

AI 시대 지능형 홈 에너지 빅데이터 활용 및 마이데이터 보안 기술 동향

Trends in Energy Big Data Utilization and MyData Security Technology for Intelligent Homes in the AI Era

구태연 (T.Y. Ku, kutai@etri.re.kr)
박완기 (W.K. Park, wkpark@etri.re.kr)

에너지ICT연구실 책임연구원/기술총괄
에너지ICT연구실 책임연구원/실장

ABSTRACT

Recently, the establishment of intelligent home services and their supporting energy big-data platforms, which optimize residential energy consumption to address climate change and achieve carbon neutrality, has emerged as a key global trend. To resolve privacy violations arising from high-frequency energy data collection, research is being conducted on the application of energy-aware differential privacy (DP) models, 4.5th-generation fully homomorphic encryption (CKKS), and top-tier security systems based on hybrid edge cloud (HEC) architectures. Furthermore, along with enabling a new service ecosystem through open APIs and full-stack data distribution portals, lightweight blockchain-based data sovereignty technology has become a central topic in protecting individual rights in the AI era. This study provided an in-depth analysis of the latest industry trends and future development directions of intelligent home energy big-data platforms and MyData security technologies.

KEYWORDS AI, Big Data, Blockchain, Differential Privacy, HEC, OpenAPI, 경량 블록체인, 데이터 유통 플랫폼, 데이터 주권, 융합 보안, 지능형 홈

I. 서론

최근 기후변화 대응과 탄소중립 실현은 전 세계적인 시대 과제로 자리 잡았으며, 이를 달성하기 위해 국가 전체 에너지 수요의 근본적인 혁신이 요구되고 있다. 특히 국내 건물 에너지 사용량의 약 60%

를 차지하는 주거 영역(공동주택 및 단독주택)에서의 에너지 소비 최적화는 탄소 배출 저감을 위한 필수 선결 과제이다. 과거의 주거 환경이 단순한 쉼터의 역할에 머물렀다면, 최근에는 사물인터넷과 인공지능 기술이 융합되고 글로벌 스마트홈 연동 표준인 매터(Matter)가 급격히 확산함에 따라, 주거 공간은

* DOI: <https://doi.org/10.22648/ETRI.2026.J.410403>

* 이 논문은 2026년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임[No. RS-2025-02217507, (세부3) 지능형 홈 엣지-클라우드 하이브리드 개인정보관리 강화 기술개발(한국전자통신연구원)].

단순 제어 중심에서 벗어나 거주자의 생활 패턴을 인지하고 맞춤형 서비스를 제공하는 지능형 홈으로 진화하고 있다. 이 과정에서 전기, 가스, 수도, 난방, 온수 등 5종에 달하는 복합 에너지 데이터가 실시간으로 쏟아지는, 이른바 데이터 폭발 현상이 발생하고 있다.

본고는 이러한 시대적 요구와 기술적 당위성을 바탕으로, 안전하고 상호운용성이 보장된 지능형 홈 에너지 빅데이터 유통 플랫폼 및 융합 보안 체계 조성 방안에 대해 살펴보고자 한다. 이를 위해 II장에서는 에너지 디지털 전환의 기술적 배경 및 필요성을 짚어보고, III장에서는 데이터 품질 확보 및 시스템 보안 동향에 대해 알아본다. IV장과 V장에서는 OpenAPI 기반의 유통 마켓 구축 및 인공지능 효율화 서비스 실증 사례를 논의하며, VI장에서는 거주자의 프라이버시 보호를 위한 차분 프라이버시(DP) 및 블록체인 기반 데이터 주권 기술 동향을 심도 있게 다룬다. 마지막으로 VII장에서는 본문의 내용을 종합하며 결론을 맺는다.

II. 기술 배경 및 필요성

1. 주택 에너지 데이터의 가치와 정책적 배경

지능형 홈의 5종 에너지원 중에서도 열에너지(난방 및 온수)는 지능형 홈 서비스의 품질과 거주자의 열적 쾌적성, 그리고 삶의 질을 결정짓는 중추적인 데이터 자원이다. 그럼에도 불구하고 국가 주도로 2,250만 호 규모의 대규모 스마트계량기(AMI) 보급이 우선적으로 추진되었던 전력 분야와 비교할 때, 열에너지 분야의 정보 집적 및 디지털 전환 속도는 상대적으로 매우 더디게 진행됐다. 전력 분야의 경우 한국전력이 에너지마켓플레이스(EN:TER)를 통해 160만 명 이상의 사용자를 확보하고 제3자

(3rd Party)에게 실시간 데이터를 개방하는 생태계를 이미 구축한 반면, 열에너지는 데이터 수집조차 파편화되어 있는 실정이다[1-3]. 이러한 산업 간 데이터 격차를 타개하고 제4차 산업혁명의 기반을 다지기 위해 정부는 데이터·AI 경제 활성화 계획을 발표하였다. 본 계획을 통해 2023년까지 약 10조 원을 투입해 데이터 시장 규모를 30조 원으로 키우고 10개의 AI 유니콘 기업을 육성하는 등, 사람 중심의 플랫폼 경제 실현을 위한 공격적인 정책을 추진하고 있다. 또한, 공공기관의 실시간 전력량을 확인하는 공공기관 그린버튼 사업을 민간 기업과 가정으로 확대하여 데이터 기반의 에너지 소비 효율화를 촉진하려는 움직임도 활발하다[4].

2. 지능형 홈 환경의 에너지 데이터 서비스 혁신의 필요성

에너지 분야에서 빅데이터 기반의 플랫폼 구축과 지능화 서비스 혁신이 시급히 요구되는 이유는 크게 4가지 핵심 요인으로 분석된다.

첫째, 기존 데이터 관리의 기술적 한계이다. 기존 집단에너지사업자 및 공동주택 관리사무소는 에너지 검침 데이터를 윈도우 기반의 폐쇄적인 데이터 베이스(MS SQL, Access 등)나 비표준화된 포맷으로 개별 관리하고 있다. 이러한 단순 저장 위주의 시스템은 데이터의 실시간성과 무결성을 보장하지 못하며, 결측치나 이상치가 발생해도 즉각적인 보정이 불가능하여 전체의 지능화에 심각한 장애 요인으로 작용하고 있다.

둘째, 고도화된 융합 서비스 혁신 요구의 폭발적 증가이다. 지능형 홈 환경의 수용가는 단순히 요금 정산을 위한 사용량 모니터링 수준을 넘어선 혁신을 원한다. 이들은 인공지능 분석을 통한 미터기 고장 및 오작동의 원격 진단, 밸브 이상으로 인한 난방

비 폭탄 방지, 시간대별/계절별 소비 패턴 분석, 그리고 이웃 단지와의 사용량 벤치마킹 등 경제성과 편의성을 동시에 만족시키는 고도화된 부가 서비스를 적극적으로 요구하고 있다.

셋째, 산업 연계 및 데이터 유통 생태계의 부재이다. 현재 집단에너지사업자 간 또는 외부 분석가, 서비스 개발자 등 제3자 기관과의 데이터 교류 인프라가 극히 제한적이다. 데이터가 유통될 수 있는 개방형 API(OpenAPI) 포털이나 거래 마켓이 활성화되지 않아 계절별·시간대별 차등 요금제(TOU) 유도나 잉여 열에너지 축열조 저장 등의 최적 스케줄링을 통한 신규 비즈니스 창출 기회가 원천적으로 차단되어 있다.

넷째, 정책 지원 및 표준화 확립의 시급성이다. 에너지 데이터의 체계적인 수집은 제로에너지빌딩 인증제나 그린리모델링과 같은 국가 탄소중립 실현 및 건물 에너지 효율화 정책 추진 과정에서 핵심 인프라로 작용한다. 이를 투명하게 뒷받침하기 위해서는 이기종 검침 인프라를 아우를 수 있는 플랫폼 요구 사항, 참조 구조, 데이터 모델 및 인터페이스에 대한 공통된 기술 규격(단체 표준 등) 확립이 필수적이다.

3. 개인정보 보호 데이터 플랫폼 현황

국내외 기술 격차와 프라이버시 이슈를 살펴보면 독자적인 플랫폼 확보의 당위성은 더욱 명확해진다. 현재 국내 빅데이터 처리 및 유통 분야의 기술 수준은 세계 최고 수준인 미국 대비 약 79.5%(기술 격차 1.35년) 수준에 머물러 있다. 미국은 일찍이 연방정부 차원의 2억 달러 규모 빅데이터 연구 개발 이니셔티브를 추진하였으며, 글로벌 기업들이 주축이 된 ODPI(Open Data Platform Initiative)를 통해 개방형 데이터 유통 생태계를 선점하였다. 반면, 국내 시장은 해외 기업의 솔루션을 단순 응용하거나 오픈

소스 기반의 소규모 시스템 구축에 머물러 해외 기술 종속 우려가 큰 실정이다.

더욱이 지능형 홈에서 수집되는 1시간 단위의 고빈도 에너지 데이터는 거주자의 기상 시각, 샤워 시간, 외출 및 귀가 여부 등 매우 민감한 사생활을 직접 노출할 수 있는 심각한 프라이버시 침해 위험을 내포하고 있다. 따라서 단순한 중앙집중형 클라우드 데이터 수집을 넘어, 엣지 환경에서 개인정보 유출을 자율적으로 검증하고 실시간 차폐 및 감사 추적을 수행하는 마이데이터 및 데이터 주권 보안 기술의 융합이 절실하다.

III. 지능형 홈 에너지 빅데이터 보안 및 프라이버시 보호 동향

앞서 언급한 바와 같이 고빈도 데이터가 악의적으로 유출될 경우, 거주자의 외출 여부를 파악한 범 죄에 악용되거나 무단 프로파일링의 대상이 될 수 있다. 이에 따라 최신 빅데이터 플랫폼들은 방대한 센서 데이터가 외부 클라우드에 일방적으로 집중될 경우 발생할 수 있는 대규모 프라이버시 유출을 막기 위해, 데이터 처리의 무게 중심을 분산시키는 하이브리드 엣지-클라우드 아키텍처를 전면적으로 도입하는 추세이다. 데이터 처리가 분산되어 민감한 원시 데이터는 로컬 엣지에서 우선 가공 및 필터링되므로 외부 전송을 원천 차단할 수 있다. 이와 함께 데이터의 통계적 활용성과 개인정보 보호라는 두 가지 상충하는 목표의 균형을 달성하기 위해 강력한 프라이버시 보호 기술이 접목되고 있다.

1. 추론위험인자통제및차분프라이버시 모델 도입

연속적인 시계열 데이터로부터 개인의 특정 행

위를 유추해 내는 추론 위험을 원천 차단하기 위해, 최근 산업계에서는 개별 세대의 원시 데이터(Raw Data)를 그대로 노출하는 대신 시간적·공간적으로 묶은 집계(Aggregation) 데이터를 활용하거나, 수학적으로 프라이버시를 보장하는 차분 프라이버시(DP: Differential Privacy) 모델을 적극 적용하는 사례가 증가하고 있다[5]. 차분 프라이버시는 데이터 연산 및 질의 과정에 라플라스 또는 가우시안 분포에 기반한 정교한 임의의 노이즈를 의도적으로 삽입하는 기술이다. 이를 통해 데이터베이스에 특정 개인의 데이터가 포함되거나 제외되더라도 전체 통계 분포의 변화량이 엡실론(ϵ)이라는 극히 작은 파라미터 한계치 내로 통제되도록 보장함으로써, 공격자가 어떠한 외부 보조 지식을 동원하더라도 특정 개인의 존재나 민감 속성을 역추산할 수 없도록 강력한 보호막을 제공한다. 더불어, 각 엡지 디바이스의 데이터 기여도와 민감도에 따라 프라이버시 예산(Privacy Budget, ϵ)을 동적으로 탄력 할당하는 기법과 데이터 불균형 환경에서도 연합학습의 수렴 속도를 보장하는 FedProx 알고리즘이 새롭게 결합되어 모델 학습 성능을 극도로 고도화하고 있다[6].

2. 고도화된 비식별 프로세스 및 4.5세대 완전동형암호 기반 보호

수집된 정보가 오픈 데이터 마켓에서 외부 정보와 결합하여 특정 개인을 찾아내는 재식별을 막기 위해, 단순 가명 처리를 넘어 개별화, 재연결, 추론이라는 3대 프라이버시 위협 인자를 상시 모니터링하는 시스템이 대두되고 있다. 이와 함께 특정 개인이 식별될 확률을 $1/k$ 이하로 낮추는 k -익명성, l -다양성, t -접근성 등의 정량적 보호 지표를 복합 적용하여 비식별 조치의 적정성을 평가한다. 나아가, 데이터가 엡지에서 클라우드로 전송되어 분석되는 전

과정에서 유출을 원천 방지하기 위해 완전동형암호(FHE: Fully Homomorphic Encryption)의 도입이 본격화되고 있다. 특히 실수 연산과 근사 연산을 지원하여 암호화된 상태 그대로 로지스틱 회귀나 서포트 벡터 머신(SVM)과 같은 머신러닝 학습 및 추론이 가능한 4.5세대 동형암호 기술(CKKS 등)이 도입되었다. 이는 기존 기술 대비 성능이 대폭 개선되어 스마트폰 모바일 환경에서도 실시간 처리가 가능한 수준으로 발전하며 지능형 홈 데이터 분석의 핵심으로 부상하고 있다[7].

더불어, 다가오는 양자 컴퓨팅 시대의 암호 해독 위협에 선제적으로 대응하기 위해 완전동형암호와 결합 가능한 양자내성암호(PQC: Post-Quantum Cryptography) 알고리즘을 엡지 디바이스 통신에 경량화하여 적용하는 연구가 최신 보안 트렌드로 급부상하고 있다.

3. 글로벌 연동 인증 및 다계층 시스템 보안 체계

네트워크 및 시스템 보안 측면에서는 데이터 생성부터 저장에 이르는 전 구간(E2E) 방어 체계가 필수 요건이다. 최근에는 지능형 홈 기기들의 파편화를 막고 강력한 상호운용성을 보장하는 글로벌 스마트홈 연동 표준인 매터 기반의 기기 인증 체계가 적극 활용된다. 기기 생성 시 발급되는 DAC(Device Attestation Certificate)와 홈 네트워크 연동 시 발급되는 NOC(Node Operational Certificate) 기반의 X.509 인증서를 통해 신뢰할 수 있는 기기 간의 안전한 통신만을 허용한다. 데이터 전송 구간에서는 IPSec 기반 VPN 터널링과 SSL/TLS(HTTPS, HTTP/2) 암호화 채널이 의무적으로 적용되며, 이용자 접근 구간에서는 고유 API Key 기반의 강력한 인증과 역할 기반 접근 제어(RBAC)가 활용된다. 서버 구간에서는 API

호출 로그와 윈도우 보안 이벤트를 실시간 분석해 이상 징후 발생 시 자동 통보하는 지능형 관제 체계를 갖추며, 핵심 하드웨어 및 스마트카드에는 국제 보안 평가 공통 기준인 CC(Common Criteria) EAL7(국제용) 및 EAL4(국내용) 수준의 최고 등급 인증을 지향하여 플랫폼 신뢰도를 극대화하고 있다[8]. 이는 내부 네트워크망조차 기본적으로 신뢰하지 않고 모든 자원 접근을 지속적으로 검증하는 제로 트러스트 아키텍처(ZTA: Zero Trust Architecture)의 핵심 기반 체계로 작용한다[9,10]. 특히, 기기 간의 단순 인증을 넘어 거주자의 프라이버시 권리 보장을 정량화하기 위해 지능형 홈 프라이버시 지수(HPI: Home Privacy Index) 산출 모델이 새롭게 도입되고 있다. 이 지수는 $HPI=wAA+wBB+wCC+wDD+wEE$ 의 수식으로 정의되며, 데이터 수집 통제성(A), 이용 투명성(B), 저장·보안성(C), 제공·연계 통제성(D), 사용자 인식·제어성(E)의 5대 속성을 표준화하여 평가한다.

또한, 매터 표준과 더불어 W3C 표준 기반의 분산 식별자(DID: Decentralized Identifier) 및 검증 가능한 크리덴셜(VC: Verifiable Credentials)이 지능형 홈 옻지 아키텍처에 융합되고 있다[11]. DID 주체와 상호작용을 하는 모든 메타데이터는 JSON-LD 형태의 DID 문서(Document)로 구성되며, 이 문서 내의 verificationMethod 속성을 통해 공개키를 관리하고, 프래그먼트(#) 식별자를 통해 암호학적 서명 검증을 정밀하게 수행함으로써 중앙 서버 없이도 무결점의 기기 인증을 구현한다[12,13].

IV. 데이터 공유·유통 플랫폼 및 신서비스 생태계 조성 동향

지능형 홈 빅데이터 플랫폼의 궁극적인 지향점은 방대한 데이터를 단순히 수집하고 보관하는 데이터 웨어하우스 수준을 넘어, 이를 표준화된 형태로 외

부에 개방하여 새로운 경제적 가치를 창출하는 데이터 유통 포털 및 신서비스 생태계 조성에 있다. 최근 산업계에서는 다양한 제3자 서비스 기업과 데이터 분석가들이 복잡한 인프라 지식 없이도 쉽게 데이터에 접근하고 거래할 수 있도록 범용적인 개방형 API를 고도화하고 있으며, 이러한 플랫폼 간의 상호운용성을 보장하기 위한 국가 차원의 단체표준 제정이 활발히 전개되고 있다.

1. 확장성 기반의 다계층 OpenAPI 아키텍처 설계 동향

이기종 소프트웨어 애플리케이션 간의 원활한 상호작용을 보장하기 위해, 최근의 빅데이터 플랫폼들은 REST 아키텍처 스타일을 범용적으로 채택하고 있다. 특히 실시간 데이터 검색 및 전송 시 처리 부하를 줄이기 위해, 복잡한 계층 구조를 가진 XML 방식보다는 경량화되고 기계와 사람 모두 읽기 쉬운 JSON 데이터 교환 형식을 표준으로 채택하는 추세가 두드러진다. 산업계에서는 통상 수집 종의 빅데이터 OpenAPI를 상용화하여 개방하고 있으며, 이는 크게 3가지 계층의 관점으로 세분화되어 제공된다.

- **수용가(세대) 관점의 API:** 공동주택 개별 세대의 온수, 난방, 전기, 수도, 가스 등 5종 에너지 정보 수집 및 조회를 지원한다. 대표적으로 특정 기간(시작~끝 시간)의 누적 및 순시 사용량을 요청하는 세대별 에너지 정보 수집 API 등이 JSON 구조의 POST/GET 메서드로 빈번하게 활용된다.
- **단지 설비 운영 관점의 API:** 단지 내 기계실 및 공용부 주요 설비의 상태를 실시간으로 모니터링한다. 열교환기 정보 수집 API, 온수펌프 상태 API, 열손실 정보 수집 API 등을 통해

설비의 이상 유무와 에너지 효율 진단 서비스를 지원한다.

- **공급자 발전 및 지역 벤치마크 관점의 API:** 열 에너지를 생산하는 사업자 관점에서 열병합발전기(CHP) 운영 API, 월간/일간 엔진 운영 정보 API 등을 제공한다. 또한, 특정 지역 단위로 데이터를 검색할 수 있는 지역별 에너지 소비 벤치마크 지표 산출 API를 통해 광역 단위의 수급 현황 조회를 돕는다.

2. 국내외 기술 표준화 동향

과거 열에너지 분야는 단지마다 수집 데이터 포맷과 통신 구조가 상이하여 데이터 유통에 큰 장애를 겪어 왔다. 이를 근본적으로 해결하기 위해 국내에서는 한국정보통신기술협회(TTA) 산하 PG424를 통해 「공동주택 열에너지 데이터 플랫폼」 단체 표준(TTAK.KO-10.1439) 제정이 강력히 추진되었다[14]. 이 표준은 단순 가이드라인을 넘어, 이기종 인프라의 레거시 DB를 수용할 수 있도록 크게 3개의 부(Part)로 세분화하여 제정되었다. 제1부(일반 요구사항)는 플랫폼 전 주기의 기능적/비기능적 요구사항을, 제2부(참조구조)는 현장 설비부터 데이터 어댑터(DAU)와 빅데이터 서버 간의 논리적/물리적 참조 아키텍처를, 제3부(데이터 모델 및 인터페이스)는 5종 에너지 검침 데이터의 세부 필드 규격과 교환 인터페이스 명세를 엄격히 정의하여 투명한 교환 생태계의 청사진을 마련했다. 국제적으로는 ITU-T SG13 및 SG17을 중심으로 신뢰 기반 사용자 동의 관리 프레임워크(Y.trcm-fr) 및 PII(개인식별정보) 동의 관리에 관한 국제 표준 제안이 활발히 전개되며, 제3자 서비스 제공자와 지능형 홈 허브 간의 안전한 글로벌 데이터 공유 체계가 확립되고 있다[15].

표준화된 아키텍처 위에서 유통되는 데이터 자원은 원시 센서 데이터부터 AI 추론 결과까지 다양하며, 각 특성에 맞춘 데이터 생명주기 및 스토리지 계층화 관리가 필수적이다. 고빈도(초 단위)로 발생하는 시계열 원시 데이터는 생성 즉시 7일간 Hot 계층(SSD)에 보관된 후, 30일간 Warm 계층(HDD)으로 이동하며, 이후 Cold 및 Archive 단계로 전환되거나 자동 삭제되는 정교한 보존 정책이 적용된다.

더불어, 비식별화 및 차분 프라이버시 적용으로 인해 생성된 재현 데이터가 오픈 마켓에 유통될 때, 원본 자료의 특성을 얼마나 유사하게 재현했는지 평가하는 유용성 지표와 원본이 노출될 가능성을 재는 노출 위험도 지표 간의 트레이드오프가 정밀하게 관리된다. 주로 평균 제곱 오차(MSE: Mean Squared Error), MAE, RMSE 지표를 활용하여 예측값과 실제값 간의 차이가 0에 수렴하도록 모델의 성능을 평가 및 입증하고 있다.

3. 폴스택 데이터 유통 포털 및 개발자 친화적 샌드박스 환경

데이터 개방 생태계를 완성하기 위해, 최근 플랫폼 기술은 단순한 API 엔드포인트 제공을 넘어 데이터 카탈로그 등록 — 권한 공유 — 스마트 컨트랙트 매칭 — MyData 연계로 이어지는 완결된 형태의 폴스택 데이터 유통 전용 포털 솔루션을 활발히 도입하고 있다. 유통 포털 내에서 데이터의 거래 유형은 수집된 형태 그대로를 제공하는 원 데이터 거래, 전처리 과정을 통해 결측치 및 이상치가 정제된 가공 데이터 거래, 그리고 과거 사용량 분석 기반의 미래 예측 결과를 제공하는 지능형 분석 서비스 등 세 가지 형태로 세분화되어 유통된다. 포털 이용자는 메타데이터가 체계적으로 카탈로그화된 신규 공유 데이터나 조회수/거래 상위 데이터를 직관적인 대

시보드에서 확인하고, 보유한 디지털 포인트를 차감하여 데이터를 안전하게 구매(매칭)할 수 있다. 구매한 데이터에 대해서는 최댓값, 최솟값, 평균, 편차 등의 기초 통계를 기간 및 지역별로 자동 생성하며, 인공지능 예측 모델을 결합한 시각화 그래프를 클라우드 환경에서 즉시 제공하는 기술이 적용되고 있다. 또한, 제3자 개발자의 생태계 진입 장벽을 낮추기 위해 Swagger UI 기반의 대화형 API 명세서를 제공하고, 서비스 배포 전 알고리즘을 안전하게 사전 검증할 수 있는 샌드박스 환경을 지원하고 있다[16]. 더불어, 최근에는 분석 서비스 결과에 대한 수용가와 개발자의 신뢰도를 높이기 위해, AI가 특정한 요금 예측이나 결측치 보정 결과를 도출한 근거를 논리적으로 설명해 주는 설명 가능한 AI 기술의 도입도 새롭게 요구되고 있다[17]. 그뿐만 아니라 유통되는 데이터의 신뢰성을 담보하기 위해 완전성, 일관성, 유효성을 실시간으로 측정하는 자동화된 데이터 품질 평가 파이프라인이 샌드박스 내에 기본 탑재되는 추세이다.

V. 인공지능 기반 융합 서비스 시나리오 및 데이터 마켓 활용 사례

단일 플랫폼 구축을 넘어 데이터 마켓플레이스의 효용성이 입증됨에 따라, 전력 분야의 에너지마켓플레이스(EN:TER) 성공 사례를 적극 벤치마킹하여 에너지 특화 개방형 포털을 통해 축적된 고품질 데이터가 다양한 인공지능 기반 융합 서비스로 부가가치를 창출하고 있다.

1. 결측치/이상치 자동 보정을 위한 디노이징 마스크 오토인코더

실제 현장에서는 통신 장애, 센서 결함 등으로 데

이터 누락(결측)이나 비정상적 값(이상치)이 빈번히 발생한다. 기존 시스템은 단주기/장주기 결측 특성에 따라 비학습형 방식인 선형 보간법과 과거 평균법을 자동 선택하는 모듈을 운영했다. 그러나 최근에는 딥러닝 기반의 차세대 전처리 기술인 디노이징 마스크 오토인코더 기술이 적극 도입되고 있다. DMAE는 입력 데이터에 인위적 노이즈를 가한 뒤 원본을 복원하도록 학습하는 DAE의 강인함과 입력값의 특정 구간을 가리고 주변 데이터와의 유사성을 통해 결측치를 복원하는 MAE의 장점을 결합한 혁신적인 아키텍처이다. 이 기술은 데이터 결측률이 50%를 상회하는 열악한 희소 환경에서도 정상 데이터의 잠재적 특징을 강력히 추출하여 획기적인 복원 성능을 달성하며, 결과적으로 AI 수요 예측 모델의 정확도를 극대화한다. 나아가, 복원된 데이터 기반의 예측 결과에 대해 사용자의 신뢰도를 높이기 위해, 딥러닝 모델의 블랙박스 의사결정 과정을 시각적으로 해석해 주는 설명 가능한 AI(XAI) 분석 기법이 추가 결합되어 서비스의 실효성을 더하고 있다.

2. 공급 시설 최적화 및 인공지능 자원 스케줄링

열병합발전기, 대형 보일러, 히트펌프 등 지역난방 공급 설비의 과거 운영 데이터와 기상 정보를 결합하여, 머신러닝 기반으로 수용가의 미래 수요를 정밀하게 예측한다. 예측된 수요 패턴을 바탕으로 설비의 가동 시간을 최적화하며, 특히 심야 시간대 등의 저렴한 잉여 에너지를 대형 축열조에 저장해 두었다가 피크 시간대에 방열 시점을 결정하는 고도화된 에너지 자원 스케줄링 및 모델예측제어 기술이 실현되어 공급망 전체의 에너지 손실을 극적으로 절감하고 있다.

3. 수용가 에너지 벤치마크 및 모바일 이상 진단 라이프케어

데이터 분석 결과를 수용가에게 직접 환원하는 B2C 서비스도 확산 중이다. 단지 내에서 주거 면적이나 가구원 수가 유사한 세대들을 클러스터링하여 일/주/월 단위 소비 패턴을 시각화하고, 비교 순위를 제공함으로써 게임화 요소를 통한 자발적인 에너지 절약을 유도한다. 그리고 스마트폰 모바일 앱을 통해 평소 패턴을 벗어난 급격한 유량 증가 시, 미터기 결측이나 세대 내 난방 밸브 오작동(누수 등)에 의한 난방비 폭탄 등의 이상 징후를 실시간으로 자동 진단하고 알림을 통보해 주는 혁신적인 라이프케어 서비스로 그 영역이 빠르게 확장되고 있다. 또한, 대규모 언어 모델(LLM) 기반의 대화형 에너지 비서가 새롭게 도입되어, 사용자가 자연어로 에너지 절감 팁이나 요금 원인 분석을 요청하면 즉각적으로 맞춤형 인사이트를 제공하는 초개인화 지능형 서비스가 본격적인 실증 단계에 진입했다.

VI. AI 시대, 지능형 홈 데이터 주권 기술 동향

인공지능 생태계가 대규모 언어 모델(LLM)과 가정 내 가전 기기에서 직접 구동되는 온디바이스 AI를 중심으로 급격히 확장되면서, 정보 주체가 자신이 생성한 데이터를 마치 가방이나 휴대폰 같은 개인의 물리적 자산처럼 배타적으로 통제하는 데이터 주권 확보 기술이 산업계의 핵심 화두로 떠올랐다. 지능형 홈 환경에서는 거주자의 상세한 생활 패턴, 음성 대화 이력, 영상 모니터링 기록 등 가장 내밀하고 민감한 개인정보가 대규모 AI 모델 학습의 기초 자산으로 수집·활용되기 때문에 데이터의 수집, 유통, 활용 범위와 목적을 정보 주체가 직접 제어하

는 일은 개인의 권리 보호와 AI 모델 자체의 편향성 및 신뢰성 확보를 위해 무엇보다 중요하다.

이에 따라 기존 중앙집중식 플랫폼 서버 환경에서 소수 거대 기업이 데이터를 일방적으로 독점하던 형태를 탈피하여, 개인이 데이터의 제공 기간과 세부 목적을 직접 설정하고, 생성부터 폐기까지 전 생애주기 이력을 역추적하며, 학습 데이터 제공 기여도에 따른 정당한 경제적 보상까지 자동 정산받는 웹3(Web3) 및 블록체인 기반 데이터 주권 프레임워크가 업계의 차세대 표준으로 강력히 발전하고 있다. 이러한 패러다임 전환을 실현하기 위한 구체적인 3대 융합 기술 동향은 다음과 같다.

1. 디지털 낙인: 워터마크와 해시 검증 엔진 기반 소유권 증명

디지털 환경에서 복제와 위조가 쉬운 데이터의 소유권을 명확히 증명하기 위해서는, 내 물건에 지워지지 않는 이름을 새기듯 데이터 원본 자체에 고유한 디지털 낙인을 찍는 기술이 필수이다. 이를 위해 텍스트, 이미지, 시계열 센서 데이터 등 각 데이터 유형의 특성에 최적화된 비가시적 오픈소스 워터마크 알고리즘과 데이터의 고유한 수학적 지문 역할을 하는 해시 추출 엔진이 결합되어 활용된다. 데이터가 가정 내 IoT 기기에서 생성되거나 클라우드로 업로드되는 즉시, 데이터가 외부망으로 유출되기 전인 엣지 환경의 백그라운드에서 자동 서명 및 해시 생성 에이전트가 작동하여 원본 데이터에 대한 고유 해시값을 즉각적으로 추출한다. 이후 원본 해시값과 사용자의 소유권 정보, 타임스탬프(메타데이터)를 매핑하는 스마트 컨트랙트 인터페이스를 통해, 위변조가 불가능한 블록체인 분산 원장에 안전하게 앵커링함으로써 해당 시점에 해당 개인이 생성한 최초 원본임을 법적·기술적으로 확정

짓는다. 여기에 더해, 데이터 거래와 권한 관리를 자동으로 수행하는 스마트 컨트랙트의 무결성을 보장하기 위해, 코드를 배포하기 전 수학적 증명을 통해 잠재적 취약점을 원천 검증하는 정형 검증 도구의 선제적 도입이 시스템 안전성의 필수 표준으로 자리매김하고 있다.

또한, 데이터의 원본 내용이나 민감 정보 자체를 노출하지 않고도 해당 데이터가 유효한 소유권을 가졌음을 수학적으로 완벽히 증명할 수 있는 영지식 증명(ZKP: Zero-Knowledge Proof) 기술이 새롭게 결합되어 블록체인상의 프라이버시 보호 수준을 한 차원 더 끌어올리고 있다. 이후 데이터가 분석 기관이나 다른 제3자로 유통되거나 거대 AI 학습 환경에 투입될 때는 오프체인 탐지 모듈이 선제적으로 가동되어 워터마크를 추출하고 해시값을 대조한다. 특히 수십~수백 기가바이트에 달하는 대용량 데이터셋이나 수억 개의 파라미터를 가진 대형 AI 모델 환경에서도 검증 과정으로 인한 병목 현상과 속도 저하를 막기 위해, 데이터의 일부 구간만을 무작위로 추출하여 검증하는 샘플링 기반 무결성 검증 기술이 적극 도입되어 위·변조 여부를 실시간으로 탐지한다. 검증 엔진의 대조 결과 일치 혹은 변조의 심을 판별한 검증 결과 자체를 다시 한번 온체인에 기록함으로써 생태계 참여자 누구나 언제든지 검증할 수 있고 누구도 결과를 부정할 수 없는 강력한 무결성 기반의 시스템 신뢰를 구축하고 있다.

2. 데이터 투명성 확보를 위한 AI-BOM과 하이브리드 출처 추적

데이터가 여러 단계의 전처리, 결합, 비식별화 등 복잡한 가공 과정을 거치더라도 최초 원본 생성자가 누구인지, 중간에 어떠한 변형 모델이 적용되었는지 그 발자취를 투명하게 추적하기 위해 AI-

BOM 표준 스키마 도입이 급격히 확대되고 있다. 이는 마치 대형마트에서 판매되는 가공식품에 원산지와 영양성분표를 표기하듯, 해당 AI 학습 데이터가 어떤 재료로 어떻게 만들어졌는지 이력을 명세화하는 혁신적인 가계부 기술이다. 그러나 페타바이트(PB) 규모에 달하는 대규모 학습 데이터를 모두 용량이 제한된 블록체인 네트워크에 직접 기록할 경우, 심각한 트랜잭션 지연 및 블록 용량 초과 문제가 발생한다. 이를 영리하게 해결하기 위해 최신 플랫폼들은 온체인/오프체인 하이브리드 저장 구조를 채택하고 있다. 즉, 무겁고 방대한 용량의 원본 및 가공 데이터 실체는 오프체인(IPFS 등 분산 파일 시스템이나 외부 클라우드 스토리지)에 안전하게 보관하고, 데이터 무결성과 소유권을 증명하는 가벼운 핵심 해시값 및 권한 메타데이터만 온체인(블록체인)에 기록하여 처리 효율성과 보안성이라는 두 마리 토끼를 동시에 확보하는 아키텍처 방식이다 [18].

이러한 하이브리드 추적 시스템을 통해 모델 개발에 참여한 수많은 기여자 정보(Hash), 최신 C2PA(콘텐츠 출처 및 진위 확인을 위한 연합) 표준 스키마에 기반한 데이터셋 구성 요소 및 전처리 특징 명세, 스마트 컨트랙트를 통한 각 기관 간의 유통 경로 기록을 관리한다[19]. 나아가 부여된 특정 데이터 고유 식별자를 기반으로 최초 원본 생성자부터 현재 데이터를 소유하거나 학습에 사용 중인 주체까지의 전 생애주기 역추적 기록이 하나의 체인처럼 연결되어 투명하게 관리되는 추세이다. 이러한 데이터 출처와 가공 과정의 투명성은 곧 해당 데이터를 먹고 자란 AI 모델 결과물의 윤리성, 편향성 극복 수준, 그리고 산업적 신뢰도를 최종적으로 결정짓는 핵심 지표로 작용한다[20].

3. 자기결정형 데이터 관리(ABAC) 및 동적 보상 기술 체계

진정한 의미의 데이터 주권은 단순히 데이터의 흐름을 투명하게 지켜보는 것을 넘어, 정보 주체가 원할 때 능동적으로 접근 권한을 부여하고, 만약 원치 않는다면 언제든지 즉각적으로 권한을 거둬들일 수 있는 강력한 통제권에서 완성된다. 이를 기술적 시스템 수준에서 구현하기 위해 과거의 획일적인 접근 제어를 벗어난 속성 기반 접근 제어 모델이 핵심적으로 적용되고 있다. 사용자는 ABAC 인터페이스를 통해 본인 데이터의 유효 제공 기간(예: 6개월 후 자동 파기), 특정 산업군(의료 연구용, 비영리 공공 목적 등)으로 한정된 엄격한 활용 목적, 그리고 식별 위험이 큰 민감 정보(나이, 상세 위치 등)를 선별적으로 마스킹하여 제외하는 제공 범위 등을 매우 세밀하고 탄력적으로 통제한다.

특히 정보 주체가 마음이 바뀌어 데이터 제공의사를 철회할 경우, 플랫폼상의 온체인 연동 권한 철회 모듈과 규제 대응 API가 즉각적으로 작동하여, 단순히 정책상 약관으로만 금지하는 것이 아니라 시스템 아키텍처 레벨에서 해당 데이터의 외부 유통망 전송 및 현재 진행 중인 AI 학습 모델로의 스트리밍을 원천 차단하고 철회 감사 이력을 법적 증빙용으로 자동 추출한다. 나아가, 데이터 주권 프레임워크의 대미는 정보 주체가 생태계에 제공한 고품질 데이터에 상응하는 정당한 경제적 이익을 공정하게 분배받는 보상 체계에 있다.

이를 위해 각 개인이 제공한 데이터의 실질적 가치를 정밀하게 측정하는 연합학습 기여도 산정(Shapley Value) 경량화 알고리즘이 적용된다. 이 알고리즘은 데이터의 품질과 AI 모델 성능 향상에 미친 개별 한계 기여도를 95% 이상의 정확도로 분석해 가중치를 산정한다. 이렇게 정밀하게 산정된 기

여도 데이터는 온체인상의 동적 보상 비율 할당 아키텍처로 전달되며, 사전에 투명하게 약속된 수익 분배 조건과 정산 주기가 프로그래밍된 스마트 컨트랙트를 통해 사람의 개입 없이 자동 실행된다. 산출된 토큰이나 포인트 보상은 악의적으로 무의미한 데이터를 전송해 보상만 챙기려는 부정 수급 방지 검증 로직을 무사히 통과한 후, 기여자의 블록체인 기반 디지털 지갑으로 실시간 자동 정산 및 이체된다. 이때 발생하는 모든 지급 내역과 기여도 산정 근거 데이터는 영수증 형태로 블록체인상에 투명하게 앵커링되어, 추후 어떠한 분쟁이 발생하더라도 정산의 공정성을 완벽히 보장한다.

지능형 홈 영역의 이러한 선도적 기술 동향은 거주자를 수동적인 데이터 착취의 대상에서 벗어나 스스로 권리를 통제하고 가치를 누리는 능동적인 데이터 주권자로 변화시키며, 향후 투명하고 신뢰성 높은 스마트 에너지 산업 전반에 막대한 경제적 패러다임 전환을 가져올 것으로 기대된다.

VII. 결론

본고에서는 기후변화 대응과 국가 탄소중립 실현의 핵심 과제인 주거 부문의 에너지 소비 최적화를 위해, 지능형 홈 환경에서의 에너지 빅데이터 유통 플랫폼 구축 및 마이데이터 보안 기술의 최신 동향을 심도 있게 살펴보았다.

과거 파편화된 사용량 정보의 단순 수집 단계를 지나, 현재의 지능형 홈 에너지 생태계는 데이터의 안전한 개방과 혁신적인 융합 서비스 창출을 목표로 빠르게 진화하고 있다. 본문에서 분석한 바와 같이, 고빈도 시계열 데이터의 프라이버시 침해를 막기 위해 차분 프라이버시(DP) 모델과 최고 등급 암호화(CC EAL7) 등 강력한 방어 체계를 구축하였다. 또한, 개방형 OpenAPI와 TTA 단체표준 기반의 유

통 포털을 통해 원활한 데이터 거래 생태계를 조성하고 있다. 더 나아가, AI 시대 정보 주체의 권리를 보호하기 위해 무결성 증명, AI-BOM 기반 출처 추적, 그리고 연합학습 기여도 산정이 결합된 블록체인 기반의 데이터 주권 프레임워크가 차세대 핵심 기술로 부상하고 있다.

이러한 선도적 기술 동향과 개방형 인프라의 확산은 건물 에너지 효율화 촉진을 통한 국가 탄소중립 정책의 든든한 기반이 될 것이며, 향후 투명하고

신뢰성 높은 스마트 에너지 산업 전반에 막대한 경제적 파급 효과와 고수익 신규 비즈니스 창출 기회를 가져올 것으로 기대된다. 이와 더불어 FedProx 기반의 연합학습 최적화, 자동화된 데이터 품질 평가(DQA), 스마트 컨트랙트 정형 검증, 그리고 LLM 기반 초개인화 비서 등 혁신 기술들이 지속적으로 융합됨에 따라 지능형 홈 플랫폼의 가치는 나날이 증대될 것이다.

참고문헌

- [1] 한국전력공사, “에너지마켓플레이스(EN:TER) 구축 및 운영 방안,” 한국전력공사, 2019.
- [2] The White House, “Big Data Research and Development Initiative,” Executive Office of the President, 2012. <https://obamawhitehouse.archives.gov/>
- [3] ITU-T, “Recommendation Y.3056: Trust-based user consent management framework in smart city networks (Y.trcm-fr),” ITU-T SG13, 2022.
- [4] 관계부처 합동, “데이터·AI 경제 활성화 계획,” 과학기술정보통신부, 2019.
- [5] Connectivity Standards Alliance(CSA), “Matter Specification Version 1.2,” CSA, 2023. <https://csa-iot.org/all-solutions/matter/>
- [6] C. Dwork, “Differential Privacy: A Survey of Results,” in Proc. Int. Conf. Theory Appl. Models Comput., (Xian, China), Apr. 2008, pp. 1-19.
- [7] J.H. Cheon et al., “Homomorphic Encryption for Arithmetic of Approximate Numbers,” in Proc. ASIACRYPT, (Hong Kong, China), Dec. 2017, pp. 409-437.
- [8] T. Li et al., “Federated Optimization in Heterogeneous Networks,” in Proc. MLSys Conf., (Austin, TX, USA), Mar. 2020, pp. 429-450.
- [9] K. He et al., “Masked Autoencoders Are Scalable Vision Learners,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., (New Orleans, LA, USA), Jun. 2022, pp. 16000-16009.
- [10] Common Criteria, “Common Criteria for Information Technology Security Evaluation (ISO/IEC 15408),” CCRA, 2022. <https://www.commoncriteriaportal.org/>
- [11] S. Rose et al., “Zero Trust Architecture,” NIST Special Publication 800-207, 2020. DOI 10.6028/NIST.SP.800-207
- [12] W3C, “Decentralized Identifiers (DIDs) v1.0,” W3C Recommendation, 2022. <https://www.w3.org/TR/did-core/>
- [13] W3C, “JSON-LD 1.1: A JSON-based Serialization for Linked Data,” W3C Recommendation, 2020. <https://www.w3.org/TR/json-ld11/>
- [14] 한국정보통신기술협회(TTA), “공동주택 열에너지 데이터 플랫폼,” TTAK.KO-10.1439, 2020.
- [15] C2PA, “C2PA Technical Specification Version 1.3,” Coalition for Content Provenance and Authenticity, 2023. <https://c2pa.org/>
- [16] L.S. Shapley, “A Value for n-Person Games,” Contributions to the Theory of Games II, vol. 2, 1953, pp. 307-317.
- [17] OpenAPI Initiative, “OpenAPI Specification v3.1.0,” Linux Foundation, 2021. <https://spec.openapis.org/oas/latest.html>
- [18] D. Gunning et al., “XAI-Explainable artificial intelligence,” Sci. Robot., vol. 4, no. 37, no. eaay7120, 2019. DOI: 10.1126/scirobotics.aay7120
- [19] S. Goldwasser, S. Micali, and C. Rackoff, “The knowledge complexity of interactive proof systems,” in Proc. ACM Symp. Theory Comput., (Providence, RI, USA), May 1985, pp. 291-304.
- [20] J. Benet, “IPFS - Content Addressed, Versioned, P2P File System,” arXiv preprint, 2014. DOI: 10.48550/arXiv.1407.3561