LCA 실무 AX 포트폴리오 · 2026

LCI-Match
BOM → 국가 LCI DB
AI 자동 매칭

목록 자재 입력 값마다 국가 LCI DB의 데이터셋 후보를 찾아 주고,
근거를 보여 준 뒤 사람이 승인하는 반자동 매칭 도구

이정빈 · 2026.10



LCI-Match 포트폴리오

CONTENTS

01

- LCA와 문제 정의
- LCA 개요와 규제 동향
- Pain Point와 접근 방식

02

- 평가 기준과 화면
- 신뢰도 3단계와 보조 규칙
- 화면 흐름

03

- 검증 실험 설계
- 가상 BOM 2묶음·3단계 난이도
- 채점 기준과 측정 지표

04

- 1차 실험과 개선
- 1위 정답률·신뢰도 문제
- 시도한 6가지와 채택 4가지

05

- 최종 결과와 고도화
- 가상 BOM 2000행 결과
- 한계와 고도화 계획

LCA 개요

LCA는 국제 표준(ISO 14040·14044)으로 절차가 정해진 4단계 평가이며, 수출·조달 규제에 대응하는 기본 도구가 되고 있음

LCA (Life Cycle Assessment, 전과정평가)

개요	제품의 원료 채취부터 생산·사용·폐기까지 전 과정의 자원 투입과 배출을 모아 환경 영향을 수치로 평가하는 방법
가이드라인	ISO 14040·14044 (원칙·절차) · ISO 14067 (탄소발자국) ISO 14025 (환경성적표지·EPD)
활용 규제	EU 배터리법 (탄소발자국 신고) · EU 에코디자인 규정 (제품 여권) EU 탄소국경조정제도 (CBAM) · 국내 환경성적표지

목적 및 범위 정의

- 평가 대상, 기능 단위, 시스템 경계를 정한다

전과정 목록분석

- 자원·에너지 투입과 배출 데이터를 모으고 LCI DB와 연결한다
- LCI-Match가 돕는 단계

전과정 영향평가

- 모은 목록을 기후변화 등 영향 범주로 환산한다

전과정 해석

- 결과의 의미와 한계를 해석하고 개선점을 찾는다

LCA 평가 대응 기업 현황

환경성적표지 인증

- 기업 546 → 987개사
- 제품 4,807 → 10,231개

2021 → 2025년, 5년 사이 약 2배

LCA 정부 대응 현황

국가 LCI DB 확대

- 2022년부터 현장 데이터 기반으로 새로 개발, 2028년까지 1,000개 구축 목표

해외 상호인정

- 2024년 노르웨이와 환경성적표지 상호인정 협정 체결, 유럽 수출 제품에 활용

출처: 한국환경산업기술원 환경성적표지 인증 통계(2025), 데일리환경(2024.11.11), 임팩트온(2025.9.17). EU 배터리법 탄소발자국 신고는 하위 법령 지연으로 시행일 미정(2026.10 기준).

주목한 Pain Point

LCA 실무에서는 BOM의 행마다 맞는 LCI 데이터셋을 찾아 연결해야 한다. 이름이 제각각이고 비슷한 후보가 많아 시간이 많이 들어 AX가 필요한 영역이라고 생각함

1 의미는 같지만 이름이 다름

BOM

SPCC

t1.2 · kg · 0.57

LCI DB 후보

냉간압연강판
2022 · 환경부(KEITI)

이름에 SPCC 없음

'SPCC'는 강종 약어라서 이름 글자만 비교해서는 '냉간압연강판'을 찾기 어려움

2 같은 이름의 후보가 여럿

BOM

용융아연도금강판

0.8T · kg · 1.1

LCI DB 후보

용융아연도금강판
2022 · 환경부(KEITI) 2022

용융아연도금강판
2003 · 환경부 2003

이름이 같아도 연도·기관별로 따로 있어서, BOM만 보고는 선택하기 정하기 어려움

3 기준 단위가 맞지 않음

BOM

자동차용 케이블 12V

kg · 0.8

LCI DB 후보

자동차용 케이블(12~60V)
2024 · 환경부(KEITI)

기준 1 m

BOM은 무게(kg), 후보의 기준은 길이(m). 바꾸려면 m당 무게 같은 정보가 따로 필요함

후보 이름·연도·기준 단위는 국가 LCI DB에서 발췌. 출처: 한국환경산업기술원(KEITI) 국가 LCI DB

문제해결 접근 방식

후보를 좁히고 근거와 신뢰도를 함께 보여 준다. 최종 연결은 사람이 비교해 승인함

01 업로드·정리	02 AI 후보 검색	03 판정	04 최종 승인 처리	05 내보내기
BOM 양식(xls·xlsx·csv)을 읽고 품명·규격·단위를 정리 공정 이름이면 구분을 추정	뜻 검색(bge-m3)으로 후보 100개를 찾고, 사전·역할 구분·같은 이름 묶기로 순서를 정리	글자·뜻 순위와 뜻 점수로 신뢰도(높음·보통·낮음)를 매기고, '낮음'이면 '맞는 후보 없음'	후보 3개를 나란히 비교해 승인·반려·제외. 최종 결정은 사람이 한다	승인한 행만 XLSX·CSV로 내보내기

구현 아키텍처

PostgreSQL 16 + pgvector

Ollama bge-m3 임베딩

FastAPI

React 19 · Vite

LCI MATCH의 평가 기준 1/2

LCI MATCH가 추천한 후보를 얼마나 믿을 수 있는지 3단계 신뢰도로 알려 주고, 등급별 목표를 미리 정했음

글자 일치 순위 TF-IDF 모델
품명 글자가 얼마나 겹치는가

+

뜻 일치 순위 bge-m3 모델
품명의 뜻이 얼마나 가까운가

+

뜻 유사도 점수
0~1점 · 기준 0.55

▶

두 순위와 점수를 합쳐
신뢰도 3단계 설정

신뢰도 3단계 설정 기준

높음 가장 믿을 만함

아래 세 가지를 모두 만족

- 글자 — DB와 글자 일치 1위
- 뜻 — DB와 뜻 일치 1위
- 점수 — 뜻 유사도 점수 0.55 이상

보통 사람이 확인할 것

아래 중 하나에 해당

- a — 글자·뜻 순위가 둘 다 3위 안
- b — 한쪽만 1위 (다른 쪽도 상위 100위 안)
- c — 뜻 점수 0.55 미만 ('높음' 조건은 충족)

낮음 맞는 후보 없음으로 처리

위에 해당하지 않으면

- 예 — 한쪽 순위표에 아예 없음

화면에는 '맞는 후보 없음'으로 표시

신뢰도 목표 설정 ❯

정답률 90% 이상
'높음'의 정답률 (오답률 10%이하)

'높음' 비율 45% 이상
전체 중 '높음' 등급 행 비율

LCI MATCH의 평가 기준 2/2

신뢰도 점수만으로는 가릴 수 없는 오답을 보조 규칙 4가지로 줄인다. 규칙은 후보의 순서만 바꾸고 최종 승인은 사람이 한다.

왜 보조 규칙이 필요한가

- 검색 점수만으로는 약어·소재/공정·연도를 못 가린다
- 1차 실험의 실제 오답 유형에서 규칙을 만들었다
- 규칙은 후보를 지우지 않고 순서만 바꾼다

지키는 원칙

- 최종 연결은 언제나 사람이 승인한다
- 모든 규칙은 최종 평가 전에 고정했다
- 규칙이 적용된 이유를 후보 옆에 표시한다

① 약어 사전

하는 일

- 규격 기호·수지 약어를 국가 DB 이름으로 풀어 검색어에 붙인다

예시

- SPCC → 냉간압연강판
- PP → 폴리프로필렌

② 역할 구분

하는 일

- 소재 행이면 공정 데이터를, 공정 행이면 소재 데이터를 뒤로 보낸다

예시

- ABS 수지 펠릿
- → 'ABS 용착(공정)'을 뒤로

③ 같은 이름 묶기

하는 일

- 같은 이름의 여러 판은 하나로 묶고 최신 연도를 대표로 보여 준다

예시

- 용융아연도금강판
- 2003·2005·2022년판 → 2022년

④ 맞는 후보 없음

하는 일

- 신뢰도 '낮음'이면 추천 대신 '맞는 후보 없음' 표시

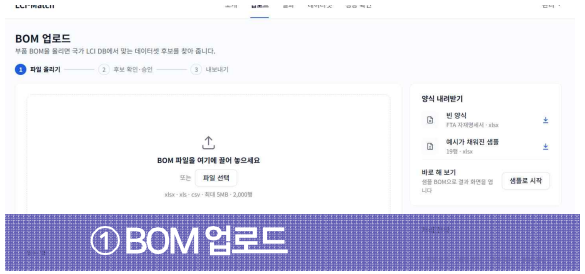
예시

- 분체도장 처리

→ DB에 도장 공정 없음

화면 프로세스

BOM 파일을 올리면 행마다 후보 3개와 신뢰도를 보여 주고, 사람이 근거를 비교해 승인한 행만 내보낸다.



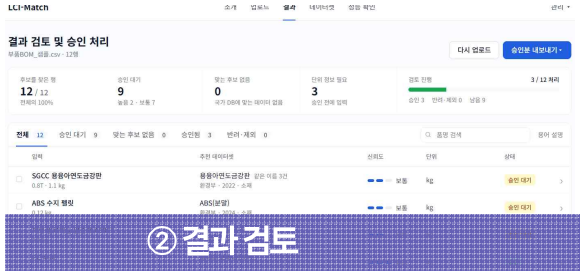
① BOM 업로드

FTA 양식을 그대로 올리면 품명·규격·단위 열을 자동 인식

엑셀·CSV, 한 번에 최대 2,000행

자재 목록 정리

소재/공정 구분이 비어 있으면 품명으로 추정



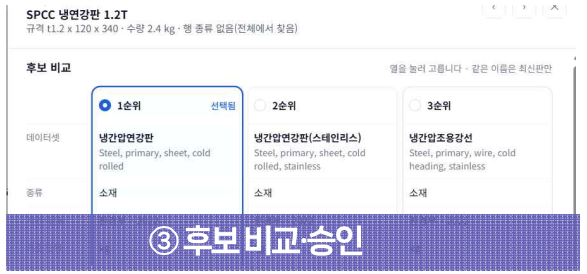
② 결과 검토

행마다 1위 후보와 신뢰도(높음·보통·낮음)를 한 줄로 표시

승인 대기·맞는 후보 없음·단위 확인 필요로 걸러 보기

확인할 행부터 보기

'높음'은 훑어보고, '보통'은 3개 중 고른다



③ 후보 비교·승인

후보 3개의 기관·연도·기준 단위·데이터 품질을 나란히 비교

승인한 행만 엑셀로 내보내기(환산값·참고 배출량 포함)

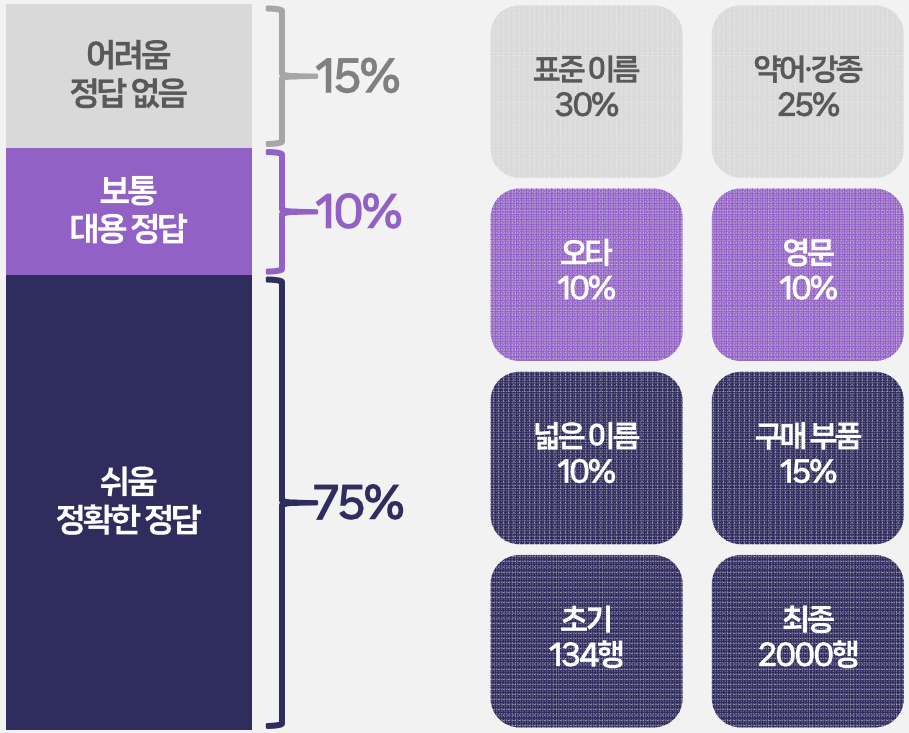
사람이 최종 승인

추천 근거(글자·뜻 순위, 사전)를 함께 표시

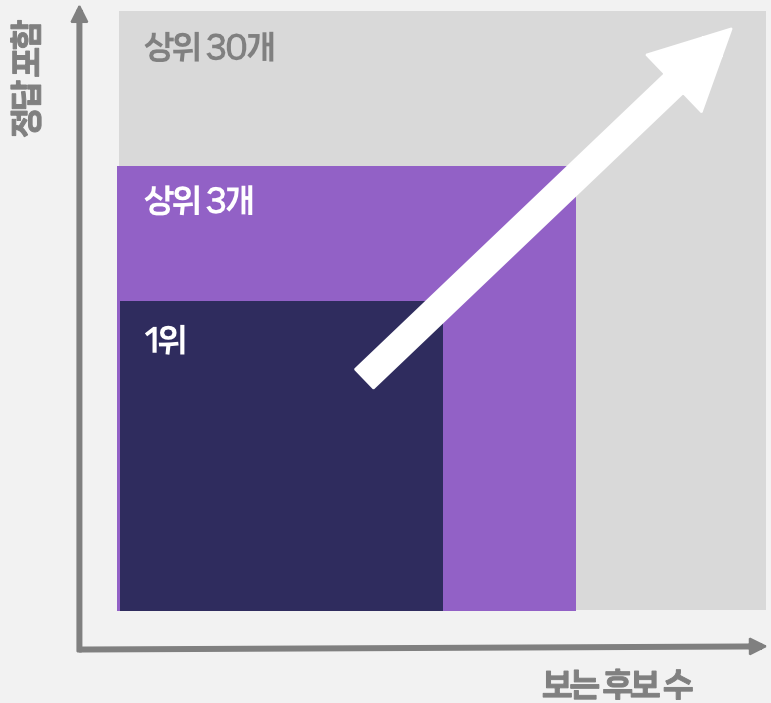
LCI MATCH 검증 실험 설정

가상 BOM을 튜닝용과 검증용으로 나눴다. 튜닝용 134행으로 규칙을 고르고, 규칙을 고정한 뒤 검증용 2000행을 한 번만 측정했다.

실험용 임시 BOM Data : 3단계 난이도



채점 기준과 기대 결과



- 채점 기준 ①: 같은 이름이면 정답
- 채점 기준 ②: 최신 연도 판까지 맞아야 정답
- 목표: '높음' 오답률 10% 이하
- 목표: '높음' 비율 45% 이상
- 지표: 1위·상위 3개 정답률
- 지표: '없음' 판정률, 95% 구간

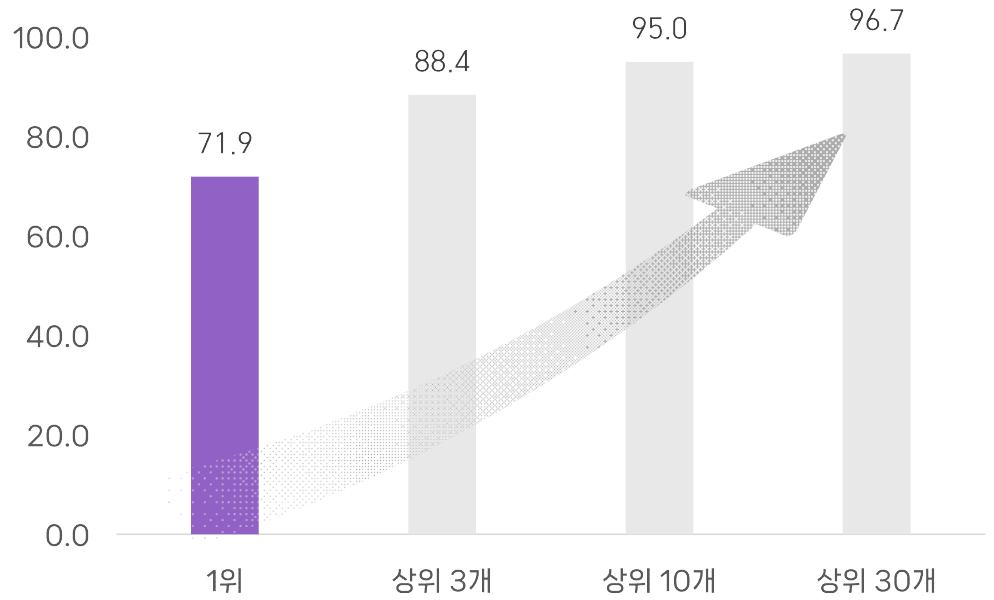
막대: 정답 유형 비율 · 타일: 입력 표기 비율(검증용 2000행 기준). 정답과 입력은 모두 가상으로 작성했음

1차 실험 결과 : 후보 1순위의 정답률 저하 문제

개선 전 방식으로 측정했다. 정답은 대부분 후보 안에 있었지만 1위가 아니었고, '높음' 정답률이 목표에 못 미쳤다.

① 1위 정답률	② 상위 3개 안 정답률	③ '높음' 정답률	④ '높음' 비율
71.9%	**88.4%**	**85.3%**	**50.7%**
정답이 있는 121행 중 87행	121행 중 107행	목표 90% 이상 · 미달	목표 45% 이상 · 달성

| 정답이 상위 몇 위 안에 있나 (%) |



정답은 대부분 후보 안에 있다

상위 10개 안 95.0%, 상위 30개 안 96.7%

하지만 1위로 올리지 못한다

1위가 틀린 34행 중 28행은 정답이 2~10위 → 문제 1/2

'높음'인데 틀린 행이 있다

'높음'으로 표시된 68행 중 10행이 오답 → 문제 2/2

①② 정답이 있는 121행 기준, ③④ 전체 134행 기준(정답이 없는 13행에 후보를 내면 오답).

1차 실험 결과 문제 1/2

1위가 틀린 행을 하나씩 열어 보니, 검색 점수만으로는 가릴 수 없는 세 가지 유형이 반복됐다.

① 소재·공정 혼동



입력: 소재



1위: 공정

이름이 비슷하면 소재와 공정 데이터를 구분하지 못했다.

예) ABS 수지 펠릿 → ABS 용착(공정)
예) PP 사출품 도어트림 → 필름 PP 압출

② 약어·다른 표기



입력: 약어



1위: 다른 자재

약어·관용 표기를 국가 DB의 정식 이름과 잇지 못했다.

예) 시너 → 시멘트
예) LLDPE 필름 그레이드 → 폐LDPE 재활용

③ 같은 이름의 옛 판



최신판: 3위



옛 판: 1위

같은 이름이 여러 해에 있으면 옛 판이 앞섰다.

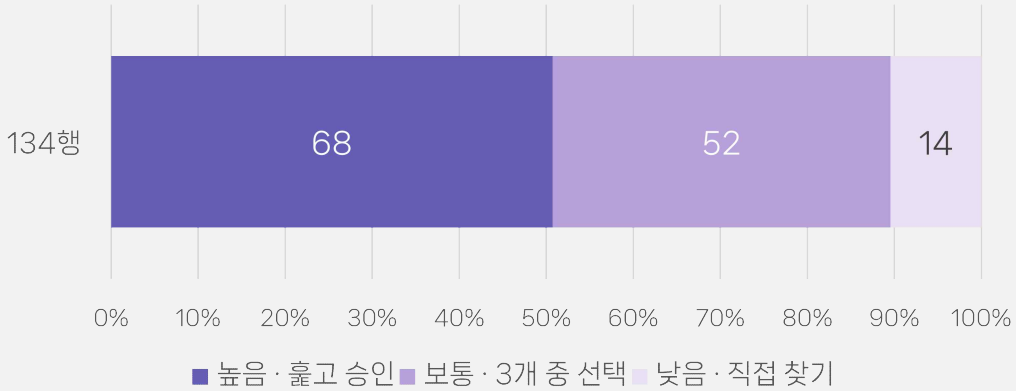
예) SGCC 용융아연도금강판 →
2003년판(톤 단위)이 1위, 2022년판은 3위

튜닝용 134행, 개선 전 방식에서 1위가 틀린 실제 행. 위 예시는 모두 개선 뒤(해결 방법 장) 1위가 정답으로 바뀌었음

1차 실험 결과 문제 2/2

신뢰도는 정답률이 아니라 사람이 얼마나 확인할지를 나누는 표시다. 그런데 바로 승인할 '높음'에서 10행이 틀렸다.

신뢰도별 행 수 = 사람이 할 일



③ '높음' 68행 → 정답 58행 · 오답 10행

'높음' 정답률 85.3% (목표 90% 이상 · 미달)

'높음'인데 틀린 사례

'높음' 오답 10행

시너

- 1위: 시멘트
- 관용 표기를 못 읽음
- 개선 후 해결

ABS 수지 펠릿

- 1위: ABS 용착(공정)
- 소재·공정 혼동
- 개선 후 해결

실리콘 고무 가스켓

- 1위: 실리콘 메탈
- DB에 없는 자재
- 개선 후에도 남음

AVSS 저압 전선

- 1위: 알루미늄 전선용 로드
- 전선 규격 코드
- 개선 후에도 남음

튜닝용 134행, 개선 전 방식. 정답이 없는 행에 후보를 낸 경우도 오답 처리

1차 실험 문제 해결 방법

오답 유형마다 해결책을 세워 튜닝용 134행에서 비교했다. 효과가 확인된 4가지만 채택하고, 성능이 오르지 않은 AI 모델 2가지는 버렸다.

문제 해결 접근 방법

1차 실험에서 찾은 문제

- 1위 오답
- 높음 오답
- 없음 판단

① 소재·공정 혼동
- ABS 수지 펠릿 → ABS 용착

② 약어·다른 표기
- 시너 → 시멘트

③ 같은 이름의 옛 판
- 용융아연도금강판 2003년판이 1위

시도한 방법

규칙 보완
- 약어 및 단어사전 구축 (①, ②)
- 역할 구분 (①)
- 같은 이름 묶기 (③)

검색·AI 모델
- 뜻 검색 중심 + 점수 조정
  - 글자·뜻 결합 대신 뜻 검색을 기본으로
  - '높음'은 뜻 점수 0.55 이상일 때만
- 재순위 AI · 소형 LLM 판정

결과

채택 : 규칙 3가지
- 1위 정답률 71.9% → 77.7%
- 새로 맞힘 7행 · 새로 틀림 0행

채택: 뜻 검색 + 점수 조정
- 1위 85.1% · '높음' 오답 6.6%
- ('높음' 비율 45.5%)

탈락: AI 모델 2가지
- 재순위: 1위 오히려 −0.8%p
- LLM: 5번 물으면 51~61% 바뀜

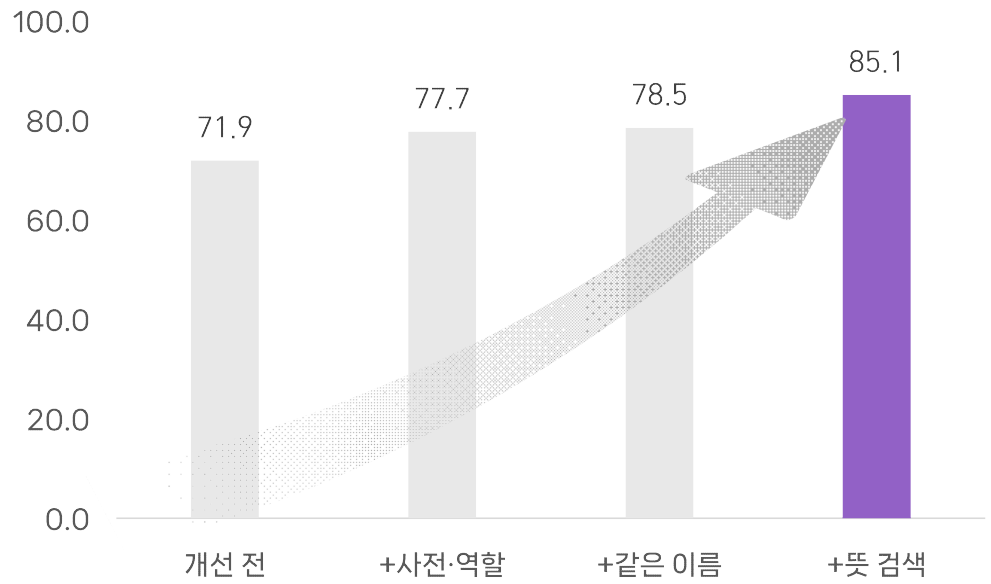
1위 정답률은 정답이 있는 121행 기준. 사전 항목 일부는 튜닝용 데이터를 본 뒤 만들어, 그 몫을 빼면 사전 효과는 +1.7%p로 추정한다.

개선 효과 : 1차 134행, 개선 전 → 후

같은 데이터·같은 지표로 비교했다. ①·③이 크게 올랐고, ④는 '높음'을 엄격하게 매긴 만큼 줄었지만 목표를 지켰다.

① 1위 정답률	② 상위 3개 안 정답률	③ '높음' 정답률	④ '높음' 비율
71.9% → **85.1%**	88.4% → **90.9%**	85.3% → **93.4%**	50.7% → **45.5%**
+13.2%p	+2.5%p	목표 90% 이상 · 달성	목표 45% 이상 · 달성

| 개선 단계별 ① 1위 정답률 (%) |



사전·역할 구분: +5.8%p

약어를 풀고 소재/공정이 반대인 후보를 뒤로 보냈다

뜻 검색 중심: +6.6%p

같은 이름 묶기(+0.8%p) 뒤 뜻 검색으로 바꿨다

'높음'을 더 엄격하게

뜻 점수 0.55 이상만 '높음' → ③ 오르고 ④ 내려감

사전 항목 일부는 튜닝용 데이터를 본 뒤 만들어 과적합 몫이 있다(빼면 사전 효과 +1.7%p 추정). 그래서 처음 보는 검증용 2000행으로 다시 쟀다(다음 장).

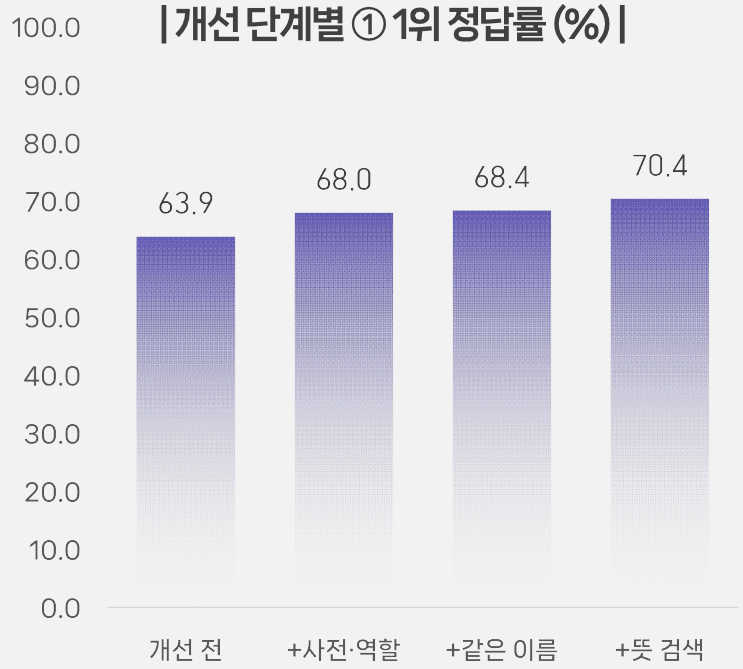
개선 효과 : 추가 실험 2000행, 개선 전 → 후

검증용 가상 BOM 2000행

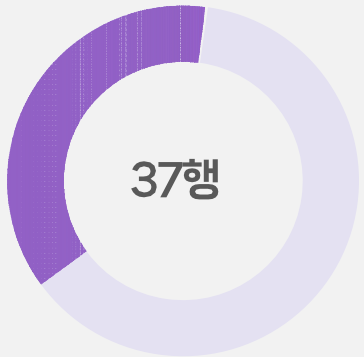
대용량 데이터 실험을 위해 임시 BOM 2000행 data로 늘려 실행결과 초기 설정과 개선 후 결과 비교

① 1위 정답률	② 상위 3개 안 정답률	③ '높음' 정답률	④ '높음' 비율
63.9% → 70.4%	78.6% → 81.8%	81.2% → 90.1%	45.7% → 37.0%
+6.5%p	+3.2%p	목표 90% 이상 · 달성	목표 45% 이상 · 미달

| 개선 단계별 ① 1위 정답률 (%) |



| 100행당 사람이 할 일 |



'높음' 37행: 훑어보고 승인

3개 중 고르기 45행 · 직접 찾기 18행



사이트 '성능 확인' 화면 (가상 BOM 2000)

- ③ 달성: '높음' 정답률 90.1% (목표 90%)
- ④ 미달: '높음' 비율 37.0% (목표 45%)
- 결론: 오답은 줄였지만 사람 몫이 아직 많다

①② 정답이 있는 1700행 기준, ③④ 전체 2000행 기준. 2000행은 약어 25%·정답 없음 15%로 더 어렵게 만든 데이터라 튜닝용 134행과 수치를 직접 비교하지 않는다.

프로젝트 한계 및 고도화 계획

AI를 활용한 가상 데이터 기준으로, 한계가 극명하여 향후 실무 BOM 기준으로 긍정적인 결과를 도출할 계획

데이터의 한계

- 정답·입력 모두 임시 data
- 같은 재료 반복, 실제 표본은 2000보다 작음
- 실제 회사 BOM은 아직 미측정

'높음' 비율 부족

- '높음' 37% (목표 45%)
- 100행 중 63행은 사람이 고르거나 찾음
- 없는 자재 300행 중 100행만 '없음' 판정

한계 4가지

약한 입력 유형 (1위 정답률)

- 약어·강종 코드 56.8%
- 넓은 이름(상위 개념) 44.0%
- 에너지·유틸리티 37.5%

아직 못 한 것

- LLM 판정은 불안정해 제외
- 소재+공정이 함께 필요한 부품 자동 분리
- 배출계수는 국가 DB의 32%만 제공



고도화 계획

- ① 규격·약어 사전 확장 (KS·JIS 강종, ISO 1043 수지)
- ② 회사별 사전 + 승인 이력 재사용
- ③ 모호한 입력은 올리기 전에 보완 안내
- ④ 실제 BOM(비식별)으로 다시 측정
- ⑤ openLCA로 결과 내보내기 연계