



• 특집 • 자율제조(Autonomous Manufacturing) II

# 스마트 유지보수를 위한 문서 계층 구조 추출과 로컬 LLM 다단계 변환을 활용한 정비교범의 S1000D 모듈화

## S1000D Modularization of Legacy Maintenance Manuals Using Style-based Hierarchy Extraction and Multi-step Transformation with a Local LLM

엄주명<sup>1,#</sup>, 안영우<sup>2</sup>, 이승욱<sup>3</sup>, 허선웅<sup>3</sup>, 서영탁<sup>3</sup>  
Jumyung Um<sup>1,#</sup>, Youngwoo An<sup>2</sup>, Seung Uk Lee<sup>3</sup>, Seon Ung Heo<sup>3</sup>, and Yeong Tak Seo<sup>3</sup>

<sup>1</sup> 경희대학교 산업경영공학과 (Department of Industrial & Management System Engineering, Kyung Hee University)

<sup>2</sup> 경희대학교 실감AX융합학과 (Department of XR-AI Convergence, Kyung Hee University)

<sup>3</sup> 한화시스템즈 (Hanwha systems)

# Corresponding Author / E-mail: [jayum@khu.ac.kr](mailto:jayum@khu.ac.kr), TEL: +82-31-201-5443

ORCID: 0000-0002-8040-6144

KEYWORDS: Smart Maintenance (스마트 유지보수), Maintenance manual (정비 교범), Large language model (대규모 언어 모델), Document modularization (문서 모듈화)

*As smart factories evolve, maintenance manuals need to be transformed from static documents into machine-readable and reusable digital assets. However, many legacy manuals are still in unstructured formats, such as Hangul word-processor files, which complicates their updating, reusability, and adaptability to changing product configurations. This paper presents a framework for converting these legacy manuals into S1000D-based documents. It combines style-based hierarchy extraction with rule-guided multi-step transformation using a local large language model (LLM). First, the style information within the Korean documents is analyzed to identify the hierarchical structure of the manual and extract content at various document levels. Next, this extracted content is converted into S1000D XML modules through the local LLM, utilizing category-specific rule files, XML tag definitions, and example templates. To enhance structural consistency and minimize errors, different prompts and rule sets are applied based on the document hierarchy level. A case study involving a maintenance manual for a high-angle limit switch module demonstrates that the proposed method can maintain document structure while generating reusable S1000D-style outputs from legacy technical documents. This approach lays a practical foundation for creating continuously updatable and context-reconfigurable maintenance guidance in smart manufacturing environments.*

Manuscript received: April 7, 2026 / Revised: April 17, 2026 / Accepted: May 7, 2026

### 1. 서론

최근 제조 산업에서는 스마트팩토리 전환이 가속화되면서, 설비 유지보수 작업도 정적인 문서 열람 중심에서 실시간 현장 지원 중심으로 변화하고 있다. 특히 유지보수 작업은 설비 상태, 고장 유형, 작업 이력, 작업자의 숙련도에 따라 필요한 정보가

달라지므로, 고정된 순서의 교범만으로는 현장 대응에 한계가 있다. 또한 제품 개발 과정에서는 설계 변경이 반복적으로 발생하므로, 기술문서 역시 이에 맞추어 지속적으로 갱신될 필요가 있다. 따라서 유지보수 교범을 단순히 문서로 보관하는 수준을 넘어, 변경 사항을 신속히 반영하면서도 필요한 정보를 상황에 맞게 즉시 제공할 수 있는 디지털 자산으로 전환할 필요가 있다[1,2].

스마트팩토리 환경에서는 설비 상태, 알람 이력, 작업 조건과 같은 현장 데이터가 지속적으로 생성되므로, 유지보수 정보 또한 운영 과정과 연계되어 제공될 필요가 있다. 따라서 정비교범은 단순 열람용 문서가 아닌, 실시간 생산·설비 운영 정보와 결합되어 필요한 절차와 주의사항을 상황별로 재구성할 수 있는 형태로 관리되어야 한다. 이를 위해서는 교범 내용을 기계가 읽을 수 있는 구조로 분해하고, 필요에 따라 재구성 가능한 모듈형 정보 체계로 관리하는 접근이 중요하다.

이러한 요구에 대응하기 위해 국제 기술문서 표준인 S1000D가 유지보수 정보의 구조화와 모듈화를 위한 대안으로 주목받고 있다. S1000D는 유지보수 정보를 데이터 모듈 단위로 분리하여 관리함으로써, 정보의 재사용성과 상호운용성을 높일 수 있는 장점을 제공한다. 기존의 기술문서 작성 및 변환 방법은 단발적인 변환 작업을 사람의 수작업에 의해 수행해 왔다[3,4]. 그러나 이러한 방식은 제품 개발 과정에서 문서가 지속적으로 갱신되어야 하는 현실을 충분히 지원하지 못하였고, 최종 제품 상태에 부합하는 문서를 유지하는 데 많은 어려움이 있다. 그 결과, 기술문서의 구조화와 표준화는 가능하더라도, 반복적인 설계 변경과 운영 변화에 유연하게 대응하는 체계로 발전하는 데에는 한계가 있었다.

기존 접근은 기술문서를 표준 형식으로 변환하는 데에는 의미가 있었지만, 대부분 단일 시점의 변환에 머무르는 경향이 있다. 따라서 설계 변경이 잦은 제품이나 지속적인 개정이 필요한 유지보수 교범에는 적합하지 못했다. 또한 실제 유지보수 현장에서는 동일한 설비라도 고장 유형, 경고 상태, 작업 조건, 부품 상태에 따라 우선적으로 제공되어야 할 작업 절차와 주의사항이 달라질 수 있다. 그러나 기존 방식은 변환된 문서를 고정된 결과물로 다루는 경우가 많아, 작업 시점의 맥락에 따라 필요한 모듈만 선택적으로 조합하거나 재배열하는 기능이 부족했다. 특히 복잡한 교범 구조를 한 번의 변환으로 처리하는 방식은 문서 계층, 절차 흐름, 경고·주의 정보의 관계를 안정적으로 보존하는데 한계가 있다. 이로 인해 작업자는 방대한 교범 전체를 탐색해야 하며, 필요한 정보를 즉시 찾기 어렵고 인지 부하가 증가하는 문제가 발생한다[1,5].

본 논문에서는 레거시 유지보수 교범을 모듈화 가능한 구조로 변환하고, 이를 바탕으로 설계 변경과 현장 상황에 유연하게 대응할 수 있는 상황 맞춤형 실시간 가이드 프레임워크를 제안한다. 제안 프레임워크는 LLM 기반 접근을 채택하여, 교범 내 절차, 경고, 조건, 장비 정보와 같은 이질적인 내용을 구조화된 정보 단위로 변환한다[6,7]. 특히 본 연구는 단일 프롬프트에 의존하는 일괄 변환 방식이 아니라, 문서 구조 인식, 정보 분류, 모듈 생성, 표준 형식 매핑의 과정을 단계적으로 수행하는 Multistep 변환 방식을 적용한다[7-9]. 이러한 Multistep 접근은 복잡한 교범의 계층 구조와 의미 관계를 점진적으로 정제하면서 반영할 수 있게 하며, 각 단계에서 문서 구조와 표준 규칙의 일관성을 높이는 데 기여한다. 또한 생성된 정보 모듈은 설계 변경 사항이나 작업 시점의 요구에 따라 재조합 될 수 있으므로,

기술문서를 단발성 결과물이 아니라 지속적으로 갱신·재구성 가능한 동적 자산으로 활용할 수 있게 한다. 이를 통해 작업자는 전체 교범을 순차적으로 탐색하는 대신, 현재 상황에 직접적으로 관련된 절차와 주의사항을 실시간으로 제공받을 수 있다.

구체적인 multistep 변환 절차와 모듈 재구성 방식은 다음 장의 상세히 설명한다. 2장은 유지보수 교범의 기존 모듈화 방법을 소개하고 3장 단발적인 문서 변환을 넘어, 설계 변경이 잦은 제품 개발 환경에서도 지속적으로 갱신 가능한 스타일 기반의 기존 교범의 계층 구조 추출 방법과 로컬 LLM을 사용한 규칙 기반의 S1000D 양식 변환 방법, 규칙 파일 및 프롬프트를 소개한다. 4장에서는 실제 제품 유지보수 교범을 활용해 S1000D 모듈 변환 과정을 설명하며 최종 결과를 비교한다. 마지막으로 5장에서는 토론과 결론으로 본 논문을 마무리한다.

## 2. Modularized Document System

S1000D는 공통 소스 데이터베이스를 기반으로 기술문서를 작성, 관리, 교환, 출판하기 위한 국제 규격이다. 핵심은 문서를 완성본 단위가 아니라 데이터 모듈단위로 분해하여 관리하는데 있으며, 이를 통해 정보의 재사용성과 문서 간 일관성을 높일 수 있다.

이 표준은 항공·방산 분야를 중심으로 발전하였으며, 복잡한 시스템과 다수의 협력사가 참여하는 제품 개발 환경에서 특히 유용하다. 항공기와 같은 대형 시스템은 부품과 서브시스템이 여러 공급망에서 개별적으로 생산·조립되므로, 기술정보를 표준화된 구조로 관리하고 변경사항을 신속히 반영할 수 있는 체계가 필수적이다.

기존의 기술문서 방식은 문서 전체를 개별적으로 수정하는 경향이 있어, 설계 변경이 잦은 제품에서는 유지관리 효율이 낮았다. 반면 S1000D는 절차, 경고, 부품정보 등의 내용을 모듈 단위로 관리하므로, 일부 정보가 바뀌더라도 관련 문서 전체에 일관되게 반영할 수 있다.

본 연구에서 S1000D의 의미는 레거시 교범을 단순히 디지털화하는 것을 넘어, 교범 내용을 재사용 가능한 모듈로 구조화하고 상황에 따라 재구성할 수 있는 기반을 제공한다는 점에 있다. 따라서 S1000D는 본 논문이 지향하는 상황 맞춤형 실시간 유지보수 가이드를 구현하기 위한 핵심 표준이라 할 수 있다.

## 3. Methodology

Fig 1는 본 논문의 전반적인 연구 흐름을 보여준다. Single-step 변환은 부품 코드와 같은 세부적인 요소를 반영하지 못하였다. 또한 전체적인 S1000D 양식의 XML 구조가 개별적으로 반영되어 계층 구조를 지키지 못하거나, 교범의 길이가 길어질 경우 입력 용량 초과로 인해 변환이 이루어지지 않았다. 따라서

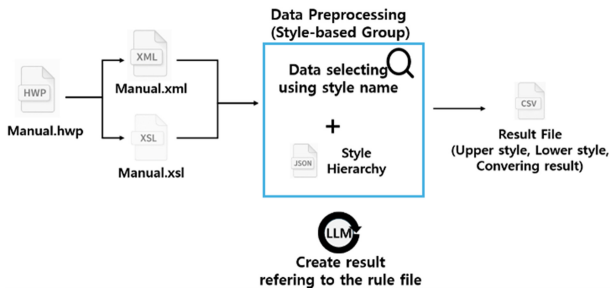


Fig. 1 Procedure of automatic conversion from local instructions into S1000D standardized documents

이를 해결하기 위해 Multistep 변환 방식을 제안한다. Multistep 변환 방식 내 입력량 조절 과정을 통해 기존 입력량 초과 오류를 해결하고, 선택적 규칙 적용 및 프롬프트 선택을 통해 잘못된 규칙을 적용시키지 않게 하여 정확도 향상하는 효과를 얻을 수 있다. 먼저 기존 비정형 교범 내에 저장된 스타일 정보를 추출한다. 한글파일을 XML 파일로 저장 시 파일 내 정보가 저장되는 XML 파일과 스타일 정보가 저장되는 XSL 파일로 나누어지며, XSL 파일에서 스타일에 관한 정보를 얻을 수 있다. 이는 기존에 정의된 스타일간 계층 구조 자료를 이용해 특정 계층에 속한 정보를 뽑는데 사용된다. 추출된 계층 정보를 사용해 교범 내 해당 계층에 속한 정보들만 모두 정해진 양식에 맞추어 정리한다. 정리된 정보들과 계층에 해당하는 규칙 파일을 LLM에게 주어 각각 알맞은 S1000D 양식 문서 구조로 변환시킨다.

3.1. 스타일 기반의 기존 교범의 계층 구조 추출

이 단계에서는 기존 기술 교범으로부터 문서의 계층 구조를 추출하는 것이다. 일반적으로 한글 형식 기술 교범은 제목, 소제목, 본문, 주의사항, 표 항목 등 서로 다른 정보 유형을 스타일 정보로 구분하고 있다[9]. 이러한 스타일 정보는 단순한 편집 요소가 아니라, 문서 내 각 내용이 어떤 역할과 수준을 가지는지를 나타내는 중요한 구조적 단서가 된다. 따라서 본 연구에서는 교범 한글 문서에 포함된 스타일 정보를 분석하여 문서의 계층 구조를 파악하고, 각 스타일에 해당하는 내용을 추출하는 방식을 채택한다

각 문단에 적용된 스타일을 식별한 후, 이를 기준으로 상위 항목과 하위 항목의 관계를 추적하여 문서의 계층 구조를 생성한다. 예를 들어, 제목, 소제목, 개요, 개요 본문, 세부 절차, 경고문, 주의문 등은 서로 다른 스타일로 구분되며, 이를 통해 원문 교범의 계층별 정보를 분리할 수 있다. Fig. 2에서는 실제 기존 교범에서 스타일 구조에 따라, 개요3-개요3 본문 스타일에 해당하는 내용이 개요2 스타일의 하위 항목으로 들어가는 것을 확인할 수 있다. 이처럼 원문 교범에서는 정보 간 관계를 각자의 스타일 정보를 통해 확인할 수 있다.

이를 통해 기존 교범 원문 내의 정보를 계층 구조를 갖는 구조화된 정보 집합으로 정리할 수 있다. 이러한 전처리 단계는 이후 S1000D 변환 단계에서 문서 내 계층 구조 파악 및 단계별 변환과정에서 사용된다.

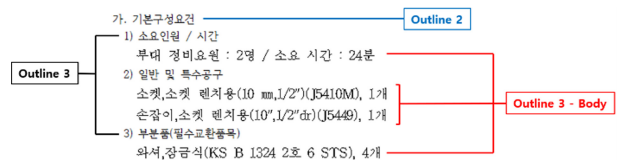


Fig. 2 Hierarchical structure of local instruction

Upper Content	예비절차	Total Converting Example	<pre> &lt;reqCondIm&gt; &lt;reqCond&gt;안테나세트 고각 0° 설정을 실시한다. &lt;/reqCond&gt; &lt;dmRefIdent&gt; &lt;dmCode modelIdentCode="Product" systemDiffCode="AAA" systemCode="00" subSystemCode="0" subSubSystemCode="0" assyCode="0000" disassyCode="00" disassyCodeVariant="A" infoCode="920" infoCodeVariant="A" itemLocationCode="A7" &lt;/dmRefIdent&gt; &lt;/dmRef&gt; &lt;/reqCondIm&gt;                     </pre>
Upper Content Tag	<reqCondGroup> </reqCondGroup>		
Lower Content Example	안테나세트 고각 0° 설정을 실시한다. (Product AAA-00-00 0000-00A-920A-A)		
No Children	<noConds/>		
Upper Content	일반 및 특수 장비	Total Converting Example	<pre> &lt;supportEquipDescr&gt; &lt;name&gt;소켓,소켓 렌치용(7 mm,1/4")&lt;/name&gt; &lt;idNumber&gt; &lt;manufacturerCode/&gt; &lt;partAndSerialNumber&gt; &lt;partNumber&gt;14707M&lt;/partNumber&gt; &lt;/partAndSerialNumber&gt; &lt;/idNumber&gt; &lt;reqQuantity unitOfMeasure="개"&gt;1&lt;/reqQuantity&gt; &lt;/supportEquipDescr&gt;                     </pre>
Upper Content Tag	<reqSupportEquips> <supportEquipDescrGroup> </supportEquipDescrGroup> </reqSupportEquips>		
Lower Content Example	소켓,소켓 렌치용 (7 mm,1/4"),(14707M), 1개		
No Children	<reqSupportEquips> <noSupportEquips/> </reqSupportEquips>		

Fig. 3 Reference data mapping category XML tags definition to S1000D examples

3.2. 로컬 LLM을 사용한 규칙 기반 S1000D 양식 변환

이 단계에서는 앞서 추출한 구조화된 교범 내용을 S1000D 형식으로 변환한다. 이를 위해 본 연구에서는 외부 서버 의존 없이 로컬 환경에서 재현 가능하게 운용할 수 있는 로컬 LLM인 Qwen3-8B를 사용하였다. Qwen3는 오픈 웨이트 기반으로 로컬 배포가 가능하고, 다국어 지원성과 지시 이행 능력을 갖추고 있어 한국어 기술문서를 대상으로 한 규칙 기반 구조화 변환 작업에 적합하다. 또한 본 연구의 과제는 자유 생성보다 규칙 파일, XML 태그 정의, 예시 문서를 참고하여 입력 내용을 표준 XML 구조로 매핑하는 유도형 변환에 가깝기 때문에, 초대형 모델보다 계산 비용과 적용 편의성 측면에서 균형이 좋은 8B급 모델을 선택하였다. 규칙 파일에는 Fig. 3과 같이 카테고리별 XML 태그 정의와 기존 S1000D 예시 문서가 포함되며, 이는 LLM이 각 정보 단위를 어떤 표준 구조에 대응시켜야 하는지 판단하도록 돕는다. Fig. 3에서 하위 항목 내 일관되게 반영되어야 하는 내용은 노란색으로, 분리되어 반영되어야 하는 세부 내용은 초록색으로 표시하였다.

LLM은 추출된 교범 내용을 입력으로 받아 해당 정보의 이름 값이 어떤 범주에 속하는지를 해석하고, 이에 규칙 파일을 참고하여 적합한 S1000D XML 구조로 재구성한다. 이 과정에서

스타일 기반으로 복원된 문서 계층은 정보 간의 상하 관계와 문맥을 유지하며 계층별 사용할 규칙 파일을 정하는 역할을 한다. 규칙 파일은 출력 형식의 일관성과 표준 적합성을 보장하는 역할을 한다. 결과적으로 제안 방법은 비정형 기존 교범을 정형화된 S1000D 문서로 변환할 수 있도록 하며, 향후 문서의 재사용성과 유지관리 효율을 높일 수 있는 기반을 제공한다.

3.3. 스타일 단계별 다른 규칙 파일 및 프롬프트 적용

본 연구에서 사용된 LLM은 로컬 모델로 장문의 교범을 한번에 읽어서 S1000D 양식으로 변환 시 부적절한 결과를 주거나 일부만 변환하는 오류가 발생한다. 따라서 변환된 S1000D 파일은 개요의 수준에 따라 다른 형식의 결과물을 제공해야 한다. 따라서 개요별로 LLM에게 제공할 규칙 파일과 프롬프트에 포함될 변환 규칙을 따로 생성하고 제공한다.

4. Case Study

기존 교범은 고각 리미트스위치 모듈 부품에 관한 교범을 사용했다. 교범은 개요2 수준인 기본 구성 요건에 관한 내용을 포함하고 있으면 개요3-개요3 본문 구조에 대해 실험을 진행하였다.

4.1. 스타일별 계층 구조 파악 및 문서 제작

교범 한글 문서에 포함된 스타일별 계층 구조를 JSON형식의 파일로 정리한다. Head에는 스타일의 상위 항목의 이름이 들어가며, Body에는 위 항목의 아래 항목의 이름이 들어간다. 고각 리미트 스위치 모듈 부품 교범의 스타일 계층 구조는 Fig. 4와 같이 이루어져 있다.

4.2. 스타일별 교범 내용 분류 및 추출

본 연구의 실험에서는 고각 리미트스위치 모듈 부품 교범 내 비정형 정보가 포함된 스타일 정보를 선택했다. 개요3-개요3 본문은 Fig. 5에서 확인 가능하다.

1) 예비절차, 2) 소요인원 / 시간, 3) 일반 및 특수 공구, 4) 물자, 5) 부분품(필수교환품목)이 개요3에 해당하며 그 외 정보들이 개요3 본문에 해당한다.

스타일 간의 계층 구조에 따라 개요3을 상위 항목, 개요3 본문을 하위 항목으로 정의하며, 이는 각각 Head와 Value으로 할당된 JSON 파일의 형태로 변환한다. 고각 리미트스위치 모듈 부품 교범 내 개요3-개요3 본문에 해당하는 내용은 변환된다. 변환된 자료는 해당 스타일에 해당하는 내용만을 담고 있어 추후 규칙 파일 탐색이나 LLM을 활용한 S1000D 변환에 효율성을 높인다.

4.3. LLM을 활용한 S1000D 변환

추출한 정보를 레벨별 S1000D 양식 변환 규칙 파일과 함께

- 1) 예비절차
  - 가) 안테나세트 고각 0° 설정을 실시한다.  
(PRODUCT-AAA-00-00-0000-00A-920A-A)
  - 나) 안테나세트 선회 0° 설정을 실시한다.  
(PRODUCT-AAA-00-00-0001-00A-920A-A)
- 2) 소요인원 / 시간  
부대 정비요원 : 2명 / 소요 시간 : 36분
- 3) 일반 및 특수 공구  
소켓,소켓 렌치용(7 mm,1/4")(J4707M), 1개  
손잡이,소켓 렌치용(5",1/4"dr)(J4749), 1개  
소켓,소켓 렌치용(13 mm,1/2")(J5413M), 1개  
손잡이,소켓 렌치용(10",1/2"dr)(J5449), 2개  
소켓,소켓 렌치용(36 mm,1/2"dr,12 pt)(J5336M), 2개  
가열 총,전기식(HL1820S)
- 4) 물자  
접착제(LOCTITE 243)
- 5) 부분품(필수교환품목)  
와셔,잠금식(KS B 1324 2호 8 STS), 2개  
와셔,잠금식(KS B 1324 2호 4 STS), 4개

Fig. 4 Hierarchical structure of part manual instruction for high-limited switch module

```

-----
head: 부모 태그
value: 자식 태그
-----
[
  {
    "head": "예비절차",
    "value": [
      "안테나세트 고각 0° 설정을 실시한다.\n(LSAM-AAA-00-00-0000-00A-920A-A)",
      "안테나세트 고각 0° 설정을 실시한다.\n(PRODUCT-AAA-00-00-0000-00A-920A-A)",
      "안테나세트 선회 0° 설정을 실시한다.\n(PRODUCT-AAA-00-00-0001-00A-920A-A)"
    ]
  },
  {
    "head": "소요인원 / 시간",
    "value": [
      "부대 정비요원 : 2명 / 소요 시간 : 36분"
    ]
  },
  {
    "head": "일반 및 특수 공구",
    "value": [
      "소켓,소켓 렌치용(7 mm,1/4")(J4707M), 1개",
      "손잡이,소켓 렌치용(5",1/4"dr)(J4749), 1개",
      "소켓,소켓 렌치용(13 mm,1/2")(J5413M), 1개",
      "손잡이,소켓 렌치용(10",1/2"dr)(J5449), 2개",
      "소켓,소켓 렌치용(36 mm,1/2"dr,12 pt)(J5336M), 2개",
      "가열 총,전기식(HL1820S)"
    ]
  }
]
-----
    
```

Fig. 5 Main contents of a specific section in a local instruction

정해진 프롬프트를 LLM에게 입력한다. Fig. 6에는 연구에서 사용된 개요 3에서 적용되는 프롬프트이다. 개요3에서 변환 시 사용되는 주요 규칙 Table 1과 같다.

XML Parsing Rule은 앞서 불러온 개요3 수준에 해당하는 룰 파일을 불러와 적용한다. 모든 정보가 적용된 프롬프트를 받은 LLM은 기존 교범 자료를 S1000D 양식으로 변환시킨다. Fig. 7과 같은 결과가 나오며 결과 파일은 CSV 형식으로 나오며 입력값 별로 출력값을 갖는다.

이 외에 타 교범을 포함하여 총 10개의 상위 항목과 17개의 하위 항목을 통해 실험을 하였다. 스타일을 통해 분리한 각 항목 별로 전문가가 변환한 파일과 비교를 통해 완전히 동일한 경우 변환 성공으로 판단하였다. 그 결과 상위 항목은 10개 전부, 하위 항목은 16개가 성공적으로 변환되었다. 이를 Table 2에서 확인할 수 있다. 상위 항목과 같은 간단한 변환은 Single-step에서

Return Only the final XML element. Do not explain analyzing Procedure or any other text.  
 \*\*\*Analyze the input data and find the appropriate XML tag using the [Xml Parsing Rule].\*\*\*

[Required constraints]  
 \*\*\*Return ONLY the final XML element. Do NOT return any other text under any circumstance.\*\*\*

[Input Data]  
 head\_data: {head}  
 value\_data: {value}

[Rules]  
 1. Match 'head\_data' with the most similar value in the 'ParentContent' column.  
 2. Use the corresponding 'Mapping XML' value as the parent XML tag.  
 ...

[Xml Parsing Example]  
 {xml\_parsing\_example}

Output Requirements:  
 - Return ONLY the final XML element. The value must follow the format defined in the 'Format' column. Do NOT explain your reasoning.  
 - Do NOT output any text except the XML.  
 - If the output contains the literal string W"nW", you must remove it and replace it with an actual line break.  
 - Must wrapped by the parent XML tag.

Fig. 6 Input prompt of large language model to convert the instruction

Table 1 Major rules to convert the local instruction into S1000D template

Rule
1 Find the parent item value in the data from the first column of the rule file.
2 The child item XML tag must be enclosed within the parent item XML tag.
3 Find the child item value in the data from the "Internal Information" column of the rule file.
4 Determine the format by referring to the example data in the same row as the data found in Rule 3.
5 If the child item value is empty, apply the value from the "No Data" column as it is.
6 The converted XML tag information must be enclosed within the parent item XML tag found in Rule 2, and the parent item XML tag must be used only once.
7 Even if there is only one child item value, it must still be enclosed within the parent item XML tag found in Rule 2.
8 If the parent item XML tag found in Rule 2 consists of multiple tags, such as <parent1><parent2> </parent2></parent1>, all of those tags must be treated as parent item tags.
9 If the item information mentions a quantity, such as "n units," add the following tag inside the internal tag: <reqQuantity unitOfMeasure = "?">n</reqQuantity>

Table 2 Summary of conversion results by hierarchical level

	Content level	Total number of items	Success number of items
Single-step convert	Upper level	10	9
	Lower level	17	6
Multi-step convert	Upper level	10	10
	Lower level	17	16

무리없이 변환했지만 제품 코드나 개수 등 한 태그 내에 항목 내용이 나뉘어져 들어가야 하는 경우, Multi-step에서 현저히 좋은 결과를 보였다.

Input Value	Output Value
1) 예비일자 가) 안테나세트 고각 0° 설정을 실시한다. (PRODUCT-AAA-00-00-0000-00A-920A-A) 나) 안테나세트 선회 0° 설정을 실시한다. (PRODUCT-AAA-00-00-0001-00A-920A-A)	<reqCondGroup> <reqCondDm> <reqCond>안테나세트 고각 0° 설정을 실시한다.</reqCond> <dmRef> <dmRefIdent> <dmCode modellIdentCode="PRODUCT" systemDiffCode="AAA" systemCode="00" subSystemCode="0" subSubSystemCode="0" assyCode="0000" disassyCode="00" disassyCodeVariant="A" infoCode="920" infoCodeVariant="A" itemLocationCode="A"/> </dmRefIdent> </dmRef> </reqCondDm> <reqCondDm> <reqCond>안테나세트 선회 0° 설정을 실시한다.</reqCond> <dmRef> <dmRefIdent> <dmCode modellIdentCode="PRODUCT" systemDiffCode="AAA" systemCode="00" subSystemCode="0" subSubSystemCode="0" assyCode="0001" disassyCode="00" disassyCodeVariant="A" infoCode="920" infoCodeVariant="A" itemLocationCode="A"/> </dmRefIdent> </dmRef> </reqCondDm> </reqCondGroup>
가) 일반 및 특수 공구 소켓,소켓 렌치용(7 mm,1/4") (J4707M), 1개 손잡이,소켓 렌치용(5",1/4" dr) (J4749), 1개 소켓,소켓 렌치용(13 mm,1/2") (J5413M), 1개	<reqSupportEquips> <supportEquipDescrGroup> <supportEquipDescr> <name>소켓,소켓 렌치용(7 mm,1/4")</name> <identNumber> <manufacturerCode/> <partAndSerialNumber> <partNumber>J4707M</partNumber> </partAndSerialNumber> </identNumber> <reqQuantity unitOfMeasure="개">1</reqQuantity> </supportEquipDescr> <supportEquipDescr> <name>손잡이,소켓 렌치용(5",1/4" dr)</name> <identNumber> <manufacturerCode/> <partAndSerialNumber> <partNumber>J4749</partNumber> </partAndSerialNumber> </identNumber> <reqQuantity unitOfMeasure="개">1</reqQuantity> </supportEquipDescr> <supportEquipDescr> <name>소켓,소켓 렌치용(13 mm,1/2")</name> <identNumber> <manufacturerCode/> <partAndSerialNumber> <partNumber>J5413M</partNumber> </partAndSerialNumber> </identNumber> <reqQuantity unitOfMeasure="개">1</reqQuantity> </supportEquipDescr> </supportEquipDescrGroup> </reqSupportEquips>
3) 용접 접착제(LOCTITE 243)	<reqSupplies> <supplyDescr> <name>접착제</name> <identNumber> <manufacturerCode/> <partAndSerialNumber> <partNumber>LOCTITE 243</partNumber> </partAndSerialNumber> </identNumber> <reqQuantity unitOfMeasure="개">2</reqQuantity> </supplyDescr> </reqSupplies>

Fig. 7 Final output of automatic instruction converter

5. 토론 및 결론

본 논문에서는 레거시 정비교범을 재사용 가능한 S1000D 기반 디지털 자산으로 전환하기 위해, 한글 교범의 스타일 정보를

활용한 계층 구조 추출과 로컬 LLM의 규칙 기반 다단계 변환을 결합한 프레임워크를 제안하였다. 기존 기술문서 변환 방식이 수작업 중심의 단발성 변환에 머무르며 설계 변경과 지속적인 문서 개정에 유연하게 대응하지 못했던 반면, 본 연구는 교범 내부에 내재된 스타일 정보를 구조 해석의 단서로 활용함으로써 문서의 계층성과 의미 관계를 보존한 상태에서 S1000D 형식으로 변환할 수 있는 방법을 제시하였다.

제안 방법은 먼저 한글 교범 문서의 스타일 정보를 분석하여 제목, 개요, 본문, 주의사항 등 문서 요소 간의 상하 관계를 추출하고, 이를 구조화된 정보 집합으로 정리한다. 이후 추출된 내용을 계층별 규칙 파일, XML 태그 정의, S1000D 예시 문서, 프롬프트와 함께 로컬 LLM에 입력하여 S1000D XML 형식으로 변환하였다. 특히 장문의 교범을 한 번에 처리하지 않고 계층 수준에 따라 서로 다른 규칙 파일과 프롬프트를 적용하는 **Multistep** 변환 방식을 통해 출력 구조의 일관성과 표준 적합성을 높이고자 하였다. 사례연구 결과, 고각 리미트스위치 모듈 부품 교범으로부터 계층 구조를 유지한 S1000D 형식의 결과를 생성할 수 있었으며, 이를 통해 본 연구의 접근이 레거시 교범의 모듈화와 표준화에 실질적으로 활용될 수 있음을 확인하였다.

본 연구는 한글 교범의 구조 해석과 LLM 기반 표준 변환을 결합한 실용적 파이프라인을 제시하였다는 점에서 의의를 가지며, 향후 반복 개정이 필요한 유지보수 교범을 효율적으로 관리하고 상황 맞춤형 실시간 정비 가이드로 확장할 수 있는 기반을 제공한다. 이를 통해 레거시 정비교범의 표준화뿐만 아니라 다만 본 연구는 특정 형식의 교범과 제한된 사례를 대상으로 수행되었으므로, 다양한 문서 형식과 산업 분야에 대한 추가 검증이 필요하다. 향후에는 다양한 포맷의 S1000D XML 변환을 통해 표준화의 일반화 가능성을 확장하고자 한다. 또한 S1000D XML 자동 검증 기능을 보완하고, 설비 상태, 알람 이력, 작업 조건과 연계한 모듈 재구성을 통해 스마트 제조 환경에서 활용 가능한 문맥 인식형 유지보수 정보체계로 확장할 예정이다.

## ACKNOWLEDGEMENT

본 연구는 2025년도 한화시스템(주)의 재원을 지원받아 수행된 연구임.

## REFERENCES

1. Sigsgaard, K. V., Agergaard, J. K., Mortensen, N. H., Hansen, K. B., Ge, J., (2023), Modular maintenance instructions architecture (mmia), *Journal of Quality in Maintenance Engineering*, 29(5), 50-67.
2. Woods, C., French, T., Hodkiewicz, M., Bikaun, T., (2023), An ontology for maintenance procedure documentation, *Applied Ontology*, 18(2), 169-206.
3. Pham, D., Dimov, S., Peat, B., (2000), Intelligent product manuals, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 214(5), 411-419.
4. Setchi, R., Pham, D. T., Dimov, S. S., (2006), A methodology for developing intelligent product manuals, *Engineering Applications of Artificial Intelligence*, 19(6), 657-669.
5. Knezevic, J., (2015), Improving quality of maintenance through simplified technical English, *Journal of Quality in Maintenance Engineering*, 21(3), 250-257.
6. Dagdelen, J., Dunn, A., Lee, S., Walker, N., Rosen, A. S., Ceder, G., Persson, K. A., Jain, A., (2024), Structured information extraction from scientific text with large language models, *Nature Communications*, 15(1), 1418.
7. Li, D., Zhao, Y., Wang, Z., Jung, C., Zhang, Z., (2024), Large language model-driven structured output: A comprehensive benchmark and spatial data generation framework, *ISPRS International Journal of Geo-Information*, 13(11), 405.
8. Polak, M. P., Morgan, D., (2024), Extracting accurate materials data from research papers with conversational language models and prompt engineering, *Nature Communications*, 15(1), 1569.
9. Morris, K. C., (2010), A framework for xml schema naming and design rules development tools, *Computer Standards & Interfaces*, 32(4), 179-184.



### Jumyung Um

Jumyung Um received the B.S. degree in Mechanical Engineering and Information Communication Engineering from Sungkyunkwan University, Suwon, Republic of Korea, in 2003, and the Ph.D. degree in industrial and management engineering from Pohang University of Science and Technology, Republic of Korea, in 2012. From 2012 to 2014, he was a Postdoctoral Researcher with the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland. After leaving EPFL, he was a Research Associate with the University of Cambridge, until 2015, and a Senior Researcher of technology initiative with SmartFactoryKL, Germany, in 2018. Since 2018, he has been an Associate Professor with the Department of Industrial and Management System Engineering, Kyung Hee University, Yongin, Republic of Korea. His research interests include the usage of augmented reality for smart operators and plug-n-produce of modular factory systems. E-mail: jayum@khu.ac.kr



**Youngwoo An**

M.S. student in the Department of XR-AI Convergence, Kyung Hee University. His research interests include the usage of augmented reality for smart operators and digital twin technologies for smart manufacturing and intelligent robotic automation.

E-mail: [kamilin12@khu.ac.kr](mailto:kamilin12@khu.ac.kr)



**Seung Uk Lee**

Researcher at Hanwha Systems. His research interest is Integrated Product Support (IPS) engineering.

E-mail: [lsw0908.lee@hanwha.com](mailto:lsw0908.lee@hanwha.com)



**Seon Ung Heo**

Researcher at Hanwha Systems. His research interest is Integrated Product Support (IPS) engineering.

E-mail: [seonung.heo@hanwha.com](mailto:seonung.heo@hanwha.com)



**Yeong Tak Seo**

Researcher at Hanwha Systems. His research interest is Integrated Product Support (IPS) engineering.

E-mail: [yeongtak.seo@hanwha.com](mailto:yeongtak.seo@hanwha.com)