추상적 요구사항을 측정 가능하고, 운영 가능하며, 개선 가능한 공학적 솔루션으로 변환합니다. 이는 데이터 품질 관리를 주관적 판단의 영역에서 객관적 과학의 영역으로 끌어올린 핵심적인 기여입니다. 마지막 섹션에서는 이러한 기술적, 표준적 우위가 어떻게 페블러스의 최종적인 비즈니스 경쟁력으로 직결되는지 종합적으로 분석하겠습니다.

## 5. 종합 가치 분석: 비즈니스 경쟁력 강화 전략

## 5.1. 서론

본 보고서의 결론에 해당하는 본 섹션은 앞서 분석한 특허의 기술적 내용, 제품 연계성, 표준 부합성이 어떻게 결합하여 페블러스의 지속 가능한 비즈니스 경쟁력을 구축하는지 종합적으로 평가합니다. 이 특허는 단순한 기술 보호를 넘어, 페블러스의 시장 지배력을 공고히 하는 다층적이고 방어 불가능한 가치 구조를 형성합니다.

## 5.2. 기술적 가치 (Technological Value)

- 강력한 지적 재산권(IP) 보호 2042년까지 보장되는 독점적 권리는 데이터 품질 진단 및 개선이라는 핵심 기술 영역에서 경쟁사의 진입을 원천적으로 차단하는 강력한 기술적 해자를 형성합니다. 이는 페블러스가 장기적인 관점에서 안정적인 연구개발과 시장 선도를 지속할 수 있는 법적 기반이 됩니다.
- 통합적 데이터 품질 관리 사이클 지원 이 특허는 진단(Diagnosis), 개선(Enhancement), 생성(Generation)을 아우르는 엔드투엔드(End-to-End) 데이터 품질 관리 사이클 전체를 지원하는 기술적 기반입니다. 이는 개별 기능만을 제공하는 경쟁 솔루션 대비 압도적인 통합 가치를 제공합니다.
- 멀티모달리티 및 Physical AI 시장 선점 특허는 이미지, 텍스트 등 다양한 모달리티를 처리할 수 있는 기반을 제공하여, 고도의 신뢰성과 정확성을 요구하는 고부가가치 Physical AI 시장을 선점하는데 핵심적인 기술적 우위를 부여합니다.

## 5.3. 비즈니스 가치 (Business Value)

- 규제 시장에서의 독점적 경쟁 우위 페블러스의 특허 기술은 ISO/IEC 5259-2가 요구하는 대표성, 유사성 등의 추상적 품질을 정량적으로 측정하고 개선하는 독점적 방법을 제공합니다. 이를 통해 생성된 데이터클리닉의 진단 보고서는 EU AI Act 및 ISO 42001이 요구하는 '\*\*감사 가능한 증적 자료(audit trail)\*\*'의 기술적 전제 조건을 완벽히 충족시킨다. 이는 규제 준수 의무가 있는 기업 고객에게 법적 리스크를 제거하는 결정적인 경쟁 우위로 작용합니다.
- 측정 가능한 고객 ROI 증명 특허 기술은 고객에게 명확한 재무적 투자 수익(ROI)을 제공하는 직접적인 근거가 됩니다. '데이터 다이어트'를 통한 GPU 및 클라우드 비용의 정량적 절감, '데이터 벌크업'을 통한 AI 모델 성능의 가시적 향상 등은 고객이 페블러스 솔루션 도입의 경제적 타당성을 쉽게 입증할 수 있게 합니다.

---

## 6. 결론: AI-Ready 데이터 플랫폼의 초석

미국 특허 12,481,720 B2는 단순히 하나의 기술을 보호하는 것을 넘어, 데이터 품질 진단, 개선, 시각화, 그리고 자율화에 이르는 페블러스의 모든 비즈니스 전략을 지탱하는 핵심 초석(Cornerstone)입니다. 이 특허는 데이터의 보이지 않는 문제를 '데이터 지도'라는 혁신적인 방법으로 가시화하고, 국제 표준이 요구하는 추상적 품질을 측정 가능한 공학으로 변환하며, 최종적으로 고객에게 감사 가능한 신뢰와 측정 가능한 ROI를 제공합니다. 이 강력한 지적 재산을 기반으로 페블러스는 경쟁사가 넘볼 수 없는 기술적 해자를 구축하고, 'AI-Ready 데이터 플랫폼 사업자'로서 미래 데이터 시장을 선도해 나갈 것임을 확신합니다.

# 참고문헌

---

## 1. 미국 특허 문헌 (U.S. Patent Documents)

- US 11,297,083 B1 (Kuppa et al., 2022) - Identifying and protecting against an attack against an anomaly detector machine learning classifier
- US 11,568,245 B2 (Chang et al., 2023) - Method and system for providing synthetic data to a machine learning classifier
- US 11,967,308 B2 (Lee et al., 2024) - Method and apparatus for processing data for machine learning model
- US 2017/0236069 A1 (Min, 2017) - Method for data augmentation using generative adversarial networks
- US 2020/0150235 A1 (Beijbom et al., 2020) - Sensor simulation using domain adaptation
- US 2021/0093210 A1 (Sinha et al., 2021) - Physiological signal processing using manifold learning
- US 2022/0343139 A1 (Passban et al., 2022) - Method for generating synthetic data for machine learning
- US 2022/0382976 A1 (Shin et al., 2022) - Data augmentation method and apparatus
- US 2022/0383570 A1 (Ling et al., 2022) - Image generation method and apparatus
- US 2023/0107415 A1 (Banerjee et al., 2023) - System for data synthesis and method thereof

## 2. 해외 특허 문헌 (Foreign Patent Documents)

- 한국 특허청(KR) 출원 문서입니다.
- KR 10-2019-0056009 A (2019) - 인공지능 학습 데이터 생성 방법
- KR 10-2022-0011979 A (2022) - 인공지능 모델을 위한 데이터 처리 시스템
- KR 10-2022-0102012 A (2022) - 합성 데이터 생성 및 멤버십 정보 제거
- KR 10-2022-0159213 A (2022) - 그래프 신경망(GNN) 잔차 연결 학습 방법

## 3. 기타 참고 문헌 (Other Publications)

- 데이터 매니폴드(Manifold), 합성(Synthesis), 이상 탐지(Anomaly Detection) 관련 주요 논문입니다.
- Manifold-based synthetic oversampling with manifold conformance estimation
  - Author: Bellinger et al. (2018)



- Look, Cast and Mold: Learning 3D Shape Manifold via Single-view Synthetic Data
  - Author: Feng et al. (2021)
  - Note: 특허에는 'Look, Evolve and Mold'로 기재되어 있으나, 동일 논문임.
- Incremental nonlinear dimensionality reduction by manifold learning
  - Author: Law and Jain (2006)
- Decoding brain states on the intrinsic manifold of human brain dynamics
  - Author: Rué-Queralt et al. (2021)
- Mathematical and Intelligent Techniques for Data Analytics in Science and Engineering
  - Author: Hao et al. (2018)
- Dataset Modification To Improve Machine Learning Algorithm Performance And Speed
  - Author: Ahmed, O. (2014)
  - Type: Master's Thesis (University of Houston)
- Sparse Subspace Denoising for Image Manifolds
  - Author: Wang et al. (2013)
- Deep Manifold Embedding for Hyperspectral Image Classification
  - Author: Gong et al. (2022)
- Manifolds for Unsupervised Visual Anomaly Detection
  - Author: Naud et al. (2020)

---

# Pebblous

Pebblous Makes Data Tangible

[contact@pebblous.ai](mailto:contact@pebblous.ai)