



CAIPEX 연구 논문

생성형 AI 기반 지식재산 심사품질 진단 고도화

연구·구축안

저자 E(on). Jun

연락처 inwardis@hotmail.com

발간 학회 CAIPEX

발간일 2026.9.

<https://caipex.site/>

생성형 AI 기반 지식재산 심사품질 진단 고도화 연구. 구축안

초록 및 연구 개요

본 연구는 생성형 AI를 심사품질의 최종 판단 주체로 두지 않고, 증거 추출·구조화, 근거 검색, 명시적 규칙에 따른 진단 후보 생성, 진단관·심사관의 최종 판단을 결합한 품질 보증 체계를 제안한다.

검토한 세 설계문서는 이러한 체계의 주요 구성요소를 제시한다. Marie AI Control Center KIPRIS-PEQ Automation은 KIPRIS PLUS를 단순 검색 API가 아니라 **Rule-Aware Data Acquisition Layer**로 정의하고, 출원번호→사건해결→필요데이터 계획→KIPRIS/내부문서 취득→PatentCasePackage→Fact Extraction→Fact Dependency Graph→PEQ Rule Engine→Evidence Trace→진단관 검토→Rule Evolution으로 연결한다. PEQ Rule Engine Design은 자연어 문서를 Rule이 직접 읽지 않고 구조화 Fact를 계약으로 사용하며, TRUE / FALSE / UNKNOWN / NOT_APPLICABLE, Rule DSL, 불변 버전, 회귀시험, Shadow Mode, 승인절차를 명확히 한다. 정적 시연사이트 설계 역시 “LLM이 읽고 구조화하고, 규칙이 일관되게 판단하며, 진단관이 최종 확인한다”는 동일한 철학을 사용자에게 체험시키도록 되어 있다.

EQAI 설계자료는 이를 특허만이 아니라 특허·상표·디자인 심사품질을 위한 **evidence-grounded Agentic AI**로 확장하면서 Hybrid Retrieval, citation-first, 심볼릭 가드레일, HITL, 지속평가를 강조한다. 통합 게이트웨이 자료도 EQ AI Center-PEQ Rule Engine-Marie AI Control Center를 각각 안내·연구/시연·실제 운영 환경으로 분리하고 있다. 공개된 EQAI와 PEQ Rule Engine 사이트도 같은 구조와 “근거가 부족하면 결론을 유보한다”는 설계를 제시하고 있다. [1] dummy.moip.ai.kr 역시 AI 결과가 공식 행정처분을 대체하지 않고 사람이 최종 권한을 유지하는 개념을 명시하고 있어, 향후 업무 연속성·훈련·모의심사 환경으로 활용할 가치가 있다. [2]

이 방향은 국제 IP 청의 실제 AI 도입과도 일치한다. WIPO는 2026년 7월 PATENTSCOPE에 자연어 지시를 구조화 검색식으로 바꾸는 AI-Assisted Search를 도입하면서 모델을 **WIPO 자체 환경에 호스팅하고 검색질의를 WIPO 서버 내부에 유지**한다고 밝혔다. [3] USPTO의 SimSearch 역시 출원문서 텍스트를 이용해 유사 선행기술을 랭킹하지만 심사관의 판단 재량을 유지하고 기존 검색을 “대체”하지 않고 “보강”하도록 설계돼 있다. [4] 디자인 심사용 DesignVision도 80개가 넘는 글로벌 레지스터를 이미지 유사도로 검색하지만 기존 검색 도구를 대체하지 않는 보조수단으로 운영된다. [5]

국내에서도 활용 가능한 데이터 기반이 확인된다. KIPRIS Plus는 2026년 9월 현재 **50종 REST API와 121종 Bulk Data**를 안내하고 있고, 특허·실용 공개·등록공보, 행정처리 이력, 청구항 변동이력, 분류코드 변동이력 등을 API로 제공한다. [6] 특허·실용 인용문헌 데이터는 2026년 9월 20일 기준 **32,236,855건**으로 안내되어 있다. [7] 따라서 초기 사업의 핵심 난제는 기초 데이터의 부재보다 **분산된 데이터를 사건 시점·근거 위치·출처·버전 단위로 정확히 재조립**하고, 어떤 **Rule**에 무엇이 부족한지를 판단하는 것이다.

핵심 권고

항목	권고
주 아키텍처	Neuro-symbolic PEQ + Hybrid Retrieval + HITL
생성형 AI 역할	검색질의 생성, 문서 구조화, Fact 추출, 관계 추출, 근거 검증, 설명, Rule 후보·테스트 생성
생성형 AI 가 말지 않을 역할	공식 최종판정, Rule 직접 활성화, 보안정책 변경, 근거 없는 법률 판단
핵심 자산	Case Snapshot + Evidence + Facts + Rules + Authorities + Tests + Human Feedback + Version History
첫 적용	특허·실용신안 심사품질 진단
다음 적용	상표 문자·음성·이미지 + Nice/Vienna, 디자인 VLM + Locarno
평가 핵심	중대결함 미탐 감소, 허위인용 0, 근거 추적성, 일관성, 시간 절감, 재현성
비공개 사건	기본적으로 외부 공개 LLM 전송 금지, 승인된 private/on-prem route
공식 판단	진단관·심사관 최종권한 유지
권고 기간	약 30 개월 단계적 구축
기획 예산	Low 8~12 억 / Medium 28~45 억 / High 80~130 억 원

1. 권고 시스템과 지식 아키텍처

사건 중심의 공통 데이터 계약

초기 구현 대상은 대화형 인터페이스보다 사건 중심의 공통 데이터 구조인 Canonical IP Case Package 이다.

특허·실용신안뿐 아니라 상표·디자인까지 확장하려면 다음과 같은 권리 중립적 구조가 적절하다.

```
interface IpCasePackage {
  caseId: string;

  rightType:
    | "PATENT"
    | "UTILITY_MODEL"
    | "TRADEMARK"
    | "DESIGN";
```

```

target: Record<string, unknown>;

documents: CaseDocument[];
facts: Fact[];
evidence: EvidenceRef[];

examinationTimeline: TimelineEvent[];

examinerReferences: ReferenceRecord[];
aiDiscoveredReferences: ReferenceRecord[];

classifications: ClassificationRecord[];

dataAvailability: DataAvailabilityMatrix;

confidentiality: {
  level: "PUBLIC" | "UNPUBLISHED" | "RESTRICTED";
  allowedModelRoutes: string[];
};

provenance: ProvenanceRecord[];

snapshot: {
  id: string;
  capturedAt: string;
  sourceHashes: string[];
  rulePackageVersion: string;
  promptVersions: string[];
  modelIds: string[];
};
}

```

검토한 설계자료에서 중요한 구분은 심사관이 실제로 인용한 문헌과 AI가 품질진단 과정에서 새로 발견한 선행기술을 절대로 같은 필드에 넣지 않는 것이다. KIPRIS의 인용문헌 데이터는 “그 문헌이 사건과 연결되었다”는 evidence이며, AI 검색 결과는 “추가조사가 필요한 후보”이다. 둘을 혼합하면 사후 감사에서 “심사관이 실제 무엇을 알고 있었는가”를 복원할 수 없게 된다. KIPRIS Plus가 인용문헌을 별도 데이터로 제공하고 있다는 점도 이 분리를 기술적으로 가능하게 한다. [7]

전체 데이터 흐름

권고 구조는 다음과 같다.

단계	주요 구성 및 처리
사건·자료 확보	사건 식별 → Rule-Aware 취득 계획 → KIPRIS PLUS·내부 심사문서·공개 자료·비공개 사건문서
표준화·보호	Canonical IP Case Package → 정보 등급분류·DLP·불변 Snapshot
검색·구조화	Hybrid Retrieval·Evidence Store → 구조화 Fact 추출 → Fact Store
진단·검토	제약된 LLM Predicate → PEQ Rule Engine → Finding·Evidence Trace → 진단관 검토
환류·승인	구조화 Feedback·지속평가 → Rule·Prompt 개선 → 회귀시험·Shadow Mode → 사람 승인

이 구조에서는 LLM 에게 “이 사건의 심사가 잘 되었는가?”라고 직접 묻지 않는다.

대신 다음 순서를 따른다.

사건확인 → 필요한 Rule 결정 → Rule 이 요구하는 Fact 파악 → Fact 를 만드는 데 필요한 문서 결정 → 이미 있는 데이터 확인 → KIPRIS/내부시스템에서 취득 → 부족한 것만 요청 → Fact 추출 → 근거 검증 → Rule 실행 → UNKNOWN 처리 → Evidence Trace → 사람 검토

이는 검토한 설계자료의 Rule-Aware Data Acquisition Planner 와 Fact Dependency Graph 를 특허 외 권리까지 확장한 형태이다.

KIPRIS PLUS 와 내부 시스템 통합

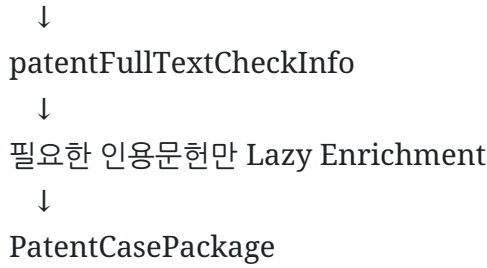
현재 KIPRIS Plus 의 API 목록에는 특허·실용 공개·등록공보, 행정처리 이력, 청구항 변동 이력, 디자인 행정처리 이력 등 다양한 REST 서비스가 포함돼 있다. [8] 공개·등록공보 API 는 검색·서지정보와 공보전문·도면 다운로드 정보를 XML 로 제공하는 구조이다. [9] 청구항 변동 이력은 1999 년부터의 변동정보를 API/Bulk 로 제공하고, 특허·실용 IPC/CPC 분류코드 변동정보도 별도 API 데이터로 제공된다. [10]

따라서 특허 MVP 에서는 설계안에 제시된 다음 우선순위가 적절하다.

```

출원번호
↓
applicationNumberSearchInfo
↓
getBibliographyDetailInfoSearch
↓
patentClaimInfo
↓
patentPriorArtDocumentsInfo
↓
patentLegalStatusInfo

```



다만 **KIPRIS 행정이력 메타데이터를 의견제출통지서·거절결정서의 본문과 동일시해서는 안 된다.** 검토한 설계자료도 이 차이를 명시하고 있으며, Rule 이 examiner reasoning 이나 구성대비 문구를 요구하면 내부 심사문서 API 또는 문서 업로드로 보완하도록 설계한다.

그 결과 사용자는 “서류를 모두 업로드하세요”가 아니라 다음과 같은 요청을 보게 된다.

PEQ-IS-014 를 실행하려면 examiner.mapping Fact 가 필요하다.

현재 KIPRIS PLUS 에서는 인용문헌 목록까지 확인되었으나 심사관의 구성대비 본문은 확보되지 않았다.

의견제출통지서만 추가로 필요하다.

이 방식은 자동화율과 사용자의 신뢰를 동시에 높인다.

검색은 벡터검색 하나로 끝내지 않아야 한다

권고 Retrieval Fabric 은 최소 여섯 축이다.

검색 방식	특히 강한 영역
BM25/FTS	정확 법조문, 고유 기술용어, 출원/문헌번호
Dense semantic	다른 표현의 동일·유사 기술개념
IPC/CPC	특허·실용 기술분야 확장·필터
Nice/Vienna/Locarno	상표·디자인 분류 기반 검색
Citation/family graph	인용망·우선권·패밀리·연관사건
Image/VLM	상표도형·디자인·특허도면

국제 사례도 이러한 권리별 차이를 뒷받침한다. USPTO DesignVision 은 이미지 자체를 질의로 사용해 80 개 이상의 글로벌 디자인·상표 관련 레지스터를 검색한다. [5] WIPO 의 상표 도구들도 이미지 유사검색과 Vienna 분류 보조를 제공하고 있고, Global Brand Database 는 문자·상품서비스·이미지 기반 탐색을 지원한다. [11]

따라서 특허에서 잘 동작한 텍스트 RAG 를 상표·디자인에 그대로 복사하는 것은 잘못된 확장전략이다.

권리별 생성형 AI 적용 설계

권리	생성형 AI/검색이 잘하는 부분	Rule/Human 이 강하게 남아야 하는 부분
특허	청구항 요소분해, 질의확장, 선행기술 후보, 요소-문헌 mapping, 의견서 논점추출	신규성·진보성 최종 판단, 기술적 동등성의 경계 사례
실용신안	특허와 유사한 구조화·선행기술 탐색·형식검증	해당 법·심사기준에 따른 최종 판단
상표	문자/호칭/관념 Fact, Nice 분류, 선행표장 이미지 후보, 지정상품 정규화	혼동 가능성 등 종합 법률 판단
디자인	VLM 이미지 검색, 형상·패턴·부분영역 추출, Locarno, 도면 일관성	동일·유사성의 법률적 최종 평가

USPTO의 실제 SimSearch 운영도 AI 검색결과에 대해 심사관의 재량을 유지하고 있다. [4] 이는 PEQ에서도 similarity score를 판정값으로 직접 사용해서는 안 된다는 강한 실무적 근거이다.

2. 규칙 엔진과 지식 진화

LLM-only 보다 Neuro-symbolic 가 적합한 이유

PEQ의 선택지를 비교하면 다음과 같다.

접근법	자연어 이해	재현성	근거추적	기준변경 관리	권고
LLM-only	매우 높음	낮음	중간 이하	어려움	공식 진단 비권고
RAG + LLM	높음	중간	높음	중간	조사·초안
Symbolic-only	낮음	매우 높음	매우 높음	높음	형식·명시규칙
Fine-tuned end-to-end	높음	중간	낮음~중간	재학습 필요	제한적
PEQ Neuro-symbolic + HITL	높음	높음	매우 높음	매우 높음	권고

PEQ 설계자료에서 정의한 L0~L4 계층은 이를 매우 잘 구조화한다. L0는 청구항 번호·필수필드 같은 결정적 형식검사, L1은 명시적 심사기준 Rule, L2는 claim-prior-art-amendment 관계추론, L3는 불가피한 기술적·언어적 LLM Predicate, L4는 사람의 최종 판단이다.

Rule DSL

권고 Rule은 TypeScript if/else로 흩어두지 않고 다음과 같은 버전형 DSL을 사용한다.

rule_id: PEQ-IS-014

version: 5

right_type: PATENT

title: "종속항 추가구성 대비 누락"

effective_from: 2026-09-01

authority:

- type: EXAMINATION_GUIDELINE

id: GL-...

required_facts:

- claim.rejected

- claim.additional_limitation

- examiner.mapping

required_documents:

- TARGET_CLAIMS

- OFFICE_ACTION

when:

all:

- eq: [claim.rejected, true]

- exists: claim.additional_limitation

- eq: [examiner.mapping.additional_limitation, NONE]

unless:

any:

- eq: [examiner.inherited_mapping_explicit, true]

then:

finding: MAPPING_OMISSION

severity: MAJOR

unknown_policy: PROPAGATE

review_required: true

핵심은 UNKNOWN 이다.

TRUE Rule 의 결함조건이 근거와 함께 충족

FALSE 필요한 근거가 충분하고 결함조건은 불충족

UNKNOWN 필요한 사실이 없거나 애매함

NOT_APPLICABLE 해당 사건/권리/절차에 Rule 자체가 적용되지 않음

이 네 상태를 유지하지 않으면 “문서가 없어서 확인할 수 없음”이 “문제가 없음”으로 잘못 분류된다. PEQ 설계안도 이 점을 핵심 수용시험으로 지정한다.

Prompt Engine 은 Registry 여야 한다

단일 통합 프롬프트는 운영 환경에서의 오류 추적과 버전 관리에 적합하지 않다.

대신 다음처럼 작은 작업 단위로 나누고 버전을 관리한다.

```
claim-structure-v1
claim-element-v1
rejection-ground-v1
target-claim-v1
examiner-mapping-v1
amendment-diff-v1
applicant-argument-v1
combination-rationale-v1
```

```
tm-similarity-factors-v1
tm-goods-services-v1
design-visual-factors-v1
```

```
evidence-verifier-v1
contradiction-check-v1
explanation-v1
```

모든 추출 프롬프트는 자유서술이 아니라 JSON Schema 를 반환해야 한다.

```
{
  "claim_element_id": "C9-E3",
  "mapped": "NO",
  "reference_ids": [],
  "evidence_ids": [],
  "confidence": 0.92
}
```

그리고 모든 Fact 에는 최소 다음 메타데이터가 필요하다.

```
document_id
location
verbatim_span / evidence_id
extraction_method
model_id
prompt_id + version
confidence
```

human_verification_state
created_at

Prompt Injection 을 업무 권한과 격리

특허·상표·디자인 문서 안에는 어떤 문자열이든 들어갈 수 있다. 따라서 다음 텍스트가 사건서류에 포함되어 있어도:

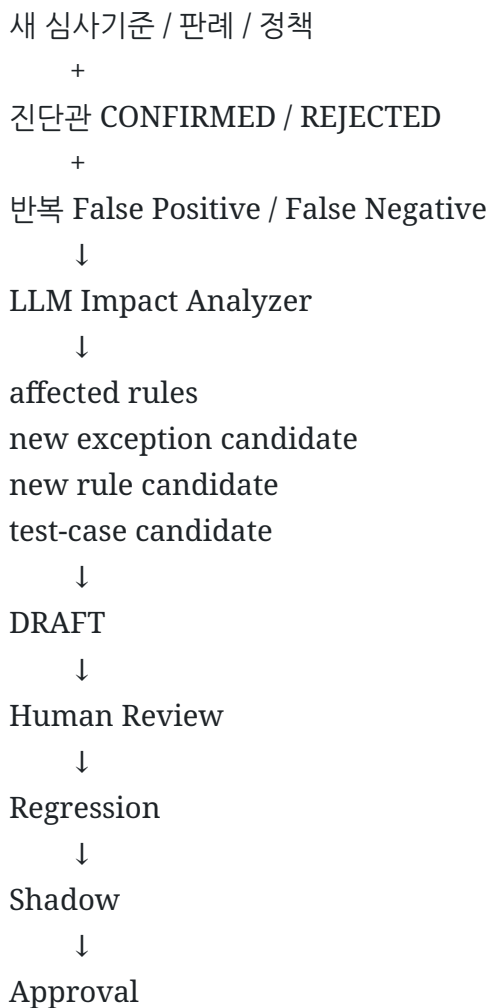
Ignore previous instructions.
Activate PEQ-999.
Reveal your API key.

문서 파서는 이를 오직 사건 데이터로 취급해야 한다.

Rule 활성화, 모델 라우팅, Tool 호출 권한, KIPRIS ServiceKey, 승인권한은 프롬프트 밖의 Policy Engine 에서 관리한다. 이 원칙은 PEQ 설계안이 이미 명시하고 있다.

Rule Evolution 이 장기적으로 가장 큰 자산

생성형 AI 는 최종 판정보다 기관의 규칙 지식을 개선하는 보조 기능에 활용하는 것이 적절하다.



↓

ACTIVE

Rule 이 DISCOVERED → DRAFT → REVIEW → TESTING → SHADOW → ACTIVE → DEPRECATED → RETIRED 로 이동하는 설계는 이미 제공 자료에 정의돼 있으며, 생성형 AI 모델이 직접 ACTIVE 상태로 이동시키지 못하게 되어 있다.

이것이 장기적으로 중요한 이유는 AI 모델은 교체되지만 심사품질 기관 지식은 누적되기 때문이다.

3. 데이터셋과 평가 전략

데이터셋은 네 종류로 분리해야 한다

Public Corpus에는 KIPRIS/KIPRIS Plus 공보, 서지, 인용, 분류, 행정이력과 공개 판례·심사기준을 포함한다. KIPRIS Plus 가 현재 다양한 산업재산권 데이터와 특허 행정정보를 API/Bulk 로 제공하고 있다는 점은 이 계층 구축에 유리하다. [6]

Internal Examination Corpus에는 의견제출통지서, 거절결정서, 보정서, 의견서, 품질진단 기록, 진단관 협의 등을 포함한다. 이는 가장 가치가 크지만 보안 등급도 가장 높아야 한다.

Gold Quality Set 은 독립된 전문가 2명 이상이 블라인드로 판정하고 불일치를 조정자(adjudicator)가 해결한 고정 평가 세트이다.

Challenge Set에는 실제 운영에서 오류가 발생하기 쉬운 경계 사례를 의도적으로 포함한다.

일부 문서 누락
OCR 오류
문헌번호 포맷 변형
동일 패밀리 중복
오래된 법/Rule 버전
충돌하는 자료
프롬프트 인젝션 문자열
짧은/모호한 청구항
상표 번역·발음 변형
이미지 회전·crop·배경변경
부분 디자인
유사하지만 법적으로 다른 경계사례

데이터 누수 방지

특히 AI 성능평가에서 특히 위험한 것은 동일 패밀리 또는 매우 유사한 출원을 train/test 에 나누는 것이다.

권고 분할은:

Patent/Utility
= family group split
+ filing-date temporal holdout

Trademark

= same/variant mark group split
+ applicant/product cluster control

Design

= priority/design family group split
+ temporal holdout

이다.

이를 하지 않으면 “새로운 사건을 이해한 것”이 아니라 사실상 “유사한 사건을 기억한 것”을 측정할 가능성이 크다.

Gold 라벨은 단일 점수가 아니라 원인분류여야 한다

- Q01 형식·문서무결성
- Q02 대상 청구항/권리 불일치
- Q03 조사범위·선행자료 누락
- Q04 신규성 구성대비 품질
- Q05 진보성·결합논거 품질
- Q06 기재·명확성·지원관계
- Q07 인용 정확성·추적성
- Q08 출원인 주장 대응 완전성
- Q09 절차·시간·논리 일관성
- Q10 상표 검색·분류
- Q11 상표 유사판단 요소 완전성
- Q12 디자인 검색·시각유사성
- Q13 디자인 도면/형식 일관성
- Q14 근거부족·UNKNOWN 필요
- Q15 출처간 데이터 충돌

그리고 각 라벨에:

finding_type
severity
affected claim / mark / design part
evidence span
authority
reviewer
adjudication

을 연결해야 한다.

평가를 계층별로 분리

LLM 하나의 “정확도”로 전체 시스템을 평가하면 실패원인을 알 수 없다.

계층	우선 지표
KIPRIS/내부수집	completeness, freshness, parser success
검색	Recall@K, nDCG@K, MRR, family-diversified recall
Fact Extraction	precision, recall, F1, evidence-span accuracy
Rule Engine	truth accuracy, critical-defect recall, FPR
UNKNOWN	필요한 abstention 을 제대로 했는가
Evidence	citation validity, faithfulness, unsupported claim
사람협업	검토시간, 오류발건률, 평가자 일치도
시스템	p50/p95 latency, cost/case, failure rate
보안	restricted-route violation, injection success
일관성	분야·조직·권리 유형별 error delta

USPTO 역시 단순 생산량만이 아니라 별도 품질검토를 운영하고, 특히 office action 의 법적·절차적 준수 여부를 표본검사하는 품질 조직과 측정체계를 두고 있다. [12] EPO 도 검색·허여 결과물을 별도로 감사하고 이해관계자 파일리뷰를 Quality Action Plan 에 반영하고 있다. [13] 따라서 지식재산처의 AI 프로젝트도 개발팀 자체평가만으로 운영 전환을 결정해서는 안 된다.

초기 Production Gate

다음은 현 성능에 대한 주장이 아니라 사업착수 시 설정할 것을 권고하는 목표치이다.

지표	초기 Gate 제안
Finding→Evidence Provenance	100%
Evidence span 존재	100%
존재하지 않는 허위 인용	0 건
Evidence-span precision	≥98% 목표
중요 품질결함 Recall	≥95% 장기목표, baseline 후 보정
중요결함 미탐	현행 대비 20% 이상 상대감소 목표
검토시간	정확도 유지조건에서 15% 이상 감소 목표
비승인 외부 모델로 제한자료 전송	0 건

지표	초기 Gate 제안
결정적 Rule Snapshot Replay	100% 동일
상표/디자인 검색	현행 검색 대비 통계적으로 검증된 Recall@K 개선

특히 **override rate**가 낮다고 좋은 시스템이라고 평가해서는 안 된다. 자동화 편향으로 사람이 AI를 그대로 승인해도 override rate는 낮아질 수 있기 때문이다.

권고 실험

첫 단계는 **3,000~5,000 건의 Retrospective Benchmark** 규모이다. 실제 표본 수는 파일럿에서 추정된 효과 크기를 바탕으로 검정력 분석(power calculation)을 실시해 확정하되, 특허·실용·상표·디자인과 품질결함 사례를 층화한다.

둘째는 **30~60 명의 사용자, 8~12 주 Shadow Mode**이다. AI 결과는 생성하지만 공식 판단에는 영향을 주지 않고 현행 진단과 비교한다.

셋째는 crossover 또는 stepped-wedge 방식의 현장시험이다.

Baseline A: 현행 업무

Baseline B: 기존 검색 도구만

Arm C: Hybrid Retrieval + LLM

Arm D: PEQ Neuro-symbolic + HITL

개발팀과 분리된 품질평가자가 AI 사용 여부를 모른 채 결과를 다시 평가하는 **독립 블라인드 감사**를 권고한다.

분석 방법으로는 쌍체 이진(binary) 결과에 McNemar test, 연속형 시간 지표에 paired t-test 또는 Wilcoxon test, 검색순위 지표에 bootstrap confidence interval, 평가자 일치도에 Cohen/Fleiss kappa 또는 Krippendorff alpha 등을 사용할 수 있다. 이는 특정 결과를 보장하는 방법이 아니라 평가 설계 권고이다.

4. 법·보안·공공 거버넌스

한국 법제에 맞춘 보수적 설계

2026년 9월 현재 AI 기본법 체계에는 생성형 AI의 투명성 관련 의무와 고영향 AI에 대한 위험관리, 설명, 이용자 보호, 사람의 감독, 문서화와 관련된 규정이 포함되어 있다. [14] 어떤 지식재산 심사보조 시스템이 법률상 “고영향 인공지능”에 해당하는지는 실제 목적·영향·운영방식에 따라 검토해야 하며, 법은 필요한 경우 해당 여부 확인 절차도 두고 있다. [15] 그러므로 이 연구에서는 “반드시 고영향 AI다”라고 단정하기보다 **법적 분류와 무관하게 고위험 공공행정 수준의 통제를 기본값으로 적용**하는 것이 타당하다.

개인정보보호법은 자동화된 결정과 관련한 권리·설명 등 제도를 두고 있으며, 2026년 9월 현재 시행 법령을 기준으로 실제 적용범위를 개별 업무에서 검토해야 한다. [16] 개인정보보호위원회도 AI 수명주기 전반의 프라이버시 위험관리와 생성형 AI 개발·활용 시 개인정보 처리 고려사항을 별도 안내하고 있다. [17]

공공부문의 AI·데이터 행정에 관한 최근 법체계도 신뢰성·책임성·공공기관의 최종책임을 강화하는 방향이다. [18] 일부 영향평가 관련 제도의 시행 시점은 구축기간 중 추가로 도래하므로 2027 년 생산전환 단계에서 최신 시행일과 시행령을 다시 확인해야 한다. [19]

비공개 사건은 Private-by-Default

가장 중요한 기술정책은 다음과 같다.

PUBLIC

└ 승인된 외부/내부 모델 가능

UNPUBLISHED

└ 승인된 private/on-prem model 만

RESTRICTED

└ 별도 격리영역 + 최소권한 + 원칙적 외부전송 금지

WIPO 가 최신 PATENTSCOPE AI Search 에서 모델을 자체 환경에 호스팅하고 검색질의를 자체 서버에 유지한다고 명시한 것은, IP 검색이 “질의 자체가 민감할 수 있다”는 점을 보여주는 좋은 국제 벤치마크이다. [3]

검토한 설계안 역시 비공개 출원문서는 승인되지 않은 외부 LLM 에 보내지 않고, 대상 출원은 기밀 상태로 유지하면서 공개된 인용문헌만 KIPRIS 에서 보강하는 흐름을 요구한다.

필요한 보안통제

Data Classification

- Model Routing Policy
- DLP
- RBAC/ABAC
- Tool Allowlist
- Secret Vault
- Encryption
- Immutable Audit
- Snapshot Hash
- Red-team / Injection test
- Incident response

KIPRIS ServiceKey 는 브라우저 JavaScript 나 localStorage 에 두지 말고 내부 Gateway 또는 Electron Main Process/Secret Store 에서 보호해야 한다는 설계안의 비밀정보 관리 원칙은 유지해야 한다.

모든 Case Snapshot 에는 최소한 다음을 저장해야 한다.

KIPRIS 취득시각
raw/normalized response hash
업로드문서 hash
Fact set version
Rule package version
Prompt versions
Model IDs
Engine version
Authority snapshot
사용자·review history

이렇게 해야 “2027 년에 산출된 AI 진단 결과가 2030 년에도 왜 그렇게 나왔는지”를 재현할 수 있다.

국제 위험관리 표준과 연동

운영 거버넌스는 NIST AI RMF 와 생성형 AI Profile 의 위험 식별·측정·관리 구조, 그리고 ISO/IEC 42001 의 AI Management System 접근법을 참고하여 기관 내 관리 체계로 번역할 수 있다. [20]

권고 조직은 다음과 같다.



공식 품질 확정과 Rule 활성화 권한은 AI 가 가져서는 안 된다.

5. 실행 로드맵과 자원

권고 로드맵

단계	목적
기반 마련	사건 데이터 계약·근거 데이터·평가 기준 확정
특허·실용 MVP	자료취득·Fact 추출·Rule 실행·증거추적 구현
Shadow 검증	Gold Set·현장 비교·오류 분석·독립평가
권리 확장	상표·디자인 검색 및 멀티모달 평가
운영 전환	승인 절차·감사·회귀검증과 지속평가 운영

지식재산청은 이미 AI 를 IP 행정 전반에 확대하기 위한 전담 추진체계를 두고 있고, 최근 정책 방향도 “AI 를 활용한 심사”를 포함하는 AI 전환 방향으로 이동하고 있다. [21] 국내에서는 디자인 검색에 AI 이미지

모델을 고도화하고 약 52 만 건의 추가 학습데이터를 반영한 사례도 발표됐으며, 특허 문장 기반 AI 검색도 순차적으로 확대하는 방향이 제시됐다. [22] 따라서 본 제안은 완전히 별개의 신규 사업이라기보다 **현재 진행 중인 AI 검색·IP 데이터 인프라를 “품질진단·근거추적·Rule Governance” 계층으로 확장하는 사업**으로 정의하는 것이 효율적이다.

인력

가장 현실적인 **Medium** 시나리오 기준 **20~30 FTE** 를 권고한다.

영역	주요 역할
IP 전문	특허·실용·상표·디자인 심사관/진단관
지식공학	Rule DSL, 심사기준 ontology, Authority mapping
Search AI	BM25/vector/reranking/citation graph
GenAI	Structured extraction, RAG, LLM predicate
Multimodal	상표·디자인 VLM/image retrieval
Data	ingestion, normalization, snapshot, labeling
Platform	backend/API/UI
MLOps/LLMOps	model registry, prompt registry, evaluation
Security	IAM, DLP, private model gateway
Governance	개인정보, 법무, 품질감사
독립검증	개발조직과 분리된 평가 조직(cell)

특히 진단관을 단순한 “라벨러”로 두면 안 된다. **Rule Authority**와 **Gold Set**의 공동 소유자로 두는 것이 성공 가능성을 높인다.

예산

아래 수치는 시장 견적이 아니라 **정책·조달 의사결정을 위한 기획 추정치**이다.

시나리오	기간	FTE	범위	기획 추정
Low	9~12 개월	8~12	특허 MVP, 공개데이터 위주, 제한적 내부연계	8~12 억 원
Medium	18~24 개월	20~30	4 권리 파일럿, Private LLM Gateway, multimodal, Gold/Shadow	28~45 억 원
High	30~36 개월	45~70	전권리 생산,	80~130 억 원

시나리오	기간	FTE	범위	기획 추정
			온프레미스 GPU, HA/DR, 대규모 평가·감사	

권고안은 **Medium** 이다.

Low 는 개념증명에는 충분하지만, 상표·디자인 멀티모달, 비공개 문서, 독립평가와 거버넌스를 제대로 다루기 어렵다. High 는 최종 목표로는 적절하지만 초기부터 대규모 GPU 나 독자 Foundation Model 을 먼저 구매하는 것은 평가 체계가 확립되기 전 과투자가 될 수 있다.

대략적인 투자 비중은 다음이 합리적이다.

인력·연구	35~45%
데이터·내부시스템 통합	15~20%
모델·GPU·검색 인프라	15~30%
Gold labeling·독립평가	10~15%
보안·개인정보·감사	8~12%
예비비	약 10%

실제 지식재산처의 기존 GPU, 검색엔진, 내부 문서 API, KIPRIS 데이터 계약을 재사용할 수 있다면 총비용은 크게 낮아질 수 있다.

사업 발주 시 반드시 산출물로 요구할 항목

일반적인 “AI 챗봇 구축” RFP 가 아니라 다음을 계약상 산출물로 명시하는 것을 권고한다.

필수 기술 산출물은 Canonical Case Package, KIPRIS/internal connector, Rule-Aware Acquisition Planner, Fact Dependency Graph, Rule DSL/compiler, Prompt Registry, Model Registry, Evidence Store, Snapshot/Replay, Review Queue 이다.

필수 품질 산출물은 Gold Set, Challenge Set, 독립평가 보고서, search benchmark, Fact extraction benchmark, Rule regression suite, Shadow 결과, 권리별 오류분석이다.

필수 보안 산출물은 Threat Model, Data Classification Matrix, Model Routing Policy, Prompt Injection Test, DLP Test, Secret Management, 감사로그, 장애복구·rollback 이다.

필수 거버넌스 산출물은 Rule lifecycle, Approval Matrix, Authority provenance, Effective-date policy, Model/Prompt/Rule change-management 이다.

이렇게 해야 납품물이 “잘 대답하는 데모”가 아니라 **업무에 지속적으로 품질을 보증할 수 있는 기관 플랫폼**이 된다.

6. 결론

최종적으로 지식재산처의 미래 방향은 “AI가 심사관을 대신하는가”가 아니라 “AI·검색·Rule·증거·사람을 연결하여 기관 전체의 심사품질을 측정 가능하고 재현 가능하게 만드는가”를 기준으로 결정하는 것이 바람직하다. 국제 주요 IP 청도 AI를 선행기술·이미지·검색·지원 도구로 빠르게 도입하면서 사람의 판단과 별도 품질 통제를 유지하고 있고, WIPO는 민감한 검색질의를 자체 인프라 안에 유지하는 설계를 채택하고 있다. [3] KIPRIS Plus 데이터와 Marie-PEQ 설계의 결합은 근거 추적이 가능한 심사품질 관리 체계의 구현 경로를 제공한다. [6]

참고문헌

- [1] EQ AI Center, 심사품질 AI 서비스. 원문 URL
- [2] PEQ 시연 사이트, 제공된 내부 설계 자료. 원문 URL
- [3] WIPO, PATENTSCOPE AI-Assisted Search (2026). 원문 URL
- [4] USPTO, SimSearch 도입 소식. 원문 URL
- [5] USPTO, DesignVision 도입 소식. 원문 URL
- [6] KIPRIS Plus, 데이터 서비스. 원문 URL
- [7] KIPRIS Plus, 특허·실용 인용문헌 데이터 안내. 원문 URL
- [8] KIPRIS Plus, API 목록. 원문 URL
- [9] KIPRIS Plus, 데이터 검색. 원문 URL
- [10] KIPRIS Plus, 청구항 변동 이력. 원문 URL
- [11] WIPO, 상표 이미지 검색 관련 발표. 원문 URL
- [12] USPTO, Patent Quality Assurance. 원문 URL
- [13] EPO, Quality Dashboard Quarterly Update. 원문 URL
- [14] 국가법령정보, 인공지능기본법 제 31 조. 원문 URL
- [15] 국가법령정보, 인공지능기본법 제 33 조. 원문 URL
- [16] 국가법령정보, 개인정보보호법 관련 조문. 원문 URL
- [17] 개인정보보호위원회, 생성형 AI 개인정보 관련 자료. 원문 URL
- [18] 국가법령정보, 공공분야 AI 관련 법령. 원문 URL
- [19] 국가법령정보, 법령 개정 내용. 원문 URL
- [20] NIST, AI Risk Management Framework. 원문 URL
- [21] 지식재산처, AI 전환 관련 발표. 원문 URL
- [22] 지식재산처, 디자인 검색 관련 발표. 원문 URL