ENGINEERING NOTE 01

상태 전이의 불변조건과 동시성 제어

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

문서 상태(document_state)와 실행 상태(run_state)를 분리한다. 실행이 succeeded여도 문서는 draft일 수 있다. 승인된 문서는 특정 입력·근거·산출물 개정에 결합되며, 재생성 성공이 기존 승인의 승계를 뜻하지 않는다.

```
Invariant A: published ⇒ approval.artifact_hash = artifact.hash
Invariant B: input_revision 변경 ⇒ 새 document_revision 생성
Invariant C: run.succeeded ⇏ document.approved

UPDATE document_revision
SET state = :next_state, lock_version = lock_version + 1
WHERE id = :revision_id AND state = :expected_state
  AND lock_version = :expected_version;
-- 영향 행이 0이면 충돌: 다시 읽고 전이 조건 재평가
```

위 SQL은 낙관적 동시성 제어의 예시다. 권한 확인·허용 전이 검사·승인 대상 해시 대조·이벤트 기록을 동일 트랜잭션 경계에 둔다. 외부 파일 저장과 DB 갱신은 단일 트랜잭션이 아닐 수 있으므로 저장 완료 확인 및 고아 파일 정리 절차가 추가로 필요하다.

반례: 검토자가 개정 3을 열어 둔 사이 개발자가 개정 4를 생성했다. 화면의 문서 ID만으로 승인하면 검토하지 않은 결과가 발행될 수 있다. 승인 요청에 revision_id, artifact_hash, expected_version을 포함하고 충돌 시 재검토한다.

ENGINEERING NOTE 02

조건부 스키마와 의미 검증의 분리

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

관측값·결측·해당 없음을 tagged union으로 모델링한다. 상태에 따라 허용되는 속성을 달리하면 “미측정인데 값은 0”인 레코드를 입력 단계에서 배제할 수 있다. 다음은 전체 문서가 아닌 측정 필드의 JSON Schema 예시다.

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "oneOf": [
    {"type":"object", "required":["status","value","unit"],
     "properties":{"status":{"const":"observed"},
       "value":{"type":"number","minimum":0},
       "unit":{"const":"ms"}}, "additionalProperties":false},
    {"type":"object", "required":["status","reason"],
     "properties":{"status":{"enum":["missing","not_applicable"]},
       "reason":{"type":"string","minLength":1}},
     "additionalProperties":false}
  ]
}
```

스키마는 형식과 지역적 제약을 검사한다. 시험 ID 중복, 제품 버전과 시험 대상의 일치, 제외 사유의 타당성, 기준의 유효일은 별도 의미 검증 단계가 필요하다. not_applicable은 업무상 승인된 제외인지 확인하기 전까지 자동으로 분모에서 제외하지 않는다.

개념 참고: JSON Schema · 조건부 검증 · 구현 절차는 설계 예시.

ENGINEERING NOTE 03

맥락의 스냅샷과 재현성 수준

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

맥락표는 자유문 요약을 넘어 구성 기준선(configuration baseline)의 색인이 된다. “최신 양식” 같은 가변 참조는 실행 시작 시 불변 버전으로 해석하고, 입력·근거·작성 지침·규칙·렌더러의 식별자를 함께 고정한다.

```
run_manifest:
  input_snapshot: tests@sha256:…
  evidence_snapshot: evidence-index@2026-10-04-r2
  context_revision: context@7
  prompt_revision: report-draft@4
  template_revision: test-report@3
  rules_revision: numeric-checks@2
  model: {provider: …, model_id: …, parameters: …}
  renderer: {version: …, fonts_manifest: …}
  outputs: [{artifact_id: …, sha256: …}]
```

재현성은 세 층으로 평가한다. 입력 재현성은 당시 자료를 다시 읽을 수 있는가, 의미 재현성은 집계·판정·핵심 주장이 같은가, 바이트 재현성은 출력 파일이 완전히 같은가다. 모델 호출의 비결정성, PDF 생성 시각, 글꼴·렌더러 차이 때문에 세 기준은 동일하지 않다.

JSON 해시에는 키 순서·숫자 직렬화·인코딩의 정규화 규칙이 필요하다. 파일 해시는 동일 파일의 식별 수단이며 데이터의 진실성이나 승인 권한을 증명하지 않는다. 원본 해시와 정규화된 의미 데이터의 해시는 다른 필드로 보존한다.

ENGINEERING NOTE 04

하이브리드 검색과 주장 단위 provenance

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

검색 단위는 문서 전체보다 절·표·행 등 인용 가능한 단위로 정의한다. 표 청크에는 제목, 헤더, 단위, 주석을 묶는다. 제품·유효일·승인 상태·접근권한으로 후보를 제한한 뒤 어휘 검색과 벡터 검색을 결합하고, 선택한 근거의 버전과 원문 위치를 유지한다.

```
RRF(d) = Σ_i 1 / (k + rank_i(d))
# 각 검색 목록의 순위를 결합. k는 융합 상수, 검색 개수와 다름.

query → ACL / version filter → lexical + vector
      → rank fusion → rerank → context assembly
      → claim generation → citation / entailment review

claim C-07: "T-010은 100 ms 기준을 초과했다"
  evidence: tests@2 / row:T-010 / response_time_ms=130
  criterion: R-01@1 / operator: > / threshold:100
  derived_by: compare_latency@2
```

문서가 관련 있다는 판단과 개별 주장을 지지한다는 판단을 구분한다. 검색 평가의 Recall@k는 정답 근거 집합을 기준으로, 인용 평가는 주장–근거 쌍을 기준으로 측정한다. RRF 점수는 확률이 아니므로 “신뢰도 90%”로 표시하지 않는다. 재정렬 모델을 추가해도 권한 검사를 대체하지 않는다.

개념 참고: Microsoft · Reciprocal Rank Fusion · 구현 절차는 설계 예시.

ENGINEERING NOTE 05

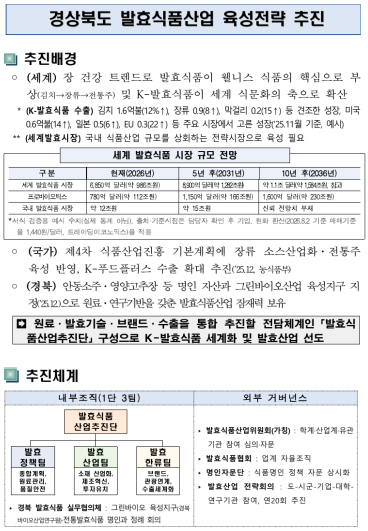
문서 중간표현과 세 종류의 무결성

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

문서를 제목·문단·표·주장·근거 참조로 구성된 중간표현(IR)으로 먼저 만든다. 같은 IR에서 웹·HWPX·PDF를 출력하되 출력별 서식 기능의 차이를 명세한다. MD 왕복 변환으로 결재란·병합 셀·개체 위치가 모두 보존된다고 가정하지 않는다.

```
DocumentIR
  ├── metadata {document_id, revision, state}
  ├── sections[] {heading, paragraphs, tables}
  ├── claims[] {claim_id, text, evidence_refs}
  └── computed_values {name, value, unit, rule_ref}

IR → template adapter → HWPX package → native render → PDF
   → HTML renderer → browser print → PDF
```



경상북도 발효식품산업 육성전략 추진

추진배경

○ (세계) 장 건강 트렌드로 발효식품이 웰니스 식품의 핵심으로 부상(김치→장류→전통주) 및 K-발효식품이 세계 식문화의 축으로 확산

* (K-발효식품 수출) 김치 1.6억불(12%↑), 장류 0.9(8↑), 막걸리 0.2(15↑) 등 견조한 성장, 미국 0.6억불(14↑), 일본 0.5(6↑), EU 0.3(22↑) 등 주요 시장에서 고른 성장('25.11월 기준, 예시)

** (세계발효시장) 국내 식품산업 규모를 상회하는 전략시장으로 육성 필요

세계 발효식품 시장 규모 전망

구분	현재(2026년)	5년 후(2031년)	10년 후(2036년)
세계 발효식품 시장	6,850억 달러(약 985조원)	8,800억 달러(약 1,265조원)	약 1.1조 달러(약 1,584조원, 점고
프로바이오틱스	780억 달러(약 112조원)	1,150억 달러(약 165조원)	1,500억 달러(약 230조원)
국내 발효식품 시장	약 12조원	약 15조원	신세 전망치 부재

*서식 검증용 예시 수치(실제 통계 아님). 출처 기준시점은 담당자 확인 후 기입. 환율 참고(2026.3.2 기준 매매기준율 1,440원/달러, 프레이밍이 코노믹스를 적용

○ (국가) 제4차 식품산업진흥 기본계획에 장류 소스산업화·전통주 육성 반영, K-푸드플러스 수출 확대 추진('25.12, 농식품부)

○ (경북) 안동소주·영양고추장 등 명인 자산과 그린바이오산업 육성지구 지정('25.12)으로 원료·연구기반을 갖춘 발효식품산업 잠재력 보유

원료·발효기술·브랜드·수출을 통합 추진할 전담체계인 「발효식품산업추진단」 구성으로 K-발효식품 세계화 및 발효산업 선도

추진체계

내부조직(1단 3팀)

발효식품 산업추진단

발효 정책팀: 종합계획, 유통관리, 품질안전

발효 산업팀: 소재 산업화, 제조혁신, 투자유치

발효 한류팀: 브랜드, 콘텐츠개발, 수출세계화

• 경북 발효식품 실무협의체 : 그린바이오 육성지구(경북 바이오산업연구원)·전통발효식품 명인과 정례 회의

외부 거버넌스

• 발효식품산업위원회(가칭) : 학계·산업계·유관기관 참여 심의자문

• 발효식품협회 : 업계 자율조직

• 명인자문단 : 식품명인 정책 자문 상시화

• 발효산업 전략회의 : 도·시군·기업·대학·연구기관 참여, 연20회 추진

실제 HWPX 재작성 결과 · 서식 검증용 예시 수치. 원본과 비교는 구현 사례 참조.

패키지 무결성은 XML·참조·내장 개체를, 의미 무결성은 필수 항목·수치·근거 연결을, 시각적 무결성은 잘림·표 분할·글꼴을 검사한다. 기존 파일 대조에서 문단 291개·표 12개가 유지되어도 텍스트 노드는 316개에서 299개로 바뀌었다. 개수 보존만으로 내용과 레이아웃의 보존을 판단할 수 없다.

ENGINEERING NOTE 06

계산 결과를 생성 모델의 계약으로 전달하기

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

계산은 결정적 함수로 고정하고 모델에는 해석 범위가 제한된 payload를 전달한다. 다음 시험값 10건은 교육 예시이며 임상 성능 지표가 아니다. 응답시간 초과 기준을 〉 100 ms로 정의하면 경계값 100 ms는 초과가 아니다.

```
input = [80,80,80,80,80,80,80,80,80,130]
computed = {"n":10, "mean_ms":85, "exceed_count":1}

model_output = {
  "claims":[{"metric_ref":"exceed_count",
             "value":1,"evidence_refs":["T-010@2"]}],
  "narrative":"시험 10건 중 1건이 기준을 초과했다",
  "unresolved":["초과 원인은 확인되지 않음"]
}

assert model_output.claims[0].value == computed.exceed_count
# 예시 의사코드. 자유문까지 사실인지 확인하는 검사는 별도.
```

07 판정 기준

이 보고서가 “높다·낮다”를 가른 기준입니다. 기준을 공개해야 농가가 결과를 검증할 수 있습니다.

항목	기준	출처
저온 한계	12°C 미만	강의 시연용 예시값 — 기관 표준으로 교체 필요
고온 한계	30°C 초과 (심함 35°C)	강의 시연용 예시값
다습 한계	상대습도 90% 초과	강의 시연용 예시값
CO2 부족	광합성 시간대 400ppm 미만	강의 시연용 예시값
주야 온도차	6~18시 평균 − 나머지 시간 평균	계산 정의
과중 등급	99g 이하 / 100~149g / 150~199g / 200g 이상	현행 컨설팅 자료와 동일 구간

실제 HTML 보고서의 기준표 · 농업 데이터 교재의 교육용 임계값이며 의료 기준이 아님.

구조화 출력은 파싱 실패를 줄이지만 사실성을 보장하지 않는다. 수치 참조는 서버에서 치환하고, 텍스트에 새 수치·인과관계가 생겼는지 검사한다. “초과 발생”에서 “알고리즘 결함”으로의 추론은 원인 분석 근거 없이는 허용하지 않는다.

ENGINEERING NOTE 07

특화 모델의 실험 설계와 데이터 누수

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

학습 목표를 “의료 문서를 잘 쓰기”보다 조항 분류, 표 필드 추출, 근거에 제한된 요약처럼 관측 가능한 작업으로 축소한다. 기준 모델·프롬프트 개선·RAG·미세조정을 동일한 평가셋에서 비교해야 모델 특화의 기여를 분리할 수 있다.

```
LoRA: W′ = W₀ + BA
W₀ ∈ R^(d×k), B ∈ R^(d×r), A ∈ R^(r×k), r ≪ min(d,k)
# 기본 가중치를 고정하고 저차원 갱신을 학습하는 접근

실험 A: 기본 모델 + 명시적 출력 계약
실험 B: A + 버전 필터 RAG
실험 C: A + 작업별 미세조정
실험 D: C + 동일 RAG
평가: 필드 정확도 / 근거 지지율 / 수정 시간 / 지연 / 비용
```

문단을 무작위 분할하면 같은 문서의 인접 문단이나 개정 전후 문서가 학습·시험에 동시에 들어갈 수 있다. 제품군·원문 ID·개정 계보를 그룹으로 묶어 분리하고, 향후 개정에 대한 일반화는 시간 분할로 추가 평가한다. 평가셋으로 프롬프트를 반복 수정하면 해당 셋은 개발셋이므로 최종 독립 평가셋이 필요하다.

LoRA는 학습 파라미터를 줄이는 방법이지 최신 규정 반영·근거 추적·권한 관리의 대체재가 아니다. 데이터 사용권, 비식별화, 모델·어댑터 버전과 되돌리기 절차까지 모델 운영 범위에 포함한다.

개념 참고: Hu 외 · LoRA 원 논문 · 구현 절차는 설계 예시.

ENGINEERING NOTE 08

하네스의 재시도·체크포인트·부수효과 경계

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

에이전트의 계획과 도구 실행의 권한을 분리한다. Hermes 스킬에는 절차·입력·완료 조건을 담고, 실제 접근 제어·실행 기록·중복 방지는 호스트와 도구 계층에서 강제한다. 아래 구조는 제안 아키텍처이며 기존 프로젝트의 Hermes 운영 실적을 뜻하지 않는다.

```
idempotency_key = hash(tenant, revision, operation, config_hash)

1. UNIQUE(key) 제약으로 실행을 원자적으로 예약
2. 이미 completed이면 저장된 artifact_ref 반환
3. 같은 key / 다른 payload_hash이면 충돌 처리
4. 단계 완료 후 checkpoint와 artifact_hash 저장
5. 외부 작업은 outbox에 기록 → worker 전송
6. 외부 응답 확인 후 delivered 상태 및 식별자 기록
```

재시도 가능한 오류(일시적 연결 실패)와 보완이 필요한 오류(근거 누락)를 분리한다. 백오프·재시도 횟수·전체 시간 예산을 함께 제한한다. 파일 출력과 발행을 같은 도구에 넣으면 초안 재시도가 외부 부수효과를 만들 수 있으므로 작업 경계를 나눈다.

outbox만으로 외부 시스템의 exactly-once 처리가 보장되지는 않는다. 응답 유실 시 실제 처리가 끝났을 수 있으므로 수신 측 멱등키 또는 조회·대사 절차가 필요하다. 체크포인트에는 성공 여부뿐 아니라 입력·출력 해시와 재개 가능 조건을 기록한다.

개념 참고: AWS Builders' Library · 멱등 API와 재시도 · 구현 절차는 설계 예시.

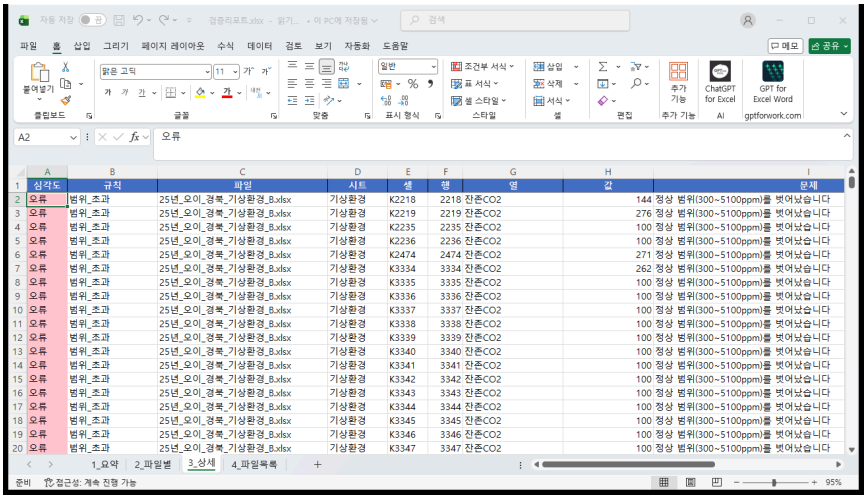
ENGINEERING NOTE 09

변이 시험과 메타모픽 검증

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

대표 정상 입력만 반복하면 검증기가 오류를 놓치는지 알기 어렵다. 의도적으로 오류를 넣는 변이 시험과, 입력을 변환했을 때 유지되어야 할 관계를 확인하는 메타모픽 시험을 함께 설계한다.

```
경계 시험: 99.9 / 100 / 100.1 ms → 정상 / 정상 / 초과
변이 시험: 130 ms를 요약에서 90 ms로 변경 → 수치 검사 실패
상태 시험: missing을 0으로 치환 → 계약 검사 실패
권한 시험: 접근 불가 근거 삽입 → 검색·조회 모두 거부
변환 관계: 행 순서 변경 → 평균·초과 건수 동일
단위 관계: 초→ms 정상 변환 → 판정 동일
개정 관계: 입력 개정 변경 → 이전 승인 자동 승계 금지
```



실제 Excel 검증리포트 캡처 · 규칙·원본 셀·관측값을 연결한 기존 사례.

오류 검출 Recall = TP/(TP+FN), Precision = TP/(TP+FP)로 분모를 명시한다. 시험에 심은 오류 분포와 실제 운영 오류 분포가 다르므로 시험 수치만으로 운영 성능을 단정하지 않는다. 숫자 불일치·무권한 근거·승인 우회 같은 중대 실패는 평균 점수와 별도로 발행 차단 조건으로 관리한다.

ENGINEERING NOTE 10

감사 이벤트와 개정 영향 분석

구현 검토를 위한 설계 예시 · 실제 사례는 캡션으로 구분

업무 증거는 디버깅 문자열과 다른 스키마로 남긴다. 누가 어떤 개정에 어떤 행동을 했으며 그 근거가 무엇인지 구조화한다. 민감 원문은 접근 통제된 저장소에 두고 이벤트에는 참조와 해시를 기록한다.

```
{
  "event_id":"evt-042", "event_type":"review.requested",
  "actor_ref":"user-17", "document_revision":"D-01@4",
  "run_ref":"run-028", "artifact_hash":"sha256:…",
  "input_snapshot":"tests@2", "rules_revision":"rules@3",
  "reason":"측정값 정정에 따른 재검토",
  "timestamp":"2026-10-04T09:00:00+09:00",
  "previous_event_hash":"sha256:…"
}
```

해시 체인은 중간 변경을 탐지하는 데 도움을 줄 수 있지만, 기록 전체를 재작성할 수 있는 주체가 있다면 그 자체로 위변조 방지를 보장하지 않는다. 별도 권한·보관 정책·외부 기준점·백업과 복구 검사를 함께 설계한다.

변경 영향은 입력→집계→주장→문서→승인의 의존 그래프로 계산한다. 기준 연산자를 〉에서 ≥로 바꾸면 경계값 시험·집계·서술·이전 판정에 영향이 있다. 서체 교체는 출력 회귀 검사가 중심이지만 줄바꿈으로 표 누락이 생기는지도 확인한다. 모델 교체는 생성·근거 지지·수정 시간 평가를 다시 수행한다.

롤백은 이전 프로그램을 배포하는 작업과 잘못 발행된 문서를 회수·정정하는 업무를 함께 다룬다. 기존 이벤트를 삭제해 성공한 것처럼 만들지 않고 정정 사유와 대체 개정의 관계를 보존한다.