

2025.08.27

정비문서 기반 LLM 이용한 지식그래프 구축 및 품질 프로젝트 보고서 생성

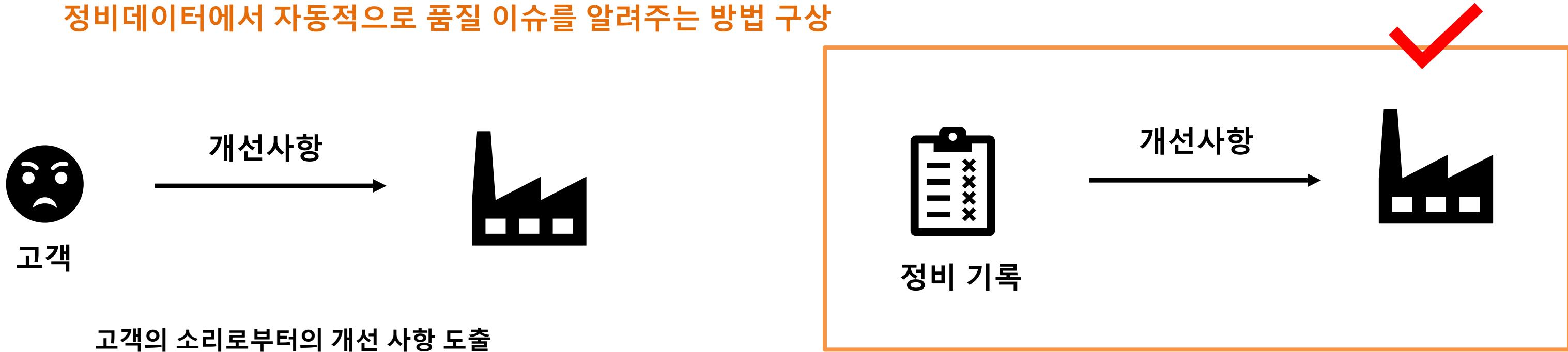
25-여름 국방 AI 해커톤 대회

서울과학기술대학교
산업정보시스템 - 김동원

01. 서론

정비데이터, LLM을 이용한 품질 개선사항 도출

정비데이터에서 자동적으로 품질 이슈를 알려주는 방법 구상



국방, 방산에서 고객의 소리를 이용한 품질 개선보다 정비데이터를 이용하는 방법이 적합하다고 판단하였으며, 정비데이터로부터 고장과 관련된 품질 개선 프로젝트를 도출하는 방법을 구상하였다.

02.

정비데이터

MainNet 항공기 정비데이터 셋 이용

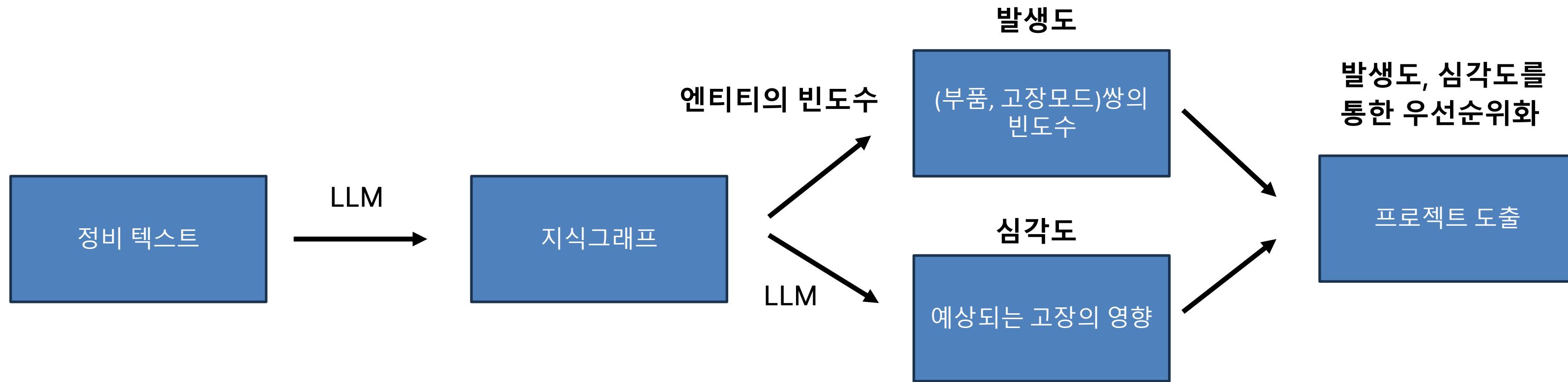
문제 - 조치 값이 있는 데이터를 이용

IDENT	PROBLEM	ACTION
100001	ENGINE IDLE OVERRIDE KILLED ENGINE.	TRIED TO ADJUST IDLE SEVERAL TIMES, WOULDN'T ADJUST.
100002	ENGINE IDLE OVERRIDE KILLED ENGINE.	REMOVED & REPLACED FUEL SERVO
100003	ENGINE IDLE OVERRIDE KILLED ENGINE.	A/C WAS RUN UP, SET IDLE SPEED, MIXTURE OK, NO LEAKS NO
100004	HAD ENGINE CHOKE & BRIEFLY LOSE POWER ON DEPARTURE. FULL THR	PERFORMED ENGINE RUN UP, FOUND CYL 2 LOWER PLUG FOULED.
100005	#2 & 4 CYL ROCKER COVER GASKETS ARE LEAKING.	REMOVED & REPLACED GASKETS.

사용한 문제/조치에 대한 데이터

03. 도출 과정

정비문서를 통해 일차적으로 지식그래프를 구축



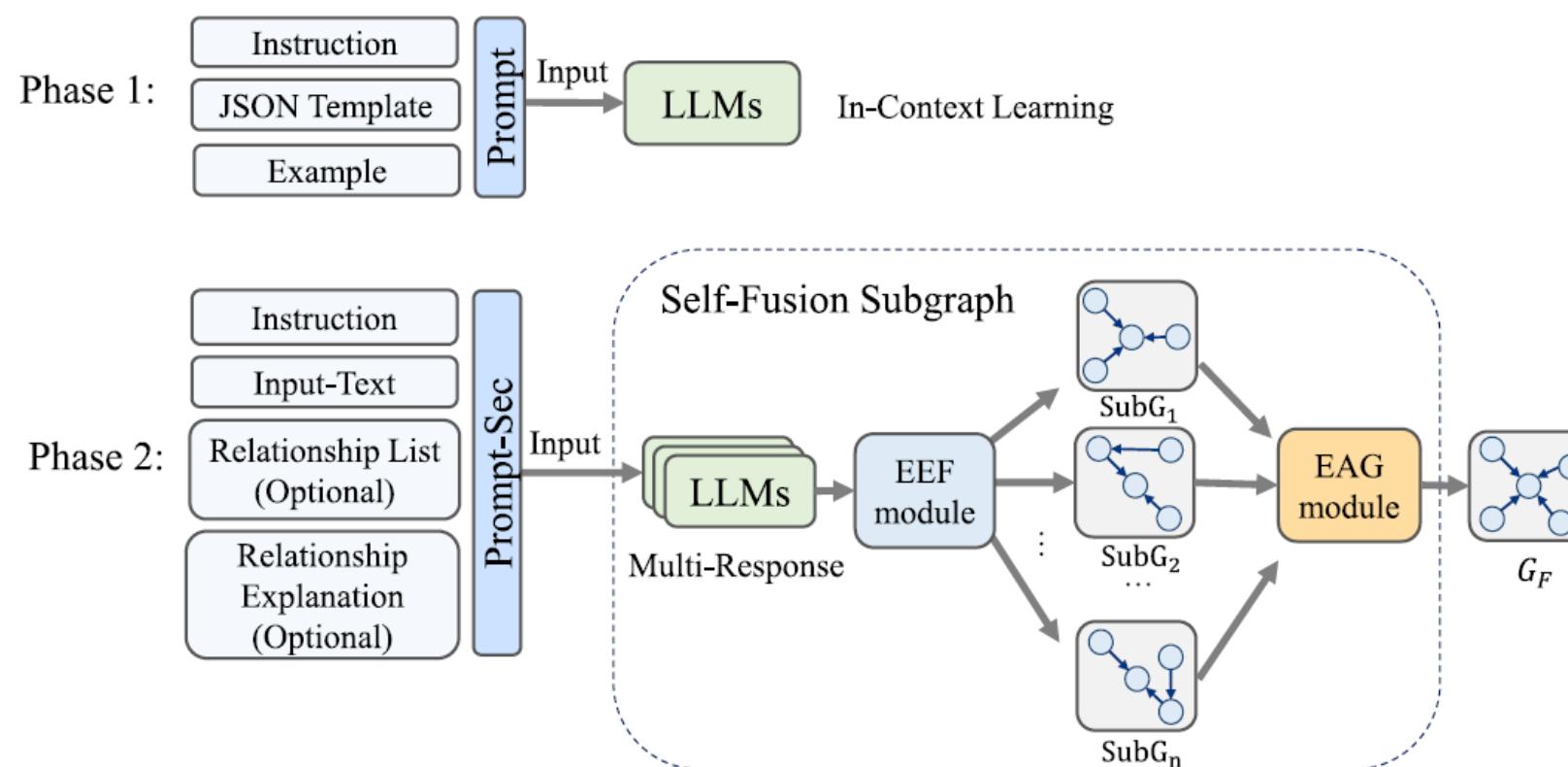
KG(지식그래프)를 사용한 목적

KG 구축 시 모든 정비 기록에서 고장모드, 부품, 영향, 조치방법, 원인에 맞는 내용을 추출할 수 있어서 동일한 고장모드 인식 간편하고 정비데이터 기반 QA시스템 등에서도 활용도가 높기 때문에 이용하였으며

KG를 바탕으로 각각의 서로 다른 고장에 대한 영향, 부품, 영향에 대한 내용을 찾아 이를 바탕으로 LLM을 이용한 추론이 더 구체적이고 정확할 것이라 판단하였습니다.

SF-GPT 방법을 통해 지식그래프 구축

범용적인 도메인 적용, 추출할 객체의 타입을 지정해줄 수 있는 장점을 가진 LLM을 이용한 지식그래프 구성 방법을 사용하였으며 SF-GPT논문에서 제안한 지식그래프 구축 방법을 사용하였다. 추가학습 필요X



두 개의 페이즈, 여러 번의 답변을 거쳐 LLM이 추출한 객체의 신뢰성을 높일 수 있는 방식이다.

첫번째 페이즈에서는 추출될 수 있는 모든 객체들을 출력을 받고, 두번째 페이즈에서는 지식그래프를 구성하는 트리플을 추출하여

첫 번째 페이즈에서의 일차적 필터링과 두 페이즈의 비교를 통한 필터링을 거쳐 가장 완벽한 지식그래프를 구축하는 방법이다.

(+ 객체 추출시 객체의 이름의 동의어를 여러 개 생성하여 이후 병합하는 과정)

SF-GPT 방법을 통해 지식그래프 구축

추출된 동의어로 객체를 병합하는 방법

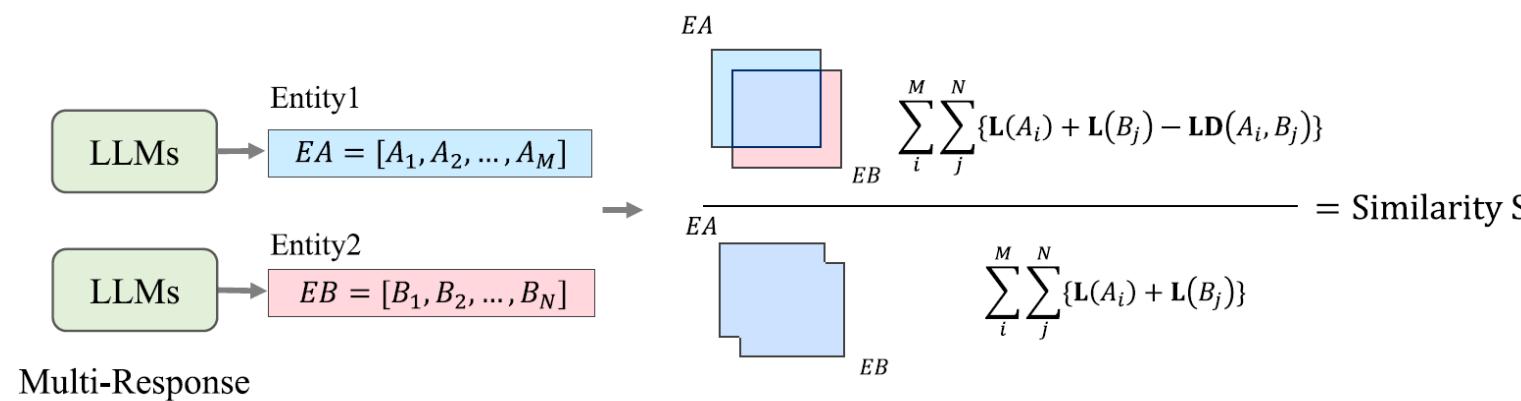
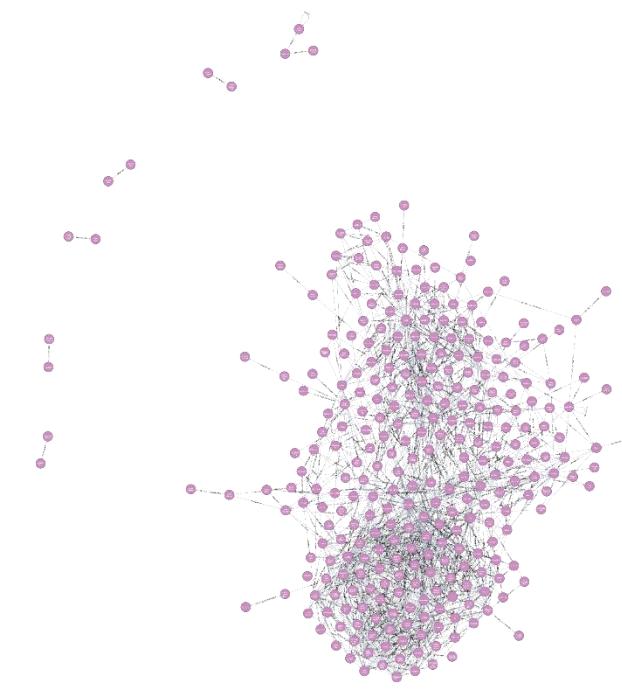


Fig. 5. Entity Alignment Algorithm Based on Entity Alias Generation.



추출 시 주어진 스키마

Schema:

Node Types (label):

- "FailureMode": specific types of failures (e.g., oil leakage, overheating, abnormal vibration)
- "Cause": root causes that trigger failure modes (e.g., seal wear, coolant shortage)
- "Effect": consequences resulting from failure modes (e.g., engine power loss, shutdown)
- "Component": physical parts related to the FailureMode or Cause (e.g., oil seal, engine, fuel pump)
- "Action": the maintenance or repair action taken to address the failure

Relationship Types (type):

- "HAS_CAUSE": (FailureMode → Cause)
- "RESULTS_IN": (FailureMode → Effect)
- "RELATED_TO_COMPONENT": (FailureMode → Component)
- "CAUSES_FAILUREMODE": (Cause → FailureMode)
- "AFFECTS_COMPONENT": (Effect → Component)
- "HANDLED_BY": (FailureMode → Action)

Link Prediction 모델을 이용한 자동적 평가 방법 사용

LLM으로 구축된 지식그래프를 평가하기 위해 보통 전문가가 직접 제작한 트리플을 기준으로 평가하지만 이번 프로젝트에서는 불가능하여 도메인에 적합성보다 지식그래프 자체의 완전성을 평가하는 방법을 사용하였다.

Link Prediction model이란?

지식그래프를 구성하는 트리플에서 head 혹은 tail에 대해 예측하는 모델.
지식그래프의 일부 트리플을 제거하고 이를 얼마나 잘 복원하는 가를 평가하는 작업

TransE모델

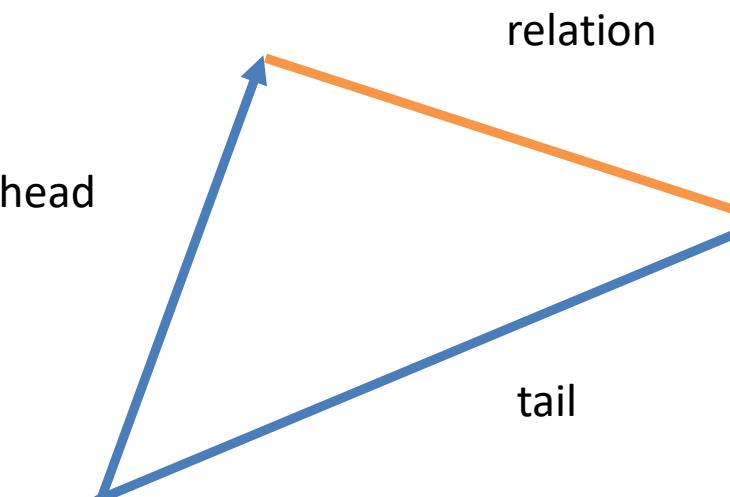
트리플의 head, tail을 저차원의 벡터로 매핑하고, relation을 벡터의 이동을 표현하여 $t \approx h + r$ 이 되도록 임베딩을 학습한다

※ 트리플 : (head, relation, tail)

(?, relation, tail), (head, relation, ?)을 예측하는 모델을

사용한 LP 모델?

transE모델을 사용하였으며 학습이 빠르며 적은 데이터셋에서도 좋은 성능을 보여서 사용하였다.



05.

KG 평가

Link Prediction 성능 = 지식그래프의 품질?

이 평가방법의 가정

- 고품질 KG는 일관성과 풍부한 패턴을 가지므로 일부 정보가 제거 되어도 잘 복원할 수 있음
- 저품질 KG는 노이즈나 불일치가 많아 예측 정확도가 낮음

정확도, 완전성을 직접 평가하지 못하고, KG의 자기 일관성을 평가하는 방법

데이터셋	MRR	Entity	relation	triple
FB15k237	0.174	14,541	237	310,116
Raw	0.048	3,601	6	8,183
Semi-Filtered	0.041	3,593	6	8,179
Filtered	0.070	1,894	6	5,291

데이터별 평가 결과

FB15k237은 대표적인 벤치마크 데이터셋이고, 모델의 성능이 논문에서 제시한 만큼 나오는지 확인,

필터링 과정에서 성능이 하락하기도 하지만 최종 데이터셋에서 가장 높은 성능을 보여 노이즈를 잘 제거한 것을 알 수 있다.

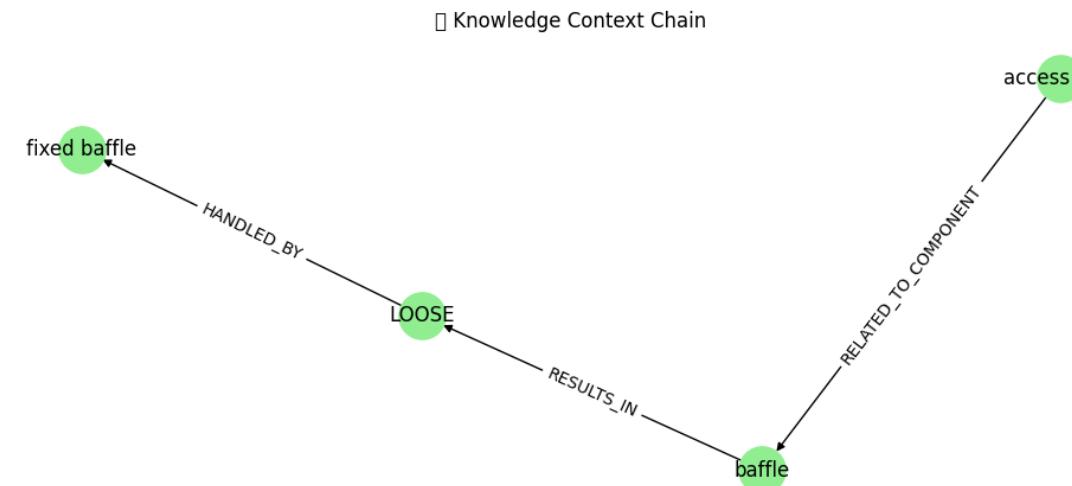
$$\text{MRR} = \frac{1}{N} \sum_{i=1}^N \frac{1}{\text{rank}_i}$$

품질 이슈 확인

고장모드 파악, 우선순위 도출

FailureMode 타입으로 추출된 고장에 대해서 지식그래프의 연관된 노드와 추출된 문장의 유사도를 바탕으로 동일 고장으로 파악하였으며, 각 고장의 발생빈도와 고장으로 인해 발생할 이벤트의 심각성 바탕으로 우선순위를 도출하였다.

고장 확인 – 발생도



서브그래프의 연관된 노드들과, 추출된 원문의 의미적 유사도를 바탕으로 동일한 고장모드를 인식하도록 하였으며 이를 바탕으로 발생빈도를 계산하여 발생도 점수를 부여하였다

점수	Probability of Failure	고장 발생률
10	고장이 거의 확실하게 발생함	1 in 2 (이하)
9		1 in 3
8	반복되는 동일한 고장	1 in 8
7		1 in 20
6	가끔 발생하는 고장	1 in 80
5		1 in 400
4		1 in 2,000
3	거의 발생하지 않음(비교적 고장이 없음)	1 in 15,000
2		1 in 150,000
1	발생하기 힘든 고장	1 in 150,000 (이상)

Chrysler, Ford, GM에서 사용하는 design-FMEA 치명도 기준표
출처-FMEA 발생도 점수표

자동차 기업에서 사용하는 FMEA 발생도 점수표를 참고하였다.

※ LEAK라는 이름의 고장에도 연관된 부품, 조치 방안에 따라 intakes leakage, oil leakage 등으로 의미가 달라질 수 있기 때문에

06.

품질 이슈 확인

고장모드 파악, 우선순위 도출

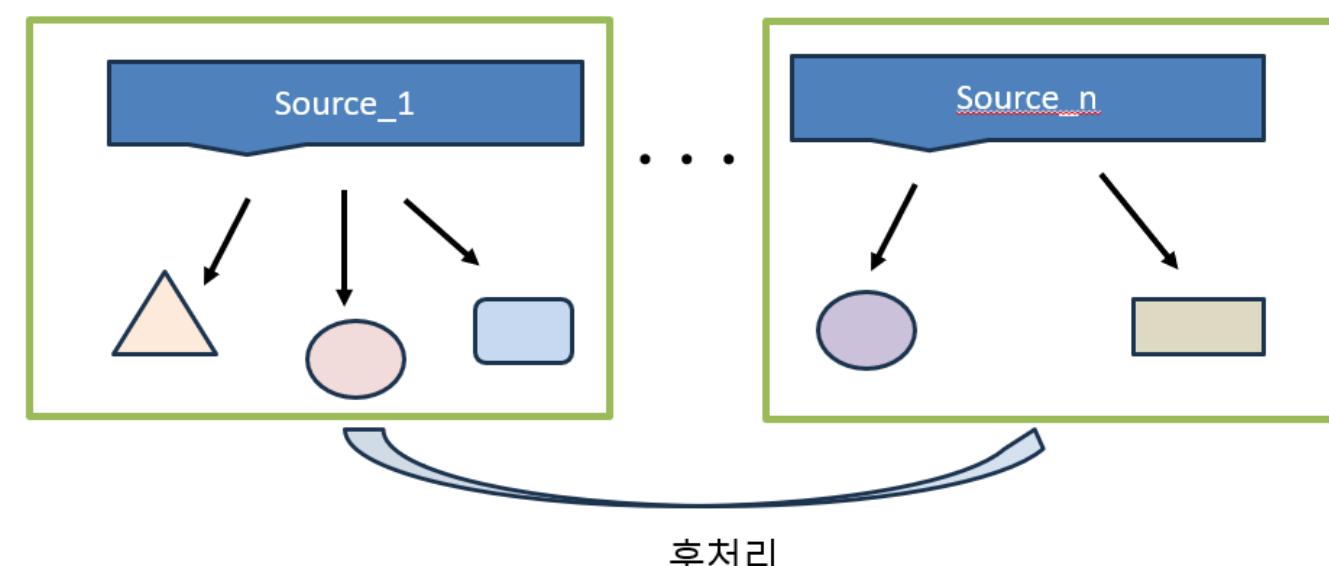
심각도

고장에 대한 서브그래프를 구성하여 LLM을 통해 고장의 영향을 추론

[Source] #4 SPARK PLUG ACCESS HOLE ON cylinder baffle REQUIRES PATCH, so the technician installed patch.
Failure mode: access hole | Related components: baffle | Handled by: patched

[Source] #4 INTAKE rocker cover gasket LEAKING, so the technician removed & replaced #4 cylinder intake rocker cover gasket.
Failure mode: gasket, LEAK | Related components: 4 | Effects: LEAK | Handled by: gasket swap

[Source] #3 & 4 ROCKER COVERS LEAKING, so the technician removed & replaced #3 & 4 cylinder rocker cover gaskets.
Failure mode: #2 & 4, LEAK | Related components: #3 & 4 cylinder | Effects: LEAK | Handled by: seal swap



분류	설명	설명 (한글 요약)
Category I - Catastrophic (파국적)	A failure which may cause death or weapon system loss (i.e., aircraft, tank, missile, ship, etc.)	사망 또는 항공기, 전차, 미사일, 함선 등 무기 시스템의 전체 손실로 이어질 수 있는 고장
Category II - Critical (치명적)	A failure which may cause severe injury, major property damage, or major system damage resulting in mission loss	중상, 주요 자산 피해, 또는 주요 시스템 손상을 야기하여 임무 실패로 이어질 수 있는 고장
Category III - Marginal (경미하지만 영향 있음)	A failure which may cause minor injury, minor property damage, or minor system damage resulting in mission delay or degradation	경미한 부상이나 손상으로 인해 임무 지연 또는 기능 저하를 유발할 수 있는 고장
Category IV - Minor (경미한)	A failure not serious enough to cause injury or damage, but results in unscheduled maintenance or repair	부상이나 손상은 없으나 예기치 않은 정비나 수리를 필요로 하는 고장

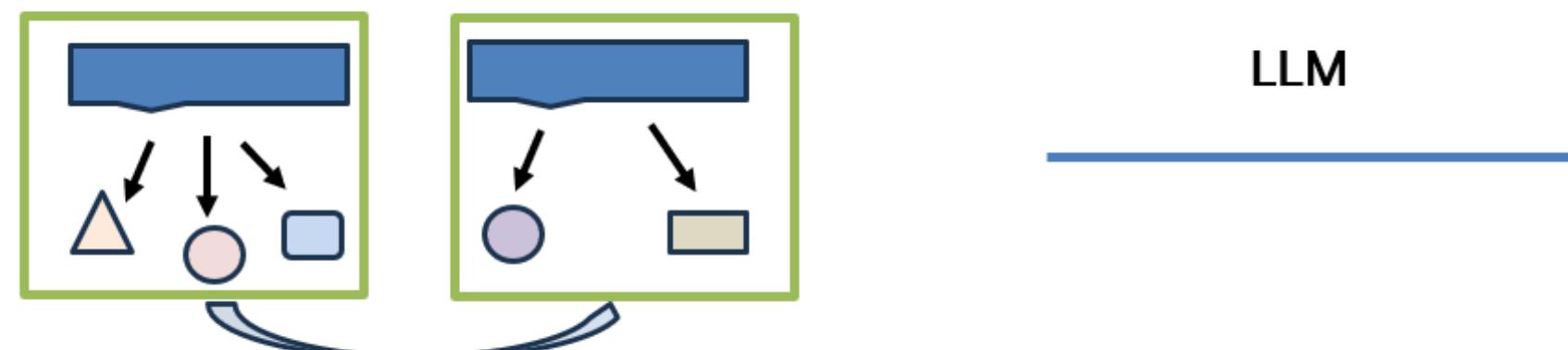
[Failure Mode, Effects, and Criticality Analysis \(FMECA\)](#)

고장의 영향의 심각도를 분류 기준에 따라 분류하여 심각도 점수를 부였으며
심각도* 발생도 점수를 기준으로 우선순위를 도출 하였다.

품질 이슈 확인

고장모드 파악, 우선순위 도출

고장의 영향 추론 출력 예시



답변:

```
{"source": 추론에 사용된 원문 문장,  
"severity_category": 심각도 분류 등급 이름,  
"severity_label": 등급 라벨,  
"rationale_short": 짧은 선정이유,  
}
```

보고서 생성

데이터프레임의 row값을 입력하여 결과 생성

우선순위 , 고장에 대한 최종 데이터프레임을 구성

head	tail	relation	occurrence	occurrence_score	severity_category	severity_score	priority	impact_sentences_all	rationale_short
5775	cover leak	R/H cylinder 3	RELATED_TO_COMPONENT	1	3	D	1	I'L/H aircraft engine cylinder #3 ROCKER COVER...	Leakage only with corrective action; no stated...
1840	absent plugs	new plugs	HANDLED_BY	2	3	D	1	I'SPARK PLUG cylender baffle PLUGS MISSING, so...	Only missing small parts; no stated impact bey...
4272	mixture issue	IDLE MIXTURE	RELATED_TO_COMPONENT	1	3	C	4	I'MIXTURE engine power control NOT HITTING FUL...	Clear operational limitation but no loss or in...
80	LEAK	seal swap	HANDLED_BY	39	5	D	1	I'#1 & 2 ROCKER COVERS LEAKING, so the technic...	Leaks noted without stated damage or operation...
3133	low power	cylinder 4 left/right	RELATED_TO_COMPONENT	2	3	C	4	I"#4 left or right aircraft cylinders COMPRESS...	Performance degradation implied; no injury or ...

각 row 값을 보고서 생성에 입력으로 사용하고, 추가로 기존 지식그래프에서 고장의 원인에 대한 정보가 있다면 추가하여 입력된다

보고서 생성

프롬프트 및 출력 예시

LLM의 프롬프트 구성

```
return f"""
당신은 항공기 정비/품질 전문가입니다. 주어진 데이터를 바탕으로 **품질 개선 프로젝트 제안서**를 작성하세요.

[입력 데이터]
- {head_role}: {head}
- {tail_role}: {tail}
- 관계(Relation): {relation}
- 발생빈도(occurrence): {occurrence_raw}
- 발생빈도 점수(occurrence_score): {occ}
- 고장 영향 등급(severity_category): {cat}
- 고장 영향 점수(severity_score): {sev}
- 우선순위(priority): {pri}
- 데이터 기반 예상 원인(cause_candidates): {cause_str}
- 관련 정비기록(impact): {impact}
- 고장의 예상 영향(Effect_rationale): {rationale}

[작성 지침]
1) 최근 발생 건수와 빈도
2) 발생사례 : 관련 정비기록(impact)에 등장하는 문장들을 요약
3) 고장 영향 설명: 고장의 예상 영향(Effect_rationale) 요약
4) 예상 원인: 먼저 **입력의 cause_candidates**에서 근거를 제시하고, 추가 가설이 필요하면 별도 단락으로 작성
   - 데이터에 없는 새로운 가설은 **"AI 생성 추정:"**으로 시작하는 문장으로 구분
5) 개선 방법 제안: 기술적/운영적 대책을 즉시/단기/중기로 구분, 프로젝트 KPI 포함

형식: 한국어, 간결한 불릿 위주, 250단어 이내.
""".strip()
```

품질 개선 프로젝트 제안서

1) 최근 발생 건수와 빈도

- 고장 모드: CRACK
- 발생 빈도: 3회
- 발생 빈도 점수: 4

2) 발생사례

- #2 엔진의 왼쪽 및 오른쪽 실린더에서 균열 발생
- 기술자가 항공기 정비 비행 작업 필요

3) 고장 영향 설명

- 균열이 있는 실린더는 주요 시스템 손상을 의미
- 교체가 필요하며, 이는 임무 순서로 이어질 수 있음

4) 예상 원인

- 입력의 cause_candidates:
 - FRONT, side baffle #1 and #3, baffling, #4 cylinder baffle
- AI 생성 추정:
 - 부품의 제조 결함 또는 설치 오류 가능성
 - 환경적 요인(온도 변화, 진동 등)도 원인으로 작용할 수 있음

5) 개선 방법 제안

- 즉시:
 - 균열 발생 부품의 즉각적인 교체 및 정비 기록 검토
- 단기:
 - 정기적인 점검 및 모니터링 프로세스 강화
 - 기술자 교육 프로그램 실시
- 중기:
 - 부품 설계 및 제조 공정 개선
 - 데이터 분석을 통한 예측 유지보수 시스템 도입

KPI

- 고장 발생률 50% 감소
- 정비 소요 시간 30% 단축
- 기술자 교육 완료율 100% 달성

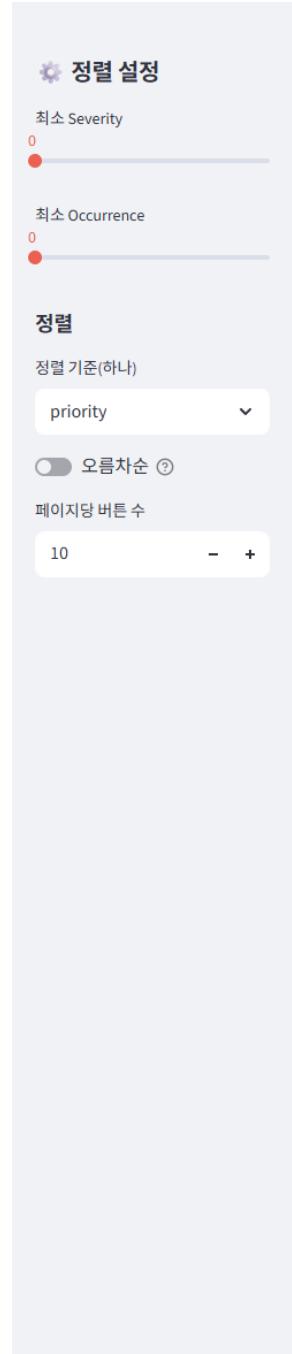
이 프로젝트는 항공기 안전성을 높이고, 운영 효율성을 개선하는 데 기여할 것입니다.

엔진의 Crack관련 고장

08.

대시보드 구성

Streamlit을 통해 대시보드를 구성



AI 보고서 생성 대시보드

⚡ AI 보고서 생성 버튼 목록

페이지

2

보고서 미리보기

최근 선택: ID 2250

품질 개선 프로젝트 제안서

1) 최근 발생 건수와 빈도

- 고장 모드: CRACK
- 발생 빈도: 3회
- 발생 빈도 점수: 4

2) 발생 사례

- #2 엔진의 원쪽 및 오른쪽 실린더에서 균열 발생
- 기술자가 항공기 정비 비행 작업 필요

3) 고장 영향 설명

- 균열이 있는 실린더는 주요 시스템 손상을 의미
- 교체가 필요하며, 이는 임무 손실로 이어질 수 있음

4) 예상 원인

- 입력의 cause_candidates:
 - FRONT, side baffle #1 and #3, baffling, #4 cylinder baffle
- AI 생성 추정:
 - 부품의 제조 결함 또는 설치 오류 가능성
 - 환경적 요인(온도 변화, 진동 등)도 원인으로 작용할 수 있음

5) 개선 방법 제안

- 즉시:
 - 균열 발생 부품의 즉각적인 교체 및 정비 기록 검토
- 단기:
 - 정기적인 점검 및 모니터링 프로세스 강화
 - 기술자 교육 프로그램 실시
- 중기:
 - 부품 설계 및 제조 공정 개선

고장에 대한 정보와 LLM이 생성한 보고서를 확인 가능

+ 우선순위, 발생도, 심각도에 따른 정렬기준을 선택할 수 있다.

결론·한계

1. LLM을 이용한 지식그래프를 구축 및 품질 이슈 확인 방법을 구상
2. 지식그래프, 생성된 보고서에 대한 전문가의 평가 부재한 것이 한계
3. LLM의 제품에, 프로세스에 대한 이해가 부족
4. 정비데이터가 특정 제품이 아닌 항공기 전반에 대한 데이터

2025.08.27

감사합니다