정 보 보 호 학 회 지
제 33 권 제6호, 2023. 12

산업제어시스템에서의 AI IDS 성능 향상을 위한 데이터 품질 연구 동향 및 제언

권 남 혁*, 김 유 신**, 우 은 규*, 정 다 훈**, 채 척**, 신 동 훈**

요 약

최근 산업제어시스템을 대상으로 하는 보안 사고가 지속적으로 증가함에 따라서 이상탐지 시스템에 대한 다양한 연구가 진행되고 있다. 특히 AI 기술의 급속한 발달과 함께 수준 높은 AI기반 이상탐지시스템이 연구되고 있다. 이러한 AI 모델은 산업제어시스템 환경에서 적용할 수 있도록 실시간의 처리가 필요하며, 데이터 세트의 학습에는 산업제어시스템 특성을 고려하는 것이 요구된다. 따라서, 데이터 세트가 산업제어시스템에서 적합하게 활용될 수 있는지 판별할 수 있는 세부 기준을 마련하게 된다면, 우수한 데이터 세트의 활용을 통해 산업제어시스템을 위한 AI 모델의 성능이 향상될 것으로 보인다. 본 논문에서는 산업제어시스템의 AI 침입 탐지시스템의 성능 향상을 위한 데이터 품질 연구의 동향을 조사하고, 향후 발전을 위한 방향성을 구체적인 평가항목을 통해 제시하고자 한다.

Ⅰ. 서 론

최근의 기술 발전으로 산업제어시스템의 중요성이 부각되고 있다. 수자원, 가스, 전기, 에너지 등의 다양한 산업 분야의 기반시설에서 핵심적으로 사용되는 산업제어시스템(Industrial Control System)은 각 분야 설비의 실시간 모니터링, 자동 제어, 데이터 수집 및 분석을 통해 생산성을 높이고 안전성을 강화함으로써 가용성을 유지하고 있다[1].

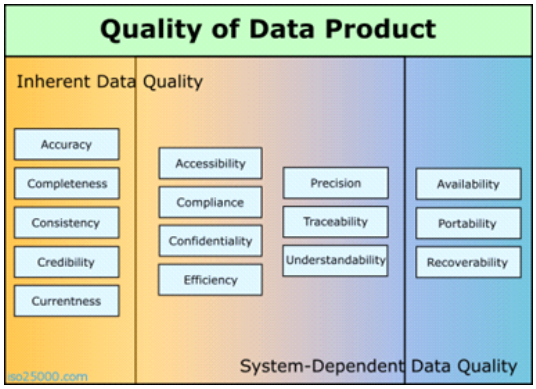
이러한 중요 기반시설과 산업제어시스템이 보안 문제에 노출되면 심각한 사회적 문제를 초래하게 된다. 기반시설의 산업제어시스템은 IT 시스템 환경과는 다르게 인터넷망과 내부 제어망이 분리되어 운영되어왔기에, 사이버 위협에 대한 위협 가능성을 낮게 평가해왔고 침입 탐지시스템(Intrusion Detection System)을 제한적으로 사용해왔다. 하지만 2010년에 이르러 이란의 우라늄 농축 시설의 원심분리기를 약 1000대 파괴한 스턱스넷(Stuxnet)과 2015년에 발생한 우크라이나 대규모 정전 사건 등 기반시설을 목표로 하는 공격 사례가 발생하기 시작했다[2,3].

이에 대응하기 위한 수많은 보안 연구가 진행되었으며, 최근에는 AI 기술의 발전과 함께 위협탐지 성능이 크게 향상된 AI 기반의 침입 탐지시스템이 등장하고 있다[4,5]. 그러나 산업제어시스템 환경에서 AI 모델을 쉽게 적용하기 어렵다는 점과 실시간 및 가용성이 중요시되는 환경에서 대량의 데이터와 긴 학습 및 처리시간을 가진 AI 모델을 적용하는 데에 어려움이 있다. 또한, 산업제어시스템만의 특성이 반영된 데이터가 충분하지 못하며 이를 위한 AI 모델 또한 현장에 적용할 수준까지 연구되지는 않고 있다.

산업 제어시스템의 AI IDS 학습을 위한 데이터의 품질이 중요하며, 현존하는 데이터의 신뢰성을 평가하고 새로이 만들어지는 데이터의 품질 향상을 위해 데이터를 평가하는 방안 또한 필요하다. ICS 분야의 특성을 고려하고 실제 환경의 공격을 반영한 우수한 데이터 세트를 생성하고 판별할 수 있는 세부 기준을 마련할 것이 요구된다. 산업제어시스템은 국가의 주요 시설에서 사용되고 있기에 보안 시스템의 성능 평가항목 및 기준에 관한 독자적인 연구가 필요하며, 국내 시설은 국내법에 따라 설계 및 운용되므로 이를 반영하

이 연구는 ETRI부설연구소의 위탁연구과제[2023-054]로 수행한 연구결과 및 과학기술정보통신부의 재원으로 한국연구재단 지원을 받아 수행된 연구임(No.NRF-2022R1G1A1011933)

* 대구경북과학기술원 융복합대학 기초학부 (학부생, knh7345@dgist.ac.kr, 학부생, ekwoo@dgist.ac.kr)
** 대구경북과학기술원 전기전자컴퓨터공학과 (대학원생, yooshin0303@dgist.ac.kr, 대학원생, hndada@dgist.ac.kr, 대학원생, cocjr0208@dgist.ac.kr, 교수, dshin@dgist.ac.kr)



〔그림 1〕 ISO/IEC 20512 데이터 품질 표준 항목〔7〕

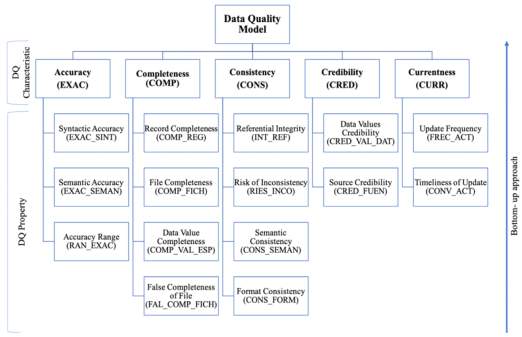
여 보안 요구사항을 파악하고 적절한 평가 기준을 제시할 필요가 있다.

본 논문은 제어시스템의 AI IDS의 성능 향상을 위한 데이터 품질 연구에 대한 동향 파악 및 핵심적인 항목 제언을 목표로 한다. 2장과 3장에서는 각각 국외동향과 국내동향에 대해 살펴본다. 4장에서는 이를 바탕으로 도출한 공통적인 평가 요소들을 파악하고, 이후 AI IDS에 특화된 필수적인 평가항목을 제언한다. 마지막 5장에서 본 논문의 결론을 맺는다.

Ⅱ. 국외동향

2.1. ISO/IEC 25012 & 25024

AI IDS에 특화된 표준은 현재 존재하지 않으나, 일반적인 전통 데이터나 빅데이터에 대한 품질 표준에 관한 연구들은 1990년대 이래로 활발히 진행되어왔다. 먼저 2008년에 발표된 ISO/IEC 25012의 경우, 컴퓨터 시스템(소프트웨어)에 관한 데이터 품질 표준으로, SQuaRE(Software product Quality Requirements and Evaluation)의 일환이다. ISO/IEC 25012는 15개의 품질 평가항목으로 이루어져 있으며, 그 항목은 크게 두 가지 카테고리(내재, 시스템-의존)로 분류된다. 항목이 내재(Inherent)적이라는 것은 데이터의 고유한 특성, 즉 데이터 도메인값이나 그 값 간의 관계, 메타 데이터 등에 관한 항목이라는 것을 의미하며, 그 예로 정확성, 완전성, 일관성 등이 있다. 시스템-의존(System-Dependent) 항목은 외부의 정보 시스템을 통해 데이터 품질이 달성·보존되는 정도를 의미하며, 그 예로 접근성, 기밀성, 복구성 등이 있다. 데이터 저장소(repository)의 기술적 특성은 그 차이가 매우 광범위하므로, 서로 다른 조직 간의 비교가 거의 불가능하여 내재적인 데이터 품질 특성만 고려되어왔지만, ISO/IEC 25012는 저장소의 특성이나 데이터 세트의 특수성과 관계없이 독립적으로 수행, 반복, 비교하는 것을 가능케 한 것이 큰 특징이다[6]. 아래는 그림 1은 ISO/IEC 25012의 항목에 관한 것이다.



〔그림 2〕 ISO/IEC 20524 데이터 품질 표준 항목〔6〕

이후, 2015년에는 ISO/IEC 25024가 발표되었다. 25024 또한 25012와 마찬가지로 SQuaRE 시리즈의 일부이며, 현재 와이즈스톤 등 여러 데이터 품질인증 기관에서 사용하고 있다. ISO/IEC 25024는 25012의 '데이터 품질 특성(data quality characteristic)' 개념과, Rodríguez et al[8], Merino[9]의 논문에서 소개된 '품질 속성(quality property)' 개념을 동시에 적용하여 25012를 개선한 표준이다[6]. '품질 속성'이란, 저장소에 포함된 데이터의 특정 측면 또는 특수성을 평가하는 방법을 나타내는 요소로, 25012의 데이터 품질 특성 항목들의 하위 항목들로서 기능한다. 위 그림 2는 ISO/IEC 25024의 항목에 관한 것이다.

2.2. Cai & Zhu의 표준

21세기 이후, 정보 기술 산업의 획기적인 발전과 클라우드 컴퓨팅이나 IoT, 소셜 네트워크 등의 출현과 함께, 빅데이터의 시대가 도래하게 되었다. 빅데이터는 V4로 알려진 4가지 주요 특성(Volume, Variety, Velocity, Value)가 있는데, 이러한 특성들 때문에 빅데이터를 사용·처리할 때 가변적이며 복잡한 데이터 세트에서 고품질의 실제 데이터를 추출하는 것이 매우 중요하다. Cai & Zhu[10] 는 기존의 데이터 품질 표준과 관련한 연구들을 총정리하고, 계층적

(hierarchical)인 새로운 빅데이터 품질 표준을 제안하였다. Cai & Zhu에 의하면 MIT의 Wang & Strong [11] 등의 데이터 품질 평가에 관한 고전적인 분석부터 Katal, Wazid & Goudar[12] 등의 빅데이터에 관한 최신 연구까지 다양한 결과를 기반으로 하여 새로운 표준을 제시하였다. 데이터 공급자뿐만 아니라 사용자들의 관점을 대폭 반영하였고, 이는 5개의 dimension (대분류)과 14개의 element(소분류)로 구성되어 있다. 5개의 dimension은 4개의 inherent 분류와 1개의 customer satisfaction 분류로 나누어져 있다. 이들은 빅데이터 품질 표준뿐만 아니라, 항목들을 이용한 구체적인 평가 절차까지 논문에서 제시하였다. 위 그림 3과 그림 4는 각각 Cai & Zhu의 빅데이터 품질 표준 항목과 그 평가 절차에 대한 것이다.

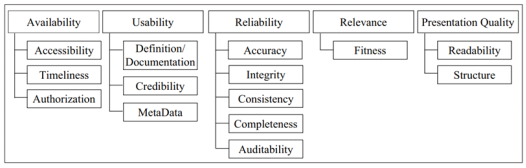


Figure 2: A universal, two-layer big data quality standard for assessment.

〔그림 3〕 Cai & Zhu의 빅데이터 품질 표준 항목 〔10〕

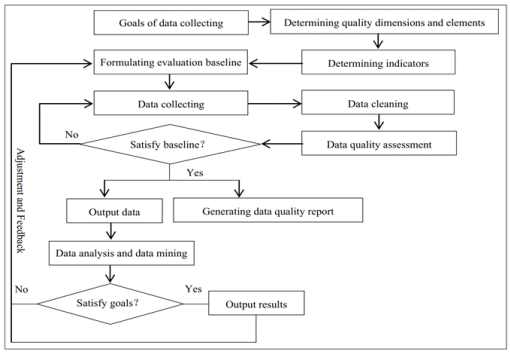
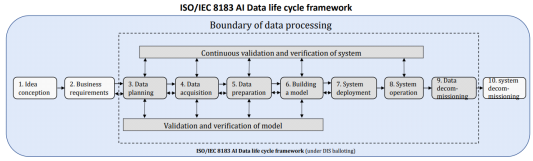


Figure 3: Quality assessment process for big data.

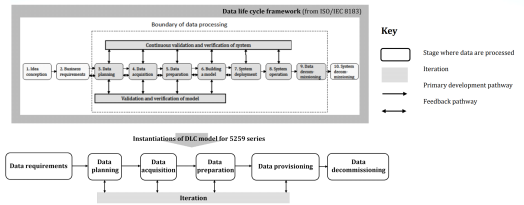
〔그림 4〕 Cai & Zhu의 빅데이터 품질 평가 절차 〔10〕

2.3. ISO/IEC JTC SC 42 5259

ISO/IEC JTC SC 42 5259는 머신러닝에 사용되는 데이터를 대상으로 그 품질을 측정하는 표준이다. 해당 표준은 ISO/IEC 8183 “인공지능 데이터 생애주기 프레임워크”와 연동되어 크게 네 가지의 세부 표준을 제시하고 있다. 첫 번째로 ISO/IEC 5259-1: 개념 소개



〔그림 5〕 ISO/IEC 8183 AI 데이터 생애주기 프레임워크 〔13〕



〔그림 6〕 ISO/IEC 5259-1에서 밝히는 표준의 적용 범위 〔13〕

는 ISO/IEC JTC SC 42 5259 표준 시리즈 내의 개별 표준 간의 연관 관계를 소개한다. 두 번째는 ISO/IEC 5259-2: 데이터 품질 측정 방법 및 품질 향상 기법으로, 정확성, 정밀도, 완전성, 대표성 등 데이터 품질 측정을 위한 항목을 제시한다. 그 외에도 ISO/IEC 5259-3: 데이터 품질관리를 위한 요구사항 및 ISO/IEC 5259-4: 데이터 품질관리 프레임워크 등을 제시하고 있다. 위 그림 5와 그림 6은 각각 AI 데이터 생애주기 프레임워크와 ISO/IEC 5259-1의 표준 적용 범위에 대한 것이다.

2.4. NERC & NASPI

NERC의 RTBPTF(Real-time Tools Best Practices Task Force)와 NASPI의 PARTF (PMU Applications Requirements Task Force)는 싱크로페이저 데이터 품질 요구사항을 부과하였다[14]. NASPI는 싱크로페이저 데이터 품질을 맥락화하여 정적 데이터의 정확성과 계보 측면에서 사용 적합성을 결정함으로, 도메인의 요구사항에 맞춰 다양한 방식으로 데이터 품질을 결정한다. 데이터 품질 요구사항은 어플리케이션에 따라 다르지만 광범위하게 문서화되어 있다.

싱크로페이저 데이터 품질 평가항목은 데이터 품질이 저하되는 원인을 분석 및 분류하여 결정되었다. 분석된 결과에 의하면, 데이터 품질이 저하되는 원인은 크게 장치, 통신, 집계자, 그리고 어플리케이션으로 분류되며, 이에 따라 평가항목인 완전성(Completeness),

정확도(Accuracy), 타당성(Plausibility) 및 가용성(Availability), 기원성(Origination), 일관성(Consistency). 품질 평가 방법(Evaluation of quality)이 파생되었다.

데이터의 완전성은 다양한 값 간의 차이에 중점을 두며, 누락된 값을 파악한다. 완전성은 장치 고장이나 패킷 손실 및 통신 링크 실패에 영향을 받을 수 있다. 정확도는 값 또는 속성의 정확성을 나타내며, 주로 총 벡터 오차(Total Vector Error, TVE)로 측정된다. IEEE 표준 C37.118에 따르면, TVE는 측정된 상부값(크기, 각도 및 주파수)과 예상된 상부값 간의 벡터 차이이다. 타당성의 경우, 시스템의 어떤 현상을 측정하는 과정(관측값)과 그 값을 계산하는 과정(예상값)이 주어진 정밀도로 효과적으로 표준 단위로 문서화되었는지, 명시된 신뢰 구간 내에 있는지를 평가하는 항목이다. 가용성은 네트워크 가용성을 의미하며, 싱크로페이저의 특성상 높은 네트워크 지연이 있으면, 데이터가 누락되거나 불완전하게 인식될 수 있고 이는 시스템의 가용성에 영향을 줄 수 있다. 따라서 네트워크 가용성은 품질에 영향을 미치는 간접적인 속성으로 간주된다. 데이터 측정에서 기원성은 장치 클래스, 표준 준수 및 데이터 조작 기법을 포함한 원산지, 지리적 공간 또는 전기적 토폴로지에 기반한 물리적 위치를 나타내는 커버리지, 그리고 장치, 집계자 또는 어플리케이션 수준에서 적용되는 변환을 포함한다. 일관성은 헤더와 데이터 프레임의 일관성, 데이터 프레임의 순서 일관성, 표준 준수, 디바이스 간 통신 일관성을 평가하는 것을 포함한다. 데이터 품질 평가에는 잘못된 데이터 세트가 있는 애플리케이션을 테스트하는 벤치마킹과 표준화를 통해 품질 저하의 영향을 평가하는 데 중점을 두고 성능에 대한 장치 보정 및 네트워크 상태의 영향을 평가하는 지표이다. NASPI의 전체 품질 평가 지표는 표 1과 같다.

〔표 1〕 NASPI의 스마트 그리드 싱크로페이저 데이터에 대한 품질 평가 지표

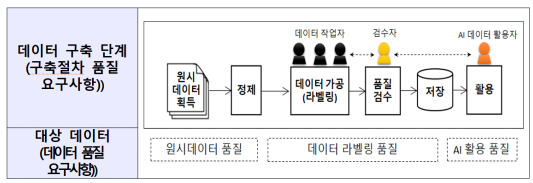
항목	내용
완전성	값의 범위가 주어졌을 때, 누락된 값에 대한 항목
정확도	IEEE 표준 C37.118에 따라 측정된 값 간의 벡터 오류(TVE)
타당성 및 가용성	값을 측정하는 과정이 효과적으로 문서화되어 있는지 타당성을 검토하며, 네트워크 가용성을 평가
기원성	데이터가 측정되는 소스의 신뢰성(원산지, 적용 범위 등)에 기반한 지표
일관성	데이터가 해당 유형의 전체 구조와 얼마나 일치하는지 평가
품질 평가	장치나 네트워크의 상태가 데이터 품질에 미치는 영향을 연구하고, 품질이 좋지 않아 어플리케이션 성능이 얼마나 감소하는지 분석

Ⅲ. 국내동향

3.1. 인공지능 데이터 품질 표준안

국내에서 진행된 데이터 품질 방안 중 하나로 2020년에 배포된 과학기술정보통신부에서 제안한 인공지능 데이터 품질 표준안이 있다. 인공지능 데이터 품질 표준안에서는 주요 품질 요구사항을 크게 2가지(데이터 품질 요구사항, 구축절차 품질 요구사항)로 나뉘어 분류했다. 데이터 품질 요구사항은 원시 데이터 품질 데이터 라벨링 품질, 인공지능 활용 품질로 3가지로 분류했고, 이 중에서 원시 데이터(원천 데이터를 정제하여 데이터 라벨링을 하는 용도로 작업이 완료된 데이터)가 적합한지 평가하기 위해 다양성, 신뢰성, 충분성, 균일성 등을 제안하고, 속성에 맞는지 파일 포맷, 동영상/이미지 해상도 및 컬러 심도 등을 고려했으며, 그리고 품질관리가 되었는지 평가하는 항목을 세웠다. 그리고 데이터 라벨링(인공지능이 학습에 활용할 수 있도록 설명정보 데이터를 추가로 부착하는 과정)을 하는 과정에서 구문 정확성과 의미 정확성을 제시하고, 인공지능 데이터 품질(인공지능 데이터가 사용자에게 유용한 가치를 줄 수 있는 수준)을 평가하기 위해 유효성을 제시했다.

구축절차 품질 요구사항은 원시 데이터가 인공지능 데이터로 활용되는 6단계의 절차로부터 4가지 단계로 압축하여 각 단계에서 필요한 품질 기준을 제시하고 있다. 각각 획득 과정(법/제도 준수, 획득 환경 및 대상 등), 정제 과정(정제 기준, 비식별화, 중복성 방지 등),



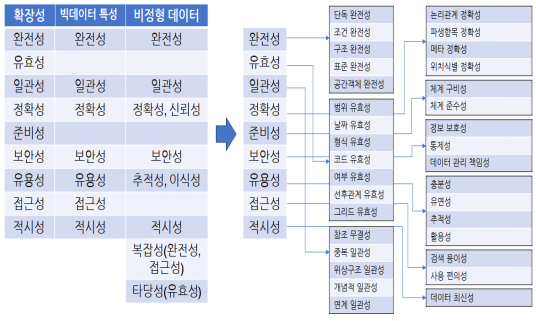
〔그림 7〕 과학기술정보통신부 인공지능 데이터 품질 개념 〔15〕

〔그림 8〕 데이터 품질관리 지표 간의 관계〔16〕

가공 과정(라벨링 포맷 및 도구, 작업 방식 등), 품질 검수(검수 기준 및 방법 등)을 제시하고 있다. 위 그림 7은 인공지능 데이터 품질 개념에 관한 것이다.

3.2. 빅데이터 플랫폼 및 센터 데이터 품질관리 가이드

이후 2021년, 한국지능정보사회진흥원(NIA)에서 빅데이터 플랫폼 및 센터 데이터 품질관리 가이드를 발간했다. 빅데이터의 특성을 고려한 기준을 제작하기 위해 빅데이터와 비정형 데이터에 적용이 가능한 범용성과 확장성을 고려하고, 빅데이터의 생애주기를 고려한 지표를 토대로 상위 수준의 지표를 데이터 품질관리 지표로서 총 9개를 선별해 구성했다. 9개의 지표는 데이터 품질관리의 기반에서부터 활용 단계까지 적용할 수 있음을 보여준다. 아래 그림 8과 그림 9는 각각 9개의 지표 간의 관계와 그에 근거해 제시된 데이터 품질 항목의 모습이다.



〔그림 9〕 빅데이터 플랫폼 및 센터 데이터 품질관리 가이드에서 제시한 데이터 품질 항목〔16〕

3.3. 인공지능 학습용 데이터 품질관리 가이드라인

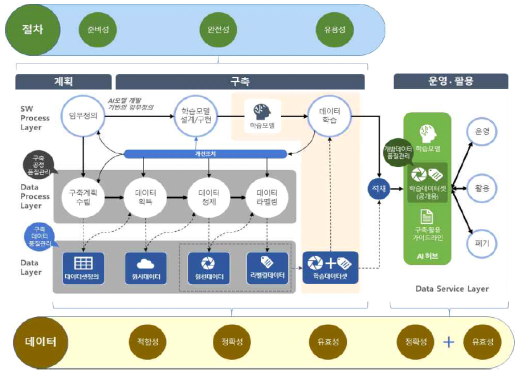
2021년 2월에는 과학기술정보통신부와 한국지능정보사회진흥원, 한국정보통신기술협회에서 인공지능 학습용 데이터 품질관리 가이드라인을 발간했다. 인공지능 학습용 데이터의 생애주기를 분석하여 구분함으로써 데이터 품질관리의 대상과 범위를 명확히 하였다. 생애주기는 크게 '구축프로세스 품질관리', '구축데이터 품질관리', '개방데이터 품질관리' 3 영역으로 구분되며 그 내용과 범위는 아래의 표 2와 같다.

인공지능 학습용 데이터 품질관리 지표는 생애주기 분석을 통해 설정되었으며, 인공지능 학습용 데이터 구축 및 품질관점에서 일치성 분석, 인공지능 학습용 데이터 분석을 통해 구축 및 활용관점을 반영한 적합성, 정확성, 유효성, 준비성, 완전성, 유용성 등 6가지 지표로 구성된다. 수립된 품질관리 지표는 절차(준비성, 완전성, 유용성)와 데이터(적합성, 정확성, 유효성)를 평가하는 데 사용되며 인공지능 학습용 생애주기인 계획, 구축, 운영 및 활용 단계의 품질관리까지 활용하여 데이터 품질 수준을 확보하는 기준이 된다. 인공지능 학습용 데이터 생애주기와 품질지표 간의 관계는 위의 그림 10과 같다.

인공지능 학습용 데이터의 품질관리 세부지표는 품

〔표 2〕 인공지능 학습용 데이터의 생애주기〔17〕

생애주기 구분	내용
구축데이터 품질관리	원시데이터, 원천데이터, 라벨링데이터 품질 검사 및 오류 개선
구축프로세스 품질관리	데이터 획득, 데이터 정제, 데이터 라벨링 등 구축 과정에서 데이터 품질 보장
개방데이터 품질관리	AI Hub에 적재된 데이터 품질관리



〔그림 10〕 인공지능 학습용 데이터 생애주기와 품질지표 간의 관계〔17〕

질관리 대상을 값, 구조, 공정으로 판단하고 검사하기 위해 품질관리 지표를 세분화하여 구성한다. 준비성은 계획 수립성과 체계 준수성, 완전성은 수집 완전성, 정제 완전성과 가공 완전성, 유용성은 사용 편의성과 유연성, 적합성은 기준 적합성, 기술 적합성과 통계적 다양성, 정확성은 의미 정확성과 구문 정확성, 그리고 유효성은 학습모델 유효성으로 구성된다.

Ⅳ. 평가항목

4.1. 공통 평가항목

본 장에서는 2장과 3장에서 조사한 국외·내 동향을 바탕으로, 여러 품질 평가방안에 제시된 항목들을 구체적으로 비교·분석하여 공통적인 요소들을 선별하고 차이점을 정리하여, 산업제어시스템 AI IDS에 특화된 새로운 데이터 품질 평가항목 체계를 최종적으로 도출한다.

4.1.1. 국외의 데이터 품질 평가 기준

국외에서는 일반적인 데이터를 대상으로 하는 표준 ISO/IEC 25012와 인공지능을 위한 데이터의 품질관리에 관한 표준 ISO/IEC JTC SC 42 5259 등 다양한 표준이 일찍이 제시되어 왔다. 위의 표 3은 이에 수록된 평가항목을 비교한 결과이다. Auditability (감사 가능성), Accuracy (정확성), Completeness (완전성), Consistency (일관성), Credibility (신뢰성), Currentness (현시성), Portability (휴대 가능성), Precision (정밀성)과 같이 데이터에 내재되어 있는 속성에 대해, 두 표준에 모두 있는 항목의 경우 일반적인 데이터 및 인공지능을 위한 데이터 모두에 중요하게 작용하는 평가항목인 것으로 추론할 수 있다. 품질 평가항목을 도출할 때, 두 표준에 모두 수록된 항목을 먼저 고려한 후, 두 표준 중 하나에만 정의된 항목에 대해서는 더 세밀한 분류의 데이터를 대상으로 하는 ISO/IEC JTC SC 42 5259에만 있는 항목, 즉 Context coverage (문맥 포괄성), Data scalability (데이터 확장성), Identifiability (신원 확인성), Relevance (관계성), Representative (대표성), Timeliness (적시성)을 먼저 고려하는 것이 합리적이다. 마지막으로 ISO/IEC 25012에만 있는 항목, 즉 Availability (적용성), Compliance (준수성), Confidentiality (기밀성),

〔표 3〕 표준 ISO/IEC 25012와 ISO/IEC JTC SC 42 5259의 공통 평가항목 대응 관계

데이터 품질 특성	ISO/IEC 25012	ISO/IEC JTC SC 42 5259
Auditability	O	O
Accuracy	O	O
Availability	O	
Completeness	O	O
Context coverage		O
Compliance	O	
Confidentiality	O	
Consistency	O	O
Credibility	O	O
Currentness	O	O
Data scalability		O
Efficiency	O	
Identifiability		O
Portability	O	O
Precision	O	O
Relevance		O
Recoverability	O	
Representative		O
Timeliness		O
Traceability	O	
Understandability	O	

Efficiency (효율성), Recoverability (회복성), Traceability (추적성), Understandability (문해성)을 추가로 고려하여 산업 제어시스템을 대상으로 하는 데이터 품질 평가항목 체계를 보강할 수 있을 것으로 기대한다.

4.1.2. 국내의 데이터 품질 평가 기준

국내에서도 4차 산업 혁명으로 일컬어지는 정보 산업 혁명의 흐름에 맞추어 여러 기관에서 표준을 제시한 바 있다. 표 4는 과학기술정보통신부에서 발간한 '인공지능 데이터 품질 표준안'과 한국지능정보사회진흥원 (NIA)에서 발간한 '데이터 품질관리 가이드'의 각 평가항목을 조사하여 공통된 사항을 평가한 항목을 대응시킨 표이다. 국외동향과 다르게 평가항목이 두 표준에 그렇게 많이 겹치지 않는 것을 확인할 수 있다. 과학기술정보통신부에서 발간한 '인공지능 데이터 품질 표준안'에서 적합성은 NIA에서 발간한 '데이터 품

질관리 가이드'와 다르게 다양성 외에도 데이터가 생성되는 시점에서의 평가 요소도 반영한 것이 특징이다. 추가로 요구사항 중 데이터 속성에 대해서도 미디어에 초점이 맞춰진 항목이 대거 수록된 것을 확인할 수 있다. 라벨링 품질관리의 경우 NIA 발간의 품질관리 가이드에서도 동일 속성의 평가항목이 존재하나 일반적인 품질관리에 관하여 서술하고 있지만, 과학기술정보통신부 발간 품질관리 가이드의 경우 라벨링을 특정하여 항목을 수록하였다.

4.2. AI IDS에 특화된 데이터 품질 평가항목

표 5는 앞서 조사한 국내 및 국외의 데이터 품질 평가 표준들에서 공통적으로 발견되는 평가항목을 바탕으로 설계한 산업 제어시스템 대상 AI IDS를 위한 데이터의 품질 평가항목 체계 예시이다. 본 체계에서는 완전성을 비롯하여 정확성, 고유성, 준수성, 최신성, 유효성, 보안성, 접근성, 일관성, 유용성 총 10개의 평가항목을 선정하였다. 각 평가항목은 세부 평가항목을 가질 수 있다. 완전성은 전체 데이터의 집합이 산업 제어시스템이 가질 수 있는 모든 값을 가졌는지를 진단하는 지표인 '데이터 완전성'과 각 데이터가 단순히 비어 있지 않은지 (null의 여부)를 진단하는 지표인 단독 완전성을 세부 평가항목으로 가질 수 있다. 준수성의 경우 데이터 세트가 데이터와 함께 제공되는 문헌에 명시된 내용과 일치하는지 진단하는 지표인 '체계 준수성'과 표기가 표준을 따르는지 진단하는 표기 준수성 등으로 세부 평가항목을 설정할 수 있다. 이러한 평가항목들은 기존 데이터 품질 평가항목들에서 산업 제어시스템의 속성에 맞추어 정의를 개정할 수 있고 그 정량적 평가 방법 역시 산업 제어시스템에 적합하게 바꾸기 용이하게 구성되어 있다.

일반적인 데이터를 대상으로 하는 평가항목 체계와의 다른 점으로는 고유성이 있다. 고유성은 주로 정상 데이터와 공격 데이터로 구성된 산업 제어시스템의 데이터 포맷을 고려하여 공격 데이터에 관한 평가를 다루는 항목이다. 고유성이 가질 수 있는 세부 평가항목으로는 공격 데이터의 양 및 공격 데이터 간의 시간 분포, 전체 실험 데이터 수 대비 수행되는 공격의 양의 비율, 그리고 공격 시나리오가 별도로 설정되어 있을 때 그 시나리오가 정상 데이터의 수에 대하여 충분한 양인지 등이 있다. 고유성의 세부 항목은 사용자의 목

〔표 4〕 인공지능 데이터 품질 표준안과 NIA 빅데이터 품질관리 가이드의 공통 평가항목 대응 관계

인공지능 데이터 품질 표준안			NIA 빅데이터 품질관리 가이드
요구사항	세부 요구사항	세부 요구사항 요소	
적합성	다양성	포괄성	충분성
		변동성	유연성
	신뢰성		-
	충분성		-
	균일성		-
	사실성		-
	편향성		-
속성	파일 포맷		형식 유효성
	미디어 해상도		-
	미디어 프레임 레이트		-
	미디어 컬러심도		-
	텍스트 어절 수		-
품질 관리	적합성		-
	기술 규격		논리 관계 정확성
라벨링 품질 관리	구문 정확성		형식 유효성
	의미 정확성		데이터 최신성
	자체 검증		체계 구비성
활용 품질 관리	유효성	달성 목표 수립	-
		자체 알고리즘 활용	-
		목표치 설정	-
		데이터 세트 간 분포 유사도	-
보안 및 관리	개인정보 유출 방지		정보 보호성
	메타데이터 관리		검색 용이성
	버전 관리		-

〔표 5〕 산업 제어시스템 대상 AI IDS를 위한 데이터 품질 평가항목 체계 예시

평가항목	세부 평가항목 예시	평가항목	세부 평가항목 예시
완전성	데이터 완전성, 단독 완전성	유효성	범위 유효성, 형식 유효성
정확성	논리 관계 정확성, 파생항목 정확성	보안성	보안 무결성
고유성	공격 횟수 다양성, 시나리오 개수 다양성	접근성	검색 용이성, 사용 편의성
준수성	체계 준수성, 표기 준수성	일관성	중복 일관성
최신성	데이터 최신성	유용성	데이터 시간 유연성, 변경 데이터 추적성

적에 따라 그 지향점이 달라질 수 있다. 이 경우 정량적 평가 방법의 세부 설정을 통해 사용자별 평가 목표를 달성할 수 있을 것으로 기대한다.

V. 결 론

본 논문에서는 산업제어시스템에서의 AI 침입 탐지 시스템 성능 향상을 목표로 하여, 데이터의 품질 연구에 대한 동향을 국외와 국내로 분류하여 조사하였다. 또한, 국외동향에서는 ISO/IEC 25012와 ISO/IEC JTC SC 42 5259를, 국내동향에서는 인공지능 데이터 품질 표준안과 NIA 빅데이터 품질관리 가이드에 수록된 평가항목들을 비교하여, 공통적인 부분과 각각의 특성에 대한 분석을 진행하였다. 이러한 분석을 바탕으로, 산업 제어시스템 분야에 적합한 데이터 품질 평가항목 10개를 최종적으로 도출하였다.

AI IDS에 특화된 표준이 현재 존재하지 않는다는 점에서, 본 논문의 국외와 국내동향에 대한 분석과, 새롭게 제시한 평가항목은 유의미한 성과라 할 수 있다. 본 논문에서 제시한 체계의 세부 평가항목들이 추가·보완된다면, 산업제어시스템의 AI IDS 성능 향상에 더욱 큰 기여를 할 것으로 사료된다.

참 고 문 헌

[1] K. Stouffer, V. Pillitteri, S. Lightman, M. Abrams, and A. Hahn, “NIST Special Publication 800-82, 2th Ed.”, *National Institute of Standards and Technology*, pp. 800-82, May 2015.

[2] R.M. Lee, M.J. Assante, and T. Conway, “Analysis of the Cyber Attack on the Ukrainian Power Grid”, *Electricity Information Sharing and Analysis Center*, Mar 2016.

[3] K.E. Hemsley, and R.E. Fisher, “History of Industrial Control System Cyber Incidents”, *NL/CON-18-44411-Rev002, Idaho National Lab*, United States, Dec 2018.

[4] C. Xu, J. Shen, X. Du and F. Zhang, “An Intrusion Detection System Using a Deep Neural Network With Gated Recurrent Units”, *IEEE Access*, vol. 6, pp. 48697-48707, Aug 2018.

[5] A. Sherstinsky, “Fundamentals of Recurrent Neural Network (RNN) and long short-term memory (LSTM) network”, *Physica D: Nonlinear Phenomena*, vol. 404, 132306, Mar 2020.

[6] F. Gualo, M. Rodríguez, J. Verdugo, I. Caballero, and M. Piattini, “Data quality certification using ISO/IEC 25012: Industrial experiences”, *Journal of Systems and Software*, 176, 110938, June 2021.

[7] ISO 25000, “Software product Quality Requirements and Evaluation (SQuaRE) Data Quality model”, *ISO/IEC 25012*, Dec 2018.

[8] M.B. Rodriguez, M. Agus, F. Bettio, F. Marton, and E. Gobbetti, “Digital Mont’e Prama: Exploring large collections of detailed 3D models of sculptures”, *Journal on Computing and Cultural Heritage (JOCCH)*, 9(4), 1-23, Sep 2016.

[9] J.M. Garcia, “Environment for the evaluation and certification of data products quality”, Ph.D. Thesis, *Universidad de Castilla-La Mancha,* April 2017.

[10] L. Cai, and Y. Zhu, “The challenges of data quality and data quality assessment in the big data era”, *Data science journal*, 14, 2-2, 2015.

[11] R.Y. Wang, and D.M. Strong, “Beyond accuracy: What data quality means to data consumers”, *Journal of management information systems*, 12(4), 5-33, Spring 1996.

[12] A. Katal, M. Wazid, and R.H. Goudar, “Big data: issues, challenges, tools and good practices“, *IEEE 2013 Sixth international conference on contemporary computing (IC3)*, Sep 2013.

[13] W. Chang, ISO/IEC JTC 1/SC 42 (AI)/WG 2 (data) data quality for analytics and machine learning (ML), *Information Technology Laboratory*, May 2022.

[14] A. Sundararajan, T. Khan, and A. Moghadasi, “Survey on synchrophasor data quality and cybersecurity challenges, and evaluation of their interdependencies”, *Journal of Modern Power Systems and Clean Energy*, 7(3), 449–467, May 2018.

[15] 과학기술정보통신부, 인공지능 데이터 품질 표준안, *과학기술정보통신부 보도자료*, Oct 2020.

[16] 이용진, 손기문, 유시형, 김은영, “빅데이터 플랫폼 및 센터 데이터 품질관리 가이드: 제 2권 데이터 관리기준 | 품질진단 방법”, *한국지능정보사회진흥원*, Dec 2021.

[17] 이용진, 손기문, 오현목, 유호진, 신다울, 박현우, 전진우, 윤주미, 오윤환, 류동주, 양원, “인공지능 학습용 데이터 품질관리 가이드라인1.0”, *한국지능정보사회진흥원*, Feb 2023.

〈저자소개〉

권 남 혁 (Namhyuk Kwon)

2021년 2월~현재 : 대구경북과학기술원 기초학부 재학

<관심분야> 정보보호, 이상탐지, 인공지능

김 유 신 (Yooshin Kim)

2023년 2월 : 대구경북과학기술원 기초학부 공학사 졸업

2023년 2월~현재 : 대구경북과학기술원 전기전자컴퓨터공학과 석박사 통합과정 재학

<관심분야> 정보보호, AI보안, 이상탐지

우 은 규 (Eungyu Woo)

2020년 2월~현재 : 대구경북과학기술원 기초학부 재학

<관심분야> 정보보호, 계산 기하학, 그래프 이론

정 다 훈 (Dahoon Jeong)

2020년 2월~현재 : 대구경북과학기술원 전기전자컴퓨터공학과 연계과정 재학

2020년 2월 : 대구경북과학기술원 기초학부 공학사 졸업

<관심 분야> 산업 보안, 동형 암호, 양자 컴퓨팅

채 척 (Chuck Chae)

2023년 2월 : 대구경북과학기술원 기초학부 이학사 졸업

2023년 2월~현재 : 대구경북과학기술원 전기전자컴퓨터공학과 석박사 통합과정 재학

<관심분야> 정보보호, 계산 기하학, 그래프 이론

신 동 훈 (Donghoon Shin)

종신회원

2007년 : 한동대학교 전산학, 전자공학 학사

2009년 : 한국과학기술원 전산학 석사

2016년 : 국과학기술원 전산학 박사.

2016년~2019년 : 가보안기술연구소 선임연구원

2019년 ~ 현재 : 대구경북과학기술원 전기전자컴퓨터공학과 교수

<관심분야> 계산 이론, CPS보안, 기반시설 보안 등