

RESEARCH ARTICLE

Explainability Requirements for Substantive Design Examination Using AI Examination Support Tools: Focusing on XAI-Based Institutional Safeguards

Joungjae Lee¹, Taeman Kim²

¹Ph.D. Program in Intellectual Property Convergence Department, Chungnam National University; Expert Advisor in Convergence IP Strategy Team, Korea Intellectual Property Strategy Agency, Republic of Korea

²Professor, Department of IP Convergence at the graduate school, Chungnam National University, Republic of Korea

Corresponding Author: Taeman Kim (taeman.kim@cnu.ac.kr)

ABSTRACT

This study examines the explainability requirements for artificial intelligence (AI) examination support tools used in substantive design examinations from an explainable AI (XAI) perspective. Substantive examination combines prior design search, similarity of articles, overall aesthetic impression, and ease of creation. Hence, AI-generated similarity scores, candidate rankings, and visualized similarity-contributing regions must support, rather than replace, legal judgment. Drawing on the Design Protection Act, Design Examination Guidelines, Korean AI legislation, recent design and AI policy research, and domestic and international materials on XAI and trustworthy AI, this study identifies three discrepancies between: image similarity and overall aesthetic impression; search objects and units, and the legal relevance of retrieved candidates; and AI similarity-contributing regions, and legally significant or dominant features. As such, it argues that explainability should not require the full disclosure of model internals. Instead, it should be designed as a procedural structure which enables examiners to connect AI outputs with legal criteria, and review and record case-specific input scope, search objects and units, recommendation reasons, similarity-contributing regions, and decisions to accept, modify, or reject AI outputs. The proposed framework includes differentiated logging, quality management, and retraining controls.

KEYWORDS

substantive examination of designs, AI examination support tools, design similarity assessment, ease of creation, search objects and search units, explainable artificial intelligence (XAI), procedural legitimacy

Open Access

Received: June 23, 2026

Revised: July 06, 2026

Accepted: September 04, 2026

Published: September 30, 2026

Funding: The authors received manuscript fees for this article from Korea Institute of Intellectual Property.

Conflict of interest: No potential conflict of interest relevant to this article was reported.

© 2026 Korea Institute of Intellectual Property



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

원저

AI 심사지원도구 도입에 따른 디자인 실체심사의 설명가능성 요건: XAI 기반 제도 보완방안을 중심으로

이정재¹, 김태만²

¹충남대학교 대학원 지식재산융합학과 박사과정, 한국특허전략개발원 IP융합전략 전문위원

²충남대학교 대학원 지식재산융합학과 교수

교신저자: 김태만 (taeman.kim@cnu.ac.kr)

차례

1. 서론
 - 1.1. 연구배경
 - 1.2. 연구목적과 범위 및 방법
 - 1.3. 선행연구와의 차별성
2. AI 심사지원도구 도입과 디자인심사의 제도적 변화
 - 2.1. 지식재산 행정의 AI 도입 흐름
 - 2.2. 디자인 실체심사에서 AI 심사지원도구의 기능적 범위와 전제
 - 2.3. AI 도입이 심사관 판단구조에 미치는 영향
3. 디자인심사에서 예상되는 AI 도구 활용상의 문제점
 - 3.1. 법적 유사판단과 AI 유사도의 불일치
 - 3.2. 과의존과 인간감독 문제
 - 3.3. 입력자료 보호, 검색대상·검색단위 및 데이터 투명성 문제
 - 3.4. 출원인의 의견제출 대응 문제
 - 3.5. 사후 검증·구제와 책임소재 문제
4. 설명가능 AI(XAI)의 제도적 기능과 국내외 규범적 근거
 - 4.1. XAI의 개념과 공공행정상 의미
 - 4.2. 국내 AI 관련 법제의 투명성·신뢰성 규율과 디자인심사에 대한 시사점
 - 4.3. 국내외 XAI 선행연구와 비교 기준의 시사점
 - 4.4. 기술적 설명가능성과 법적 설명가능성의 구분
5. XAI 기반 디자인심사 지원체계의 제도적 요건
 - 5.1. AI 심사지원도구의 역할과 기본원칙
 - 5.2. AI 검색·산출과 법적 판단의 불일치 유형 및 설명가능성 요건
 - 5.3. 로그 기록, 품질관리 및 재학습 통제
6. 정책적 제언
7. 결론

국문초록

본 연구는 AI 심사지원도구가 지식재산 행정 전반에서 확대되는 흐름을 배경으로, 출원디자인의 실체 심사에서 AI 검색·추천·시각화 도구가 활용되는 경우 요구되는 설명가능성 요건을 XAI 관점에서 검토한다. 디자인 실체심사는 선행디자인 조사, 물품의 동일·유사 판단, 전체적 심미감 및 창작비용이성 판단이 결합되는 영역이므로, AI가 제시하는 이미지 유사도, 후보 추천 및 AI 유사도 기여영역 표시가 법적 판단을 대체해서는 안 된다. 이에 본 연구는 디자인보호법, 디자인심사기준, 국내 AI 관련 법제, 최근 디자인·AI 정책연구 및 국내외 XAI·AI 신뢰성 자료를 분석하고, AI 검색·산출과 법적 판단의 불일치를 이미지 유사도와 전체적 심미감, 검색목적별 검색대상·검색단위와 후보의 법적 관련성, AI 유사도 기여영역과 법적 중요부분 또는 지배적 특징의 세 핵심 유형으로 구체화한다. 연구 결과, 디자인 실체 심사 AI 도구의 설명가능성은 모델 내부 공개보다 사건별 입력자료의 처리범위, 검색대상·검색단위, 후보 추천 이유, AI 유사도 기여영역 및 심사관의 수정·배제 이력을 법적 판단측과 연결하여 검토·기록할 수 있는 절차적 구조로 설계될 필요가 있음을 제시한다.

주제어

디자인 실체심사, AI 심사지원도구, 디자인 유사판단, 창작비용이성, 검색대상·검색단위, 설명가능 인공지능(XAI), 절차적 정당성

1. 서론

1.1. 연구배경

지식재산 행정에서 인공지능 기술의 활용은 이미 정책적 실험 단계를 넘어 심사·검색·분류 및 업무지원 전반으로 확대되고 있다. 특허청은 2021년 AI 기술이 적용된 상표·디자인 이미지 검색시스템을 개통하였고, 2023년에는 「인공지능(AI) 기술을 활용한 특허행정 혁신 이행안(로드맵)(’23~’27)」을 통해 심사·심판 등 특허행정 전반에 AI 기술을 확대 적용하는 방향을 제시하였다.¹⁾ 이러한 흐름은 디자인 실체심사에서도 AI가 선행디자인 검색, 후보 추천, 유사도 산출 및 비교자료 제시를 지원하는 도구로 활용될 가능성을 높이고 있다. 디자인 실체심사는 도면 또는 이미지로 표현된 물품의 형상·모양·색채 또는 이들의 결합이 일으키는 시각적 인상, 물품의 용도와 기능, 선행디자인과의 전체적·종합적 유사 여부, 창작비용이성 판단 등이 결합되는 영역이다.²⁾ AI가 선행디자인 후보군을 검색하거나 유사도 점수, 후보 순위, AI 유사도 기여영역을 제시하더라도 그 결과가 곧바로 디자인보호법상 유사 여부나 창작비용이성 판단을 대체할 수는 없다. AI가 산출한 시각적 근접성은 기술적 검색·추천 결과에 해당하는 반면, 디자인 실체심사의 판단은 일반수요자의 관찰, 전체적 심미감, 공지부분의 비중, 지배적 특징 및 통상의 디자인의 창작 용이성 등을 종합하는 법적·규범적 판단이기 때문이다.³⁾

그러나 AI가 제시한 후보, 유사도 점수 및 기여영역은 심사관의 자료검토와 판단형성에 영향을 미칠 수 있다. 특히 이미지 유사도와 전체적 심미감, 검색대상·검색단위와 후보의 법적 관련성, AI 유사도 기여영역과 법적 중요부분 사이에 불일치가 발생할 수 있으므로, AI 산출을 법적 판단과 구분하여 검토·기록할 수 있는 절차적 장치가 필요하다.

본 연구는 AI 심사지원도구의 도입 자체가 아니라, 출원디자인의 실체심사 과정에서 AI가 심사수단으로 활용될 때 그 기술적 산출이 심사관의 법적 판단형성에 미칠 수 있는 절차적 위험을 사전에 보완할 필요가 있다는 문제의식에서 출발한다. 이에 본 연구는 설명가능 AI(Explainable AI, 이하 XAI)를 알고리즘 내부를 설명하는 기술적 방법론에 한정하지 않고, 심사관이 AI의 검색·추천·시각화 결과를 이해·검토·수정·배제하고 그 과정을 기록할 수 있도록 하는 절차적·제도적 장치로 파악한다.

1.2. 연구목적과 범위 및 방법

본 연구의 목적은 출원디자인의 실체심사에서 AI 검색·추천·시각화 도구가 활용되는 경우, AI의 기술적 산출과 심사관의 법적 판단 사이에서 발생할 수 있는 불일치를 유형화하고 AI 심사지원도구의 도입·운영에 필요한 설명가능성 요건과 운영원칙을 제도설계의 관점에서 도출하는 데 있다.

본 연구에서 ‘AI 심사지원도구’란 출원디자인의 도면·이미지 및 물품정보를 사건별 입력자료로 이용하여 선행디자인 후보의 검색·추천과 AI 유사도 기여영역의 시각화를 통해 심사관의 비교·검토를 지원하는 AI 기반 시스템을 의미한다. 여기서 ‘AI 유사도 기여영역’은 출원디자인과

1) 지식재산처, “인공지능 기반의 상표디자인 이미지 검색시스템 개통”, 지식재산처, <<https://www.kipo.go.kr/ko/kpoBultnDetail.do?menuCd=SCD0200618&ntatcSeq=18743&sysCd=SCD02&aprchId=BUT0000029>>, 작성일: 2021. 2. 15.; 지식재산처, “인공지능으로 특허 심사 더 빠르고 정확하게”, 지식재산처, <<https://www.kipo.go.kr/ko/kpoBultnDetail.do?menuCd=SCD0200618&ntatcSeq=19668&sysCd=SCD02/1000&aprchId=BUT0000029>>, 작성일: 2023. 2. 2.

2) 「디자인보호법」 제2조 제1호, 제33조 제1항·제2항; 지식재산처, “디자인심사기준”, 지식재산처예규 제4호(2025. 11. 28. 시행), 지식재산처, 2025, 177면, 183면, 241-242면.

3) 지식재산처, 위의 자료, 183면.

후보디자인의 유사도 산출에 상대적으로 크게 기여한 것으로 적용된 설명기법이 추정하여 시각화한 이미지 영역을 의미하며, 그 자체가 모델의 실제 판단근거를 완전하게 재현하거나 법적 중요부분 또는 지배적 특징을 의미하지 않는다. 또한 공개형 범용 생성형 AI 서비스나 등록 여부의 자동결정 시스템을 전제하지 않고, 기관의 통제된 처리환경에서 운용되는 디자인 이미지·도면 및 물품정보 기반의 검색·추천·시각화 모델을 대상으로 하며 구체적인 내용은 제2.2절에서 제시한다.

본 연구는 사건별 입력자료, 검색대상 자료 및 모델의 학습·검증자료를 서로 다른 데이터 층위로 구분한다. 사건별 입력자료는 개별 출원의 도면·이미지·물품정보와 심사관이 선택한 특정 형태영역을, 검색대상 자료는 국내외 디자인 공보·등록·공개자료 등 출처와 공개시점이 확인 가능한 선행디자인 자료군을, 학습·검증자료는 모델의 개발·성능평가에 사용되는 관리·승인된 디자인 이미지·도면·물품명·분류정보 및 관련 메타데이터를 의미한다. 본 연구는 학습데이터 이용의 적법성 자체가 아니라, 이들 자료가 디자인 실체심사에 활용될 때 그 범위·출처·처리환경과 AI 산출을 설명·통제하기 위한 제도적 요건을 검토한다.

연구방법은 규범적 제도설계 연구를 기본으로 한다. 디자인보호법과 디자인심사기준 및 최근 디자인·AI 정책자료를 통해 디자인 실체심사의 법적 판단구조와 AI 도입 맥락을 검토하고, 국내 AI 관련 법제와 국내외 XAI·AI 신뢰성 자료를 보조적으로 분석하였다. 구체적인 분석절차는 다음과 같다. 첫째, 디자인보호법과 디자인심사기준에서 물품의 동일·유사성, 전체적 심미감, 지배적 특징 및 창작비용이성과 관련된 판단요소를 추출하였다. 둘째, AI 심사지원도구의 입력·검색·추천·시각화 기능을 각 판단요소와 대응시켰다. 셋째, AI의 기술적 처리단위와 법적 판단단위가 일치하지 않을 수 있는 지점을 결과, 검색대상·검색단위 및 시각적 근거의 세 층위로 구분하였다. 넷째, 각 불일치 유형에 XAI 및 신뢰 가능한 AI 문헌에서 제시된 설명, 인간감독, 기록 및 품질관리 원칙을 적용하여 디자인 실체심사 지원체계의 제도적 요건을 도출하였다.

1.3. 선행연구와의 차별성

AI와 지식재산에 관한 선행연구는 먼저 AI가 창작과정에 관여하는 경우의 권리귀속, 보호적격성 및 인간의 창작적 기여를 중심으로 전개되어 왔다. 특히 디자인 분야에서는 AI를 활용하여 생성된 디자인이 현행 디자인보호법상 등록대상이 될 수 있는지, 인간의 창작적 기여를 어떻게 판단할 것인지, AI 활용 사실을 표시할 필요가 있는지 등이 정책적·법적 쟁점으로 검토되고 있다.⁴⁾

최근 지식재산처의 정책연구도 AI 활용 디자인의 등록적격성, 인간의 창작적 기여, AI 활용 표시 및 등록요건 관련 쟁점을 종합적으로 검토하였다.⁵⁾ 이러한 연구는 AI가 심사대상 디자인의 생성·창작과정에 관여한 경우를 주된 대상으로 한다는 점에서, 출원디자인의 실체심사에 AI가 심사수단으로 활용되는 경우를 다루는 본 연구와 구별된다.

생성형 AI와 학습데이터 이용에 관한 연구는 AI 모델의 개발·학습 단계에서 디자인 데이터가 이용되는 경우의 권리침해, 데이터 이용의 적법성 및 생성형 AI 산출물의 법적 평가를 중심으로 논의되어 왔다.⁶⁾ AI 모델이 어떠한 데이터를 학습하였는지, 학습데이터로 이용된 디자인에 대

4) 이정재·김태만, “생성형 AI 디자인 보호체계 연구: 요부(要部) 형성에 대한 인간의 창작적 기여 기준”, 『지식재산연구』, 제21권 제1호(2026), 93-115면; 지식재산처, “인공지능을 활용한 디자인의 법적 쟁점 연구”, 지식재산처, 2025, 20-61면.

5) 지식재산처, 위의 보고서(주 4), 19-129면. 해당 보고서는 AI 활용 디자인의 등록적격성, 인간의 창작적 기여, AI 활용 표시, 신규성, 창작비용이성 및 등록 이후의 권리관계와 분쟁 가능성을 검토한다. 보고서 표지 발행일은 2025. 12.이고, 지식재산처 누리집 공개일은 2026. 6. 8.이다.

6) 지식재산처, 위의 보고서(주 4), 105-106면.

한 권리침해가 발생할 수 있는지, 생성형 AI를 활용한 창작물의 창의성을 어떻게 평가할 것인지에 초점을 둔다.⁷⁾ 그러나 본 연구는 학습데이터 이용 자체의 적법성을 직접 판단하는 것이 아니라, 디자인 실체심사에서 사건별 입력자료, 검색대상 자료 및 학습·검증자료가 서로 다른 층위에서 활용될 때 심사관과 출원인이 AI 산출의 의미와 한계를 검토할 수 있도록 설명가능성 요건의 도출에 초점을 둔다.

XAI와 AI 신뢰성 연구는 AI 산출에 대한 설명, 인간감독, 추적가능성 및 변경관리의 필요성을 검토해 왔으며,⁸⁾ 공공부문 AI 연구는 행정기관의 AI 활용에서 투명성·책임성 및 권리구제 가능성을 강조해 왔다.⁹⁾ 이러한 논의는 본 연구의 이론적·규범적 기반이 되지만, 구체적 내용은 제4장에서 검토한다.

이상의 선행연구들은 AI 창작물의 권리귀속, 생성형 AI와 학습데이터 이용, XAI 일반론 및 공공부문 AI 통제에 관하여 중요한 논의를 제공한다. 그러나 출원디자인의 실체심사에서 AI 심사지원도구의 검색·추천·시각화 산출을 디자인심사의 법적 판단구조, 즉 전체적 심미감, 검색 목적별 검색대상·검색단위, 법적 중요부분 또는 지배적 특징의 판단과 연결하여 설명가능성 요건을 분석한 연구는 제한적이다. 본 연구는 이러한 연구공백을 보완하기 위하여 AI의 기술적 산출과 디자인 실체심사의 법적 판단 사이의 불일치를 유형화하고, 이를 심사관의 독립적 판단권, 출원인의 절차적 대응 가능성 및 사후 검증 가능성을 확보하기 위한 XAI 기반 제도요건으로 구체화한다.

2. AI 심사지원도구 도입과 디자인심사의 제도적 변화

2.1. 지식재산 행정의 AI 도입 흐름

특허청은 2021년 AI 기술이 적용된 상표·디자인 이미지 검색시스템을 개통하여 심사·심판업무에 활용하기 시작하였다. 이 시스템은 특허청이 보유한 상표·디자인 이미지를 학습데이터로 활용하여 구축되었고, 출원 이미지와 선행 등록 이미지를 비교한 검색결과를 유사도 순으로 제시하며, 복수 형상이 결합된 이미지에서 부분 이미지를 인식하여 검색할 수 있도록 지원한다.¹⁰⁾ 이후 특허청은 2023년 「인공지능(AI) 기술을 활용한 특허행정 혁신 이행안(로드맵)('23~'27)」을 발표하고, 심사·심판을 포함한 특허행정 전반에 AI 기술을 확대 적용하는 방향을 제시하였다.¹¹⁾

이러한 정책 흐름은 AI가 단순한 행정자동화 수단을 넘어 심사관의 검색·비교·검토 업무를 지원하는 방향으로 확대되고 있음을 보여준다. 한편 지식재산처의 「인공지능을 활용한 디자인의 법적 쟁점 연구」는 AI 활용 디자인의 등록적격성, 인간의 창작적 기여, AI 활용 표시, 신규성 및 창작비용이성 등을 검토하면서, AI 활용 사실의 투명한 관리와 인간 역할의 식별 필요성을 제시한다.¹²⁾ 이는 AI가 심사대상 디자인의 창작·생성과정에 관여하는 경우를 주된 대상으로 한다는

7) 조병철·방준성, “생성형 인공지능을 활용한 문화예술의 창작과 창의성 평가에 관한 연구”, 『방송공학회논문지』, 제29권 제5호(2024), 596-605면. 이 논문은 디자인 실체심사를 직접 다루는 연구는 아니지만, 생성형 AI 활용 창작물의 창의성 평가에 관한 인접 연구로 참조한다.

8) Marco Tulio Ribeiro et al., “Why Should I Trust You?: Explaining the Predictions of Any Classifier”, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135-1144; National Institute of Standards and Technology, “Artificial Intelligence Risk Management Framework (AI RMF 1.0)”, NIST AI 100-1, National Institute of Standards and Technology, 2023, pp. 15-16.

9) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 제1조, 제31조; 「인공지능 및 데이터 기반 행정 활성화에 관한 법률」 제1조, 제2조 제3호, 제3조 제5항·제7항[2026. 8. 28. 시행 예정].

10) 지식재산처, 앞의 “인공지능 기반의 상표디자인 이미지 검색시스템 개통”(주 1).

11) 지식재산처, 앞의 “인공지능으로 특허 심사 더 빠르고 정확하게”(주 1).

점에서 본 연구와 연구대상을 달리한다. 다만 AI와 인간의 역할이 결합되는 절차에서 AI 관여의 범위와 인간의 판단경위를 구분·관리해야 한다는 문제의식은, 디자인 실체심사에서 AI 검색·추천·시각화 산출과 심사관의 법적 판단을 구별·기록하려는 본 연구의 XAI 기반 제도설계와 접점을 가진다.

이러한 정책과 개발사업이 디자인심사 AI 도구의 구체적 운영방식이나 설명가능성 기준을 직접 제시한 것은 아니므로, 디자인심사에서 AI 검색·분류·후보 추천이 심사관의 판단과정에 미치는 영향은 별도로 검토할 필요가 있다. 디자인심사에서 AI가 특히 유용할 수 있는 영역은 선행디자인 조사이다. 디자인심사기준은 선행디자인 조사를 신규성, 창작비용이성 및 선출원 요건 등을 판단하기 위한 선행디자인 발견 절차로 설명하고, 출원디자인과 전체 또는 각 부분의 형태에 공통점을 가지는 선행디자인을 참증자료로 기재하도록 한다.¹³⁾ 이미지 기반 AI 검색은 후보자료 탐색의 폭과 속도를 높일 수 있지만, 선행디자인 후보를 많이 찾는 것과 이를 법적 거절이유로 구성하는 것은 별개의 문제이다. 특히 신규성·유사 여부 판단과 창작비용이성 판단은 검색목적과 판단자료의 성격이 다르므로, AI가 제시한 후보의 검색대상·검색단위와 법적 관련성은 심사관이 별도로 검토해야 한다. 이 문제는 제5.2절에서 AI 검색·산출과 법적 판단의 불일치 유형으로 구체화한다.

2.2. 디자인 실체심사에서 AI 심사지원도구의 기능적 범위와 전제

본 연구가 전제하는 AI 심사지원도구는 출원디자인의 실체심사에서 심사관의 선행디자인 검색, 후보 비교, 유사도 기여영역 확인 및 판단경위 기록을 지원하는 도구이다. 특정한 단일 시스템의 기술적 구조나 구현방식을 제안하려는 것이 아니라, 설명가능성 요건을 논의하기 위해 AI 심사지원도구가 수행하는 기능적 범위와 한계를 먼저 정리한다. 이에 따라 본 절에서는 AI 심사지원도구의 시스템 성격, 기본 기능, 데이터 층위, 주요 출력, 적용 단계 및 심사관의 역할을 다음과 같이 전제한다. 해외 기술연구에서도 디자인 특허의 이미지, 텍스트 및 분류정보를 결합한 검색·분류 모델이 논의되고 있어, 디자인 분야 AI 검색도구는 이미지 특징뿐 아니라 물품정보와 메타데이터를 함께 활용할 가능성을 보여준다.¹⁴⁾

<표1 디자인 실체심사 AI 심사지원도구의 기능적 범위>

구분	본 연구의 전제
시스템 성격	디자인 실체심사 지원용 AI 검색·추천·시각화 시스템
기본 기능	이미지 특징 추출, 물품명·분류정보 기반 검색, 후보 추천, 유사도 기여영역 표시, 판단경위 기록 보조
학습·검증자료	관리·승인된 디자인 이미지, 도면, 물품명, 물품분류, 공보자료 및 관련 메타데이터
사건별 입력자료	출원디자인의 전체 도면·이미지 또는 심사관이 지정한 특정 형태영역, 물품명·분류정보
검색대상 자료	국내외 디자인 공보, 등록·공개자료 및 출처·공개시점 확인이 가능한 외부 공개자료

12) 지식재산처, 앞의 보고서(주 4), 19-88면.

13) 지식재산처, 앞의 “디자인심사기준”(주 2), 241-242면. 해당 부분은 선행디자인 조사의 목적, 조사대상 및 참증자료의 기재에 관한 기준을 제시한다.

14) Zhu Wang et al., “DesignCLIP: Multimodal Learning with CLIP for Design Patent Understanding”, Findings of the Association for Computational Linguistics: EMNLP 2025, 2025, pp. 10468-10490.

주요 출력	후보디자인 목록, 후보 추천 이유, 유사도 점수 또는 순위, AI 유사도 기여영역, 공통·차이 특징
적용 단계	선행디자인 조사, 유사 여부 검토 지원, 창작비용이성 판단을 위한 후보 탐색 및 비교 지원
AI가 수행하지 않는 영역	인간창작성 판단, 권리귀속 판단, 법적 유사 여부 확정, 창작비용이성 최종판단, 등록·거절 자동결정
심사관의 역할	입력영역·검색범위 조정, 후보 채택·수정·배제, 법적 중요부분 판단, 최종 인용·자료 선택, 거절이유의 법적 구성

위 기능적 범위에서 AI는 기술적 검색·추천·시각화 기능을 수행하지만, 디자인보호법상 등록 요건에 관한 최종 판단은 심사관에게 남는다. 이 도구의 구체적인 설명가능성은 AI 산출을 심사관의 법적 검토 및 판단기록과 연결하는 구조로 이해되며, 그 제도적 의미는 제4장에서 검토한다.

디자인 실체심사에서 AI 도구의 지원과정은 크게 세 단계로 구분할 수 있다. 첫째, 입력·질의 설정 단계에서 심사관은 출원디자인의 전체 도면·이미지 또는 특정 형태영역과 물품정보를 검색 입력자료로 설정할 수 있다. 전체 디자인을 입력하는 경우에는 전체 윤곽·비례·결합관계를 중심으로 후보를 탐색할 수 있고, 특정 영역을 입력하는 경우에는 형상·모양·색채 등의 형태요소를 중심으로 검색할 수 있다. 시스템은 어떤 도면·영역과 물품정보가 질의자료로 사용되었는지를 사건별로 표시·기록할 필요가 있다. 둘째, 선행디자인 검색·후보 추천 단계에서 AI는 설정된 입력자료를 기초로 검색대상 자료에서 시각적으로 또는 의미적으로 근접한 후보를 제시할 수 있다. 이때 시스템은 신규성·유사 여부와 창작비용이성 등 심사목적에 따라 검색대상과 검색단위를 구분하여 설정할 수 있어야 한다.¹⁵⁾ 구체적인 구분과 법적 의미는 제5.2절에서 검토한다.

셋째, 비교·설명 및 기록 지원 단계에서 AI는 출원디자인과 후보디자인의 유사영역·차이영역, 반복형상, 윤곽선, 비례 및 색채 분포 등을 시각화하여 심사관의 비교·검토를 지원할 수 있다. 이때 AI가 표시한 유사도 기여영역은 법적 중요부분 또는 지배적 특징과 동일하지 않으며, 최종 인용자료의 선택과 유사 여부·창작비용이성 판단은 심사관이 수행해야 한다. 시스템은 AI 산출을 법적 결론으로 자동 전환하는 것이 아니라, 후보 추천 이유와 심사관의 채택·수정·배제 및 최종 판단근거를 구조화하여 기록하는 방식으로 지원되어야 한다.

2.3. AI 도입이 심사관 판단구조에 미치는 영향

AI 도구의 도입은 심사관의 판단구조에 영향을 미친다. AI가 상위 후보군을 제시하면 심사관이 먼저 검토하는 자료의 순서와 범위가 달라지고, 이미지 특징·유사도 점수·후보 순위라는 기술적 언어가 전체적 심미감·지배적 특징·창작비용이성이라는 법적 판단언어와 결합된다. 또한 AI 결과가 거절이유 형성에 영향을 미쳤음에도 그 입력값, 검색범위, 후보 추천 이유 및 심사관의 채택·수정·배제 이력이 기록되지 않으면 사후 검증이 곤란해진다. 이와 같이 AI 심사지원도구는 단순 검색도구가 아니라 심사관의 판단형성과정에 영향을 미치는 절차적 장치로 이해되어야 한다.

15) 「디자인보호법」 제33조 제2항; 지식재산청, 앞의 “디자인심사기준”(주 2), 180-189면, 234-246면, 제3부 제2장 및 제5부 제2장 참조. 다만 본 연구에서 사용하는 ‘디자인 단위 검색’과 ‘형태요소 단위 검색’은 「디자인보호법」 제33조 제2항이 구분하는 판단자료의 차이를 AI 검색기능의 설계단위로 전환한 저자의 제안이며, 현행 법령이나 디자인심사기준에 명시된 검색방식 자체를 의미하지 않는다.

3. 디자인심사에서 예상되는 AI 도구 활용상의 문제점

3.1. 법적 유사판단과 AI 유사도의 불일치

AI 이미지 유사도와 디자인보호법상 유사판단은 구분되어야 한다. AI 이미지 검색은 대량의 이미지 자료를 빠르게 비교하여 시각적으로 근접한 후보를 찾는 데 강점이 있지만, 그 유사도는 주로 형태의 시각적 근접성을 산출하는 기술적 지표에 해당한다. 반면 디자인보호법상 유사판단은 물품의 동일·유사성, 전체적 심미감, 공지부분의 비중, 지배적 특징, 일반수요자의 관찰을 종합하는 규범적 판단이다. 디자인보호법 제2조 제1호는 디자인을 물품의 형상·모양·색채 또는 이들의 결합으로서 시각을 통하여 미감을 일으키게 하는 것으로 정의하고, 디자인심사기준 역시 디자인의 동일·유사 여부 판단에서 물품의 동일·유사성과 형태의 동일·유사성을 함께 고려한다.¹⁶⁾ 이에 AI의 유사도 점수는 선행디자인 탐색을 위한 참고자료일 수는 있으나, 그 자체가 법적 결론은 아니다.

AI가 상위 후보를 제시하더라도 심사관은 물품류, 검색범위, 추천 이유 및 유사도 산출에 영향을 준 특징을 디자인심사의 법적 판단기준에 따라 별도로 검토해야 한다. 구체적인 불일치 유형과 설명요건은 제5.2절에서 검토한다.

3.2. 과의존과 인간감독 문제

AI 추천 결과는 심사관의 탐색 범위와 검토 순서를 사실상 조정할 수 있다. 상위 후보군이 먼저 제시되면 심사관은 그 후보군에 집중할 가능성이 높고, 낮은 순위의 자료나 AI가 검색하지 못한 자료를 간과할 수 있다. 그러나 디자인심사에서는 기술적 유사도 점수가 낮은 자료라도 법적으로 중요한 선행디자인일 수 있고, 반대로 시각적으로 유사해 보이는 자료라도 물품의 용도·기능이나 공지부분의 비중에 따라 법적 관련성이 낮을 수 있다. AI 시스템은 단일 점수나 순위만을 제공해서는 안 되며, 심사관이 검색범위를 조정하고 후보를 추가·배제하며 AI 표시영역을 수정할 수 있는 상호작용 구조가 제공되어야 Human-in-the-loop(HITL)¹⁷⁾ 원리가 형식적 장치에 그치지 않는다.

3.3. 입력자료 보호, 검색대상·검색단위 및 데이터 투명성 문제

디자인 실체심사 AI 도구에서는 사건별 입력자료, 검색대상 자료 및 학습·검증자료를 구분할 필요가 있다. 사건별 입력자료는 출원디자인의 전체 도면·이미지 또는 심사관이 지정한 특정 형태영역과 물품정보이고, 검색대상 자료는 선행디자인 후보를 탐색하는 국내외 공보·공개자료 등의 자료군이며, 학습·검증자료는 AI 모델의 개발과 평가에 이용되는 데이터이다. 이들 데이터의 성격이 구분되지 않으면 특정 출원자료가 어떤 처리환경에서 이용되고 저장·재이용되는지, AI가 어떠한 자료군을 대상으로 후보를 탐색하였는지를 확인하기 어렵다. 특히 비공개 출원자료를 외부 AI 서비스로 전송·저장하거나 후속 모델의 학습자료로 재이용하는 경우에는 출원정보의 유출과 제3자 접근 위험이 발생할 수 있으므로, 기관의 통제된 처리환경, 외부전송 통제 및 사건별 입력자료와 학습·검증자료의 분리 원칙이 필요하다. 그러나 출원디자인을 AI 검색의 입력자료로 이용하였다는 사실 자체가 곧바로 디자인보호법상 신규성 상실을 의미하는 것은

16) 지식재산처, 앞의 “디자인심사기준”(주 2), 177면.

17) 과학기술정보통신부·한국정보통신기술협회, “2024 신뢰할 수 있는 인공지능 개발 안내서: 일반 분야”, 과학기술정보통신부·한국정보통신기술협회, 2024, 87면. 본 연구에서 HITL은 AI가 제시한 결과를 인간이 검토·수정·배제 또는 승인하고 최종 판단을 수행하는 운영구조를 의미한다.

아니다. 신규성 판단에서는 출원 전 공지·공용되었거나 간행물 또는 전기통신회선을 통하여 공중이 이용할 수 있게 되었는지가 문제되므로, AI 처리과정에서 실제 외부 공개 또는 공중의 이용가능성이 발생하였는지와 그 시점은 별도로 검토되어야 한다.¹⁸⁾

검색대상 자료의 법적 신뢰성도 설명되어야 한다. AI 검색도구가 신규성·유사 여부 판단을 위한 것인지 창작비용이성 판단을 위한 것인지, 공지디자인을 검색한 것인지 주지 형상·모양·색채의 후보를 탐색한 것인지가 불명확하면 검색결과와 법적 의미를 평가하기 어렵다. 또한 외부 인터넷 자료나 AI 생성 가능성이 있는 이미지가 후보로 제시되는 경우에는 자료원, 공개시점 및 원자료의 확인 가능성이 확보되지 않은 자료를 곧바로 거절이유인 인용자료로 전환해서는 안 된다.

시스템은 검색대상 기간·국가·물품분류, 자료원과 검색단위, 데이터 업데이트 시점 및 후보 자료의 출처 확인 가능성을 제시할 필요가 있다. 설명범위는 이용자와 목적에 따라 달라져야 한다. 심사관에게는 사건별 입력·검색범위와 후보 추천의 근거·한계를 검토할 수 있는 정보를, 출원인에게는 최종 인용자료와 법적 판단논리를 이해·반박할 수 있는 정보를 제공해야 한다. 구체적인 기록과 제공범위는 제5.3절에서 검토한다.

3.4. 출원인의 의견제출 대응 문제

AI 심사지원도구가 거절이유 형성에 관여할 경우, 출원인의 의견제출 및 불복 대응 가능성이 약화될 수 있다. 디자인보호법 제62조와 제63조에 따르면, 심사관이 디자인등록거절결정을 하려면 사전에 거절이유를 통지하고 출원인에게 의견서 제출 기회를 부여하여야 한다. 이는 출원인이 심사관의 판단근거를 이해하고 반박할 수 있도록 하는 절차적 장치이다. 심사관이 AI가 제시한 후보를 인용하는 경우에도 해당 후보가 어떠한 비교 단위에서 유사한지, 어떤 특징이 문제되는지, 물품의 동일·유사성과 전체적 심미감이 어떻게 판단되었는지가 설명되어야 한다. 특히 디자인은 도면과 이미지의 비교가 중심이 되므로, 필요한 경우 비교도, 차이점 표, AI 유사도 기여영역 표시자료 및 선행디자인 출처가 함께 제공될 필요가 있다. 다만 출원인에게 제공되는 설명은 AI의 모든 내부 처리이력이 아니라, 최종 인용자료와 비교관점 및 거절이유의 법적 논리를 이해·반박할 수 있는 범위로 한정하는 것이 적절하다.

3.5. 사후 검증·구제와 책임소재 문제

AI 도구가 심사과정에 활용되면 사후 검증의 대상은 인용 선행디자인과 거절이유뿐 아니라 AI 입력값, 검색범위, 모델 버전, 추천 후보군 및 심사관의 채택·수정·배제 이력까지 확장된다. AI가 부적절한 후보를 제시하고 심사관이 이를 수용한 경우뿐 아니라, 법적으로 관련성 있는 후보를 배제한 경우에도 최종 결정의 타당성이 문제될 수 있다.

그러나 AI 결과와 심사관의 판단 사이에 불일치가 존재한다는 사정만으로 곧바로 위법이나 책임이 인정되는 것은 아니다. 핵심은 심사관이 디자인보호법상 판단기준에 따라 AI 결과를 독립적으로 검토하고, 필요한 추가검색·수정·배제를 거쳐 합리적인 판단근거를 형성하였는지에 있다. 거절결정은 재심사 또는 거절결정불복심판으로 다룰 수 있고, 설정등록된 디자인은 법정 무효사유가 있는 경우 무효심판의 대상이 된다.¹⁹⁾ 이에 로그는 책임을 자동 확정하는 자료가

18) 「디자인보호법」 제33조 제1항. 신규성 판단은 출원 전 국내외에서 공지·공용되었거나 간행물 또는 전기통신회선을 통하여 공중이 이용할 수 있게 된 디자인인지 여부를 기준으로 하므로, AI 검색의 사건별 입력 자체와 공중 이용가능성 발생 여부는 구별하여 검토할 필요가 있다.

19) 「디자인보호법」 제64조, 제120조, 제121조. 각 조문은 재심사, 디자인등록거절결정 등에 대한 심판 및 이

아니라, 심사관의 검토경위와 최종 결정의 적법성을 사후에 복원하는 자료로 기능해야 한다.

4. 설명가능 AI(XAI)의 제도적 기능과 국내외 규범적 근거

4.1. XAI의 개념과 공공행정상 의미

XAI는 일반적으로 AI 시스템의 결과와 그 도출과정을 사람이 이해할 수 있도록 설명하는 기술 또는 방법론을 의미한다. 그러나 공공행정 영역에서 XAI는 단순한 기술적 해석 가능성을 넘어선다. 행정기관이 AI를 활용하여 국민의 권리·의무와 관련된 판단을 보조하는 경우, 설명가능성은 행정결정의 수용가능성, 책임성, 사후 검증 가능성을 확보하는 절차적 요건이 된다.

본 연구에서 XAI 기반 제도요건은 AI 심사지원 도구와 별개의 AI 심사시스템을 의미하지 않는다. 이는 AI 심사지원 도구가 산출한 검색·추천·시각화 결과를 심사관이 이해하고 디자인보호법상 판단촉에 따라 검토·수정·배제하며, 그 판단경위를 기록·검증할 수 있도록 하는 설명 및 통제구조를 의미한다. 이에 본 연구에서 XAI는 기술적 설명가능성, 법적 설명가능성 및 절차적 기록가능성을 포괄하여 AI의 기술적 산출을 디자인 실체심사의 법적 판단과 연결하는 제도적 장치이다.

4.2. 국내 AI 관련 법제의 투명성·신뢰성 규율과 디자인심사에 대한 시사점

국내에서도 AI 활용의 투명성과 신뢰성 확보는 법제화 단계에 진입하였다. 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(이하 “AI기본법”)은 인공지능의 발전과 신뢰 기반 조성을 함께 규율하면서, 투명성 확보, 고영향 인공지능의 확인, 안전성·신뢰성 확보 및 영향평가에 관한 기준을 규정한다.²⁰⁾ 다만 AI기본법은 인공지능사업자를 중심으로 의무체계를 구성하고 있으므로, 지식재산처 내부의 디자인심사 AI 지원도구에 직접 적용되는 개별 심사규범이라고 단정하기는 어렵다. 그럼에도 AI기본법이 규정한 투명성, 인간의 관리·감독, 설명 및 문서화 원칙은 지식재산처 내부 심사지원도구에 직접 적용되는 개별 의무라기보다, 관련 운영기준을 설계할 때 참조할 수 있는 일반적 법정책 방향을 제공한다. 또한 「인공지능 및 데이터 기반 행정 활성화에 관한 법률」은 공공기관의 인공지능기반행정에 관한 제도적 근거를 마련한다는 점에서, 향후 디자인심사 AI 지원체계의 운영기준을 구체화하는 보충적 기준이 될 수 있다.²¹⁾ 이러한 법정책적 흐름은 디자인심사 AI 지원체계에서도 투명성, 인간감독, 기록가능성 및 변경관리 원칙을 운영기준에 반영할 필요가 있음을 시사한다.

4.3. 국내외 XAI 선행연구와 비교 기준의 시사점

국내 XAI 논의는 기술적 블랙박스 문제를 넘어 공공부문 AI 활용의 신뢰성 확보 문제로 확장되고 있다. 한국행정연구원은 공공부문 XAI 활용을 위해 영역별 기준 마련과 적합성 평가체계 구축을 제안하였고,²²⁾ 소프트웨어정책연구소는 학습 완료 딥러닝 모델의 설명기법과 그 한계

자인등록무효심판을 각각 규정한다.

20) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 제1조, 제31조 및 제33조부터 제35조까지. 제31조는 인공지능의 투명성 확보를, 제33조부터 제35조까지는 고영향 인공지능의 확인, 관련 책무 및 영향평가를 규정한다.

21) 「인공지능 및 데이터 기반 행정 활성화에 관한 법률」 제1조, 제2조 제3호, 제3조 제5항·제7항 및 제19조 [2026. 8. 28. 시행 예정]. 본 논문에서 디자인심사 AI 지원체계에 직접 적용되는 현행 규범이 아니라 시행 이후 공공기관의 인공지능기반행정 운영을 위한 보충적 제도 근거로 검토한다.

22) 김문현 외 2인, “설명가능한 인공지능(XAI)의 공공분야 활용을 위한 정책 방향”, 한국행정연구원, 2024.

를 검토하였다.²³⁾ 이러한 논의는 디자인심사와 같이 시각적·규범적 판단이 결합되는 행정영역에서 일반적 XAI 원칙을 그대로 적용하기보다, 심사절차와 판단기준에 맞는 설명가능성 기준을 별도로 설계해야 함을 시사한다.

국외 XAI 연구도 같은 방향을 뒷받침한다. Ribeiro 등은 개별 예측에 대한 설명이 사용자의 신뢰 판단과 오류 식별에 중요하다고 보았고,²⁴⁾ Hoffman 등은 설명이 사용자의 의미구성, 정신모형 형성 및 적절한 신뢰를 지원해야 한다고 본다.²⁵⁾ 또한 「신뢰할 수 있는 인공지능 개발 안내서」는 모델 명세, 추론 결과 설명, 사용자의 이해도 제고, 추적가능성 및 변경이력 확보를 주요 요구사항으로 제시한다.²⁶⁾ 이는 본 연구가 제안하는 데이터 출처 설명, 후보 추천 이유 설명, 심사관 수정 로그, 모델 버전관리 및 재학습 통제와 연결된다.

EU AI Act와 NIST AI RMF는 국내 디자인심사에 직접 적용되는 규범은 아니지만, AI 심사지원도구의 제도적 요건을 검토하는 비교·보조 기준을 제공한다. EU AI Act는 고위험 AI 시스템의 투명성, 정보제공 및 인간감독을 규정하고,²⁷⁾ NIST AI RMF도 설명가능성·해석가능성, 책임성·투명성을 신뢰 가능한 AI의 주요 특성으로 제시한다.²⁸⁾ 이러한 기준은 디자인심사 AI 도구의 설명정보, 심사관 통제, 추적가능성 및 품질관리 요건을 구체화하는 데 활용될 수 있다.

4.4. 기술적 설명가능성과 법적 설명가능성의 구분

디자인심사에 필요한 설명가능성은 기술적 설명가능성과 법적 설명가능성으로 구분할 수 있다. 기술적 설명가능성은 AI가 어떤 데이터, 모델, 특징값을 기초로 특정 후보를 추천했는지 설명하는 것이다. 반면 법적 설명가능성은 그 기술적 설명이 물품의 동일·유사성, 전체적 심미감, 공지부분, 지배적 특징, 창작비용이성 등 디자인보호법상 판단요소와 어떻게 연결되는지를 설명하는 것이다. Gilpin 등은 해석가능성과 설명가능성을 구분하고, 설명의 해석가능성과 완전성 사이의 균형이 필요하며 지나친 단순화는 시스템에 대한 오해를 유발할 수 있다고 지적한다.²⁹⁾

심사관에게 필요한 것은 해당 사건에서 어떤 데이터 범위가 검색되었는지, 어떤 후보가 왜 제시되었는지, 어떤 시각적 특징이 유사도 산출에 영향을 주었는지, 그 설명이 디자인심사기준상 어떤 법적 판단축과 관련되는지를 검토할 수 있는 정보이다. 디자인심사용 XAI는 단순한 히트

4-5면, 8면.

23) 최형규, “설명가능한 인공지능(Explainable AI: XAI) 연구 동향과 시사점: 학습이 완료된 딥러닝 모델에 대한 설명을 중심으로”, 소프트웨어정책연구소, 2021, 4-13면.

24) Marco Tulio Ribeiro et al., 앞의 논문(주 8), pp. 1135-1138, 1142-1143. 이 논문은 개별 예측에 대한 설명이 사용자의 신뢰 판단과 오류 식별을 지원하며, 이미지 분류에서는 super-pixel 단위로 예측근거를 제시할 수 있음을 제시한다.

25) Robert R. Hoffman et al., “Evaluating machine-generated explanations: a ‘Scorecard’ method for XAI measurement science”, *Frontiers in Computer Science*, Vol.5(2023), Article No. 1114806, 특히 pp. 1-2, 4-5, 14. 이 논문은 단서 수준 설명의 한계를 지적하고, 설명이 사용자의 센스메이킹, 정신 모형 형성 및 적절한 신뢰를 지원하는지를 평가하기 위한 Explanation Scorecard를 제시한다.

26) 과학기술정보통신부·한국정보통신기술협회, 앞의 자료(주 17), 47-48면, 75-82면, 90-94면. 해당 부분은 AI 시스템의 추적가능성과 변경이력, 모델 및 추론결과에 대한 설명, 사용자의 설명 이해도 제고 등을 신뢰성 요구사항으로 제시한다.

27) European Union, “Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence”, *Official Journal of the European Union*, 2024, Articles 13-14. 제13조는 고위험 AI 시스템의 투명성과 정보제공을, 제14조는 인간감독을 규정한다.

28) National Institute of Standards and Technology, 앞의 보고서(주 8), pp. 15-16.

29) Leilani H. Gilpin et al., “Explaining Explanations: An Overview of Interpretability of Machine Learning”, 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 2018, pp. 80-89, 특히 pp. 80-81. 이 논문은 해석가능성과 설명가능성을 구분하고, 설명의 해석가능성과 완전성 사이의 균형 필요성을 제시한다.

맵이나 유사도 점수가 아니라, AI의 기술적 산출을 심사관이 검토할 수 있는 법적 쟁점으로 번역하는 중간 장치로 설계되어야 한다.

5. XAI 기반 디자인심사 지원체계의 제도적 요건

5.1. AI 심사지원도구의 역할과 기본원칙

앞서 제2.2절에서 전제한 AI 심사지원도구는 심사관의 선행디자인 검색과 비교·검토를 지원하는 절차적 도구이며, 등록요건에 관한 최종 판단을 수행하지 않는다. 이에 AI 산출과 심사관의 법적 판단은 화면과 기록에서 구분되어야 하고, 심사관에게는 검색범위 조정, 후보의 추가·배제, 표시영역 및 거절이유 초안의 수정 권한이 보장되어야 한다. AI 산출과 심사관의 주요 수정·배제 이력은 사건별로 기록하고, 반복되는 오류와 누락은 품질관리 및 재학습 통제에 활용해야 한다.

5.2. AI 검색·산출과 법적 판단의 불일치 유형 및 설명가능성 요건

디자인 실체심사 AI 도구의 설명가능성은 입력, 검색, 산출, 법적 검토의 연결구조로 설계되어야 한다. 즉 사건별 입력자료, 검색목적과 검색범위, 후보 추천 이유, AI 유사도 기여영역 및 심사관의 채택·수정·배제 이력이 법적 판단축과 연결되어야 한다. 특히 창작비용이성 판단에서는 AI가 제시한 후보가 공지디자인 또는 그 결합에 관한 것인지, 국내외에서 널리 알려진 형상·모양·색채 또는 그 결합에 관한 것인지, 전체 디자인 단위의 추천인지 형태요소 단위의 탐색인지를 구분해야 한다. AI 검색·산출과 법적 판단의 불일치는 결과, 검색대상·검색단위 및 근거의 서로 다른 층위에서 나타날 수 있다.

본 연구는 이를 세 가지 핵심 분석유형으로 구분한다. 첫째, AI가 산출한 이미지 유사도 결과와 전체적 심미감 판단의 불일치이다. 둘째, 신규성·유사 여부 또는 창작비용이성이라는 검색 목적에 따른 검색대상·검색단위와 후보의 법적 관련성 사이의 불일치이다. 셋째, AI 유사도 기여영역과 법적 중요부분 또는 지배적 특징의 불일치이다. 즉 제1유형은 AI가 산출한 결과와 법적 결론의 차이, 제2유형은 후보를 탐색한 목적·대상·단위와 법적 관련성의 차이, 제3유형은 결과 산출에 사용된 시각적 근거와 법적 중요부분의 차이를 다룬다. 반복패턴 탐지, 후보 순위 및 거절이유 초안에서 나타나는 문제는 이들 핵심 유형이 실제 기능과 이유제시 단계에서 나타나는 보충적 불일치로 검토한다. 이하의 <그림1>부터 <그림3>은 실제 심사사례가 아니라, AI 검색·산출과 법적 판단 사이에서 발생할 수 있는 불일치 유형을 설명하기 위해 저자가 구성 및 작성한 예시도이다.

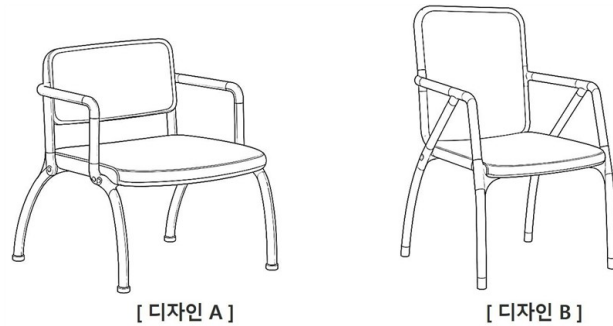
5.2.1. 이미지 유사도와 전체적 심미감의 불일치

AI 이미지 검색은 전체 윤곽, 선의 배치, 면적, 비례 및 반복형상 등 시각적 특징의 근접성을 중심으로 후보를 제시할 수 있다. 그러나 AI가 공통적인 형태요소에 높은 가중치를 부여하더라도, 그 공통요소가 전체적 심미감에 미치는 영향이 제한적이면 법적 유사판단에서의 비중은 달라질 수 있다.³⁰⁾ <그림1>의 두 의자는 공통적인 의자 골격으로 인해 AI 이미지 검색에서 시각적으로 가까운 후보로 제시될 수 있다. 그러나 좌판과 등받이의 비례, 팔걸이와 다리의 결합방식 등 세부 구성의 차이가 일반수요자에게 형성하는 전체적 인상에 영향을 준다면 법적 유사판

30) 지식재산처, 앞의 “디자인심사기준”(주 2), 183면.

단은 AI의 시각적 근접성과 다르게 이루어질 수 있다. 반대로 기술적 유사도가 낮게 나타나더라도 지배적 형상과 결합관계가 공통된다면 법적으로 유사하게 평가될 수 있다. 따라서 시스템은 단일 점수뿐 아니라 유사도 산출에 기여한 영역, 주요 공통점과 차이점 및 전체적 심미감과의 관련성을 심사관이 검토할 수 있도록 제시해야 한다.

<그림1 이미지 유사도와 법적 유사판단의 불일치 가능성>

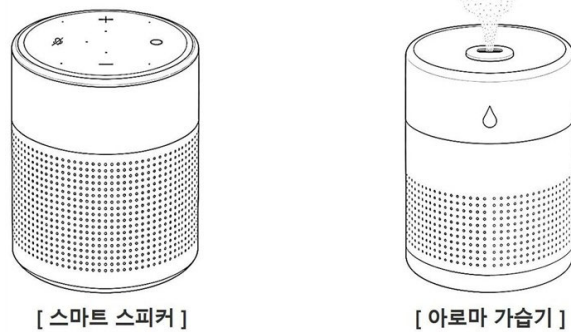


5.2.2. 검색목적별 검색대상·검색단위와 법적 관련성 판단의 불일치

AI는 이미지, 물품명, 디자인분류 및 메타데이터를 이용하여 검색범위를 설정하고 시각적으로 근접한 후보를 제시할 수 있다. 신규성·유사 여부 판단에서는 물품의 동일·유사성이 후보의 법적 관련성을 판단하는 중요한 조건이 된다. 반면 디자인보호법 제33조 제2항에 따른 창작비용이성 판단은 단순한 검색범위 확장 문제로만 설명될 수 없다. 같은 항은 판단자료를 제33조 제1항 제1호·제2호에 해당하는 디자인 또는 이들의 결합과, 국내외에서 널리 알려진 형상·모양·색채 또는 이들의 결합으로 구분하기 때문이다.³¹⁾ 이와 같이 창작비용이성 판단자료의 차이를 AI 검색기능에 반영하기 위해서는, 전체 디자인을 중심으로 후보를 탐색하는 방식과 형상·모양·색채의 요소를 중심으로 탐색하는 방식을 기능적으로 구분하는 방안을 검토할 필요가 있다. 전자는 출원디자인과 비교 가능한 선행디자인 자료를 전체 디자인 단위로 탐색하고 복수 공지디자인의 결합 가능성을 검토하는 방식이다. 후자는 완결된 디자인 자체보다 특정 윤곽, 패턴, 표면 모양, 색채 조합 등 형태요소가 국내외에서 널리 알려져 있는지를 검토하는 방식이다. 다만 AI가 형태요소를 추출하거나 반복성을 표시한다고 하여 그 요소가 곧바로 법적으로 주지형태에 해당하는 것은 아니며, 통상의 디자이너가 이를 쉽게 이용·변형·결합할 수 있었는지는 심사관이 독립적으로 판단해야 한다. <그림2>의 스마트 스피커와 아로마 가습기 사례는 이종 물품 간 시각적 근접성 자체가 곧 창작비용이성 부정을 의미한다는 사례가 아니라, AI가 확장검색 과정에서 전체 디자인 단위의 공지디자인 후보 또는 원통형 실루엣·상하 비례·하부 천공 패턴 등 형태요소 단위의 주지형태 후보를 표시할 수 있음을 보여주는 예시이다. 심사관은 해당 후보가 공지디자인인지 주지형태인지, 전체 디자인 단위의 비교인지 형상·모양·색채 요소 단위의 검토인지, 그 외관적 특징이 출원디자인의 창작수준 판단과 어떻게 연결되는지를 구별하여 판단해야 한다.

31) 「디자인보호법」 제33조 제2항 제1호·제2호. 동 조항은 창작비용이성 판단자료를 제33조 제1항 제1호·제2호에 해당하는 디자인 또는 이들의 결합과, 국내 또는 국외에서 널리 알려진 형상·모양·색채 또는 이들의 결합으로 구분한다. 다만 '디자인 단위 검색'과 '형태요소 단위 검색'은 현행 법령이나 디자인심사기준에 명시된 용어가 아니라, 법정 판단자료의 차이를 AI 검색설계에 반영하기 위하여 본 논문에서 제안하는 분석상 구분이다.

<그림2 검색목적별 검색대상·검색단위와 후보의 법적 관련성 판단의 불일치 가능성>



5.2.3. AI 유사도 기여영역과 법적 중요부분 또는 지배적 특징의 불일치

AI가 표시하는 유사도 기여영역은 적용된 설명기법이 유사도 산출에 상대적으로 크게 기여한 것으로 추정하여 시각화한 영역으로서, 디자인심사에서의 법적 중요부분 또는 지배적 특징과 동일하지 않다. 모델은 면적이 크거나 반복되는 공통형상을 중요하게 반영할 수 있지만, 디자인의 법적 유사 여부는 일반수요자의 관찰을 기준으로 전체적 심미감을 종합하여 판단하므로, AI가 표시한 영역의 면적이거나 반복성만으로 그 법적 중요성을 확정할 수 없다.³²⁾ <그림3>의 두 조명은 원형 받침대와 하부 기둥을 공통적으로 가지므로 AI가 이 영역을 유사도 산출의 중요영역으로 표시할 수 있다. 그러나 해당 부분이 관련 물품에서 통상적인 구성이라면, 갖의 형상, 암의 구조 및 전체 비례가 오히려 전체적 심미감을 좌우하는 지배적 특징으로 평가될 수 있다. 이에 시스템은 표시영역을 ‘법적 중요부분’이 아니라 ‘AI 유사도 산출에 기여한 영역’으로 명시하고, 심사관이 판단한 법적 중요부분을 별도로 기록할 수 있도록 해야 한다.

<그림3 AI 유사도 기여영역과 법적 중요부분 또는 지배적 특징 판단의 불일치 가능성>



5.2.4. 기타 기능상 불일치와 설명·기록요건

AI 검색·산출과 법적 판단의 불일치는 반복패턴 탐지, 후보군 순위 및 거절이유 초안에서도 발생할 수 있다. AI가 반복빈도가 높은 공지패턴을 핵심 특징으로 과대평가하거나, 상위 후보보다 하위 후보가 법적으로 중요한 지배적 특징을 더 직접적으로 포함할 수 있다. 또한 AI가 유사도 점수를 중심으로 거절이유 초안을 작성하더라도 물품의 동일·유사성, 공통점과 차이점, 전체적 심미감 및 창작비용이성에 관한 법적 판단이 누락될 수 있다.

32) 지식재산처, 앞의 “디자인심사기준”(주 2), 183면.

<표2 보충적 AI 검색·산출 기능의 불일치 및 최소 기록요건>

AI 검색·산출요소	불일치 가능 예시	법적 확인사항	기록사항
반복 패턴 탐지	공지패턴을 과대평가	공지성, 창작적 특징	공지근거, 배열 차이
후보군 순위	기술적 순위와 법적 관련성의 불일치	검색목적, 선행디자인 관련성	추가검색, 후보 채택·배제
거절이유 초안	유사도 중심의 이유제시	물품·심미감·창작비용이성	인용자료, 수정내용

이상의 불일치 유형은 AI 결과를 배제해야 한다는 의미가 아니라, 기술적 산출과 법적 판단을 구분하여 검토할 수 있는 설명구조가 필요함을 보여준다. 일반적인 검색·추천 이력은 시스템이 자동으로 기록하고, 심사관의 별도 사유입력은 상위 후보의 배제, 물품분류 수정, 검색범위 변경 및 AI 표시영역 수정 등 최종 판단에 영향을 미치는 주요 불일치에 한정하는 것이 적절하다.

5.3. 로그 기록, 품질관리 및 재학습 통제

앞서 살펴본 불일치는 그 유형과 중요도에 따라 검토·기록 수준을 달리할 필요가 있다. 모델 버전, 사건별 입력대상과 입력단위, 심사목적과 검색모드, 검색대상·검색단위, AI 후보군, 최종 인용자료 및 심사관의 채택·수정·배제 이력은 시스템이 자동으로 기록해야 한다. 입력자료의 외부전송·저장·재이용 여부와 모델의 변경·재학습 이력은 시스템 단위의 감사기록으로 관리할 필요가 있다. 심사관의 별도 사유입력은 상위 후보 배제, 입력영역 또는 검색대상·검색단위 변경, 물품분류 및 AI 표시영역의 수정·배제, 법적 중요부분의 별도 지정 등 최종 판단에 영향을 미치는 경우에 한정하는 것이 적절하다. 현저한 불일치나 선행디자인 누락 위험이 큰 사건은 추가검색 또는 동료검토 대상으로 관리할 수 있다. 출원인에게는 AI 후보의 모든 처리이력이 아닌 최종 인용자료와 법적 판단근거를 이해·반박할 수 있는 수준의 설명을 제공하고, 내부 로그는 판단경위의 사후 복원, 반복 오류의 개선 및 AI 도구의 품질관리에 활용해야 한다.

AI 도구의 품질관리는 검색 정확도에 한정되어서는 안 된다. 설명가능 AI 연구는 설명의 품질을 모델 성능과 별도로 평가해야 한다고 보며, 특히 Hoffman 등은 설명이 사용자의 이해, 실패 가능성 인식, 적정한 신뢰 형성을 지원해야 한다고 본다.³³⁾ 디자인심사 AI 도구의 품질관리는 후보 검색 성능뿐 아니라 설명품질, 심사관 이해가능성, 과의존 방지, 수정·배제 로그의 충실성까지 포함하여 설계되어야 하며, 다음과 같은 다층적 지표는 AI 도구의 성능과 제도적 신뢰성을 함께 평가하기 위한 것이다.

<표3 디자인심사 AI 도구의 품질관리 지표>

평가차원	예시 지표	목적
검색성능	검색목적별 상위 후보 재현율, 누락률, 오탐률	후보 탐색의 품질 확인
설명품질	추천이유 이해가능성, AI 유사도 기여영역 표시 적절성	심사관의 비판적 검토 지원
절차품질	거절이유 기재 충실도, 의견제출 대응 가능성	절차적 정당성 확보
인간감독	수정·배제율의 변화, 재검색 요청률, 검색단위 변경률	과의존 위험 점검
거버넌스	검색모드·범위 기록, 대상·단위 기록, 로그 완결성, 모델 버전관리, 재학습 검증이력	사후 검증과 책임성 확보
사용자 경험	심사관 인지부하, 사용 만족도, 교육 필요도	현장 수용성 확보

33) Robert R. Hoffman et al., 앞의 논문(주 25), pp. 1-2, 4-5, 14.

AI 심사지원도구는 지속적으로 개선되어야 하지만, 재학습은 통제되어야 한다. 심사관의 수정이력을 그대로 자동학습에 반영할 경우 특정 심사관의 판단 경향, 특정 물품류의 편향, 일시적 정책 변화가 모델에 과도하게 반영될 수 있다. 재학습은 운영부서, 기술검증부서, 법률검토부서의 검토를 거쳐 승인되어야 하며, 모델 업데이트 전후의 검색결과 변화와 설명품질 변화가 사후 평가되어야 한다.

6. 정책적 제언

지식재산청은 디자인 실체심사 AI 도구의 고도화에 앞서, 제5장에서 제시한 설명가능성·인간감독·기록 및 품질관리 요건을 반영한 운영기준을 마련할 필요가 있다. 운영기준은 AI의 기술적 산출과 심사관의 법적 판단을 구분하고, 비공개 출원자료의 외부전송·재이용, 검색목적·대상·단위의 설정, 후보 추천과 심사관의 주요 수정·배제 이력 및 모델 변경을 통제할 수 있도록 구성되어야 한다.

제도 도입은 단계적으로 추진하는 것이 적절하다. 우선 특정 물품류 또는 선행디자인 검색단계에 시범 적용하여 누락·오답, 설명품질, 심사관의 수정행태, 업무부담 및 출원인의 대응 가능성을 평가한 후 적용범위를 확대해야 한다. 심사관 교육은 AI 산출의 기술적 의미와 법적 한계, 자동화 편향, 검색범위 조정 및 수정·배제 권한을 중심으로 실시할 필요가 있다. 품질평가와 재학습은 기술부서만의 성능관리로 운영해서는 안 된다. 운영부서·기술검증부서·법률검토부서가 모델 업데이트 전후의 검색성능, 설명품질 및 심사결과에 미치는 영향을 공동 검토하고, 심사관의 수정이력은 검증과 승인 없이 자동으로 재학습에 반영되지 않도록 해야 한다.

7. 결론

본 연구는 출원디자인의 실체심사에서 AI 검색·추천·시각화 도구가 활용되는 경우 요구되는 설명가능성 요건을 검토하였다. 본 연구에서 AI는 심사대상 디자인의 생성수단이 아니라 심사관의 선행디자인 검색과 비교·검토를 지원하는 심사수단이다. 이에 논의의 초점은 AI 생성디자인의 인간창작성이나 권리귀속이 아니라, AI의 기술적 검색·산출이 물품의 동일·유사성, 전체적 심미감 및 창작비용이성이라는 법적 판단과 연결되는 과정에서 발생할 수 있는 절차적 위험에 있다.

본 연구는 AI 검색·산출과 디자인의 법적 판단 사이의 불일치를 ① 이미지 유사도와 전체적 심미감, ② 검색목적별 검색대상·검색단위와 후보의 법적 관련성, ③ AI 유사도 기여영역과 법적 중요부분 또는 지배적 특징의 세 유형으로 구체화하였다. 특히 창작비용이성 판단자료의 차이를 반영하기 위해 전체 디자인 단위의 후보탐색과 형상·모양·색채 요소 단위의 탐색을 기능적으로 구분하는 방안을 제시하였다. 이를 바탕으로 AI 도구의 설명가능성을 사건별 입력·검색·추천 결과와 심사관의 법적 검토 및 주요 개입 이력을 연결하는 절차적 구조로 제안하였다. 또한 국내 AI 관련 법제, EU AI Act 및 NIST AI RMF는 이러한 제도요건을 뒷받침하는 보조적 기준으로 활용하였다.

본 연구는 공개된 정책자료와 디자인심사의 법적 판단구조에 기초한 규범적 제도설계 연구로서, 실제 디자인심사 AI 시스템의 화면구성, 설명항목, 로그기록 방식 및 성능평가 체계를 실증적으로 분석하지 못했다는 한계를 가진다. 향후 실제 시스템이나 시범운영 자료를 바탕으로 AI 추천 후보의 검색성능과 법적 관련성, 설명유형별 심사관의 이해·수정행태 및 출원인의 대응 가능성을 검토할 필요가 있다.

그럼에도 본 연구의 핵심 기여는 AI 심사지원도구와 디자인 법적 판단 사이의 불일치를 결과, 검색대상·검색단위 및 시각적 근거의 세 층위로 유형화하고, 각 불일치에 대응하는 설명·인간 감독·기록·품질관리 요건을 제시한 데 있다. 이는 AI 심사지원도구 도입을 단순한 기술 고도화가 아니라 디자인심사제도의 절차적 정당성과 신뢰성 확보를 위한 제도설계 과제로 재구성하였다는 점에서 의의를 가진다.

참고문헌

학술지(국내 및 동양)

이정재·김태만, “생성형 AI 디자인 보호체계 연구: 요부(要部) 형성에 대한 인간의 창작적 기여 기준”, 『지식재산연구』, 제21권 제1호(2026).

조병철·방준성, “생성형 인공지능을 활용한 문화예술의 창작과 창의성 평가에 관한 연구”, 『방송공학회논문지』, 제29권 제5호(2024).

학술지(서양)

Robert R. Hoffman et al., “Evaluating machine-generated explanations: a ‘Scorecard’ method for XAI measurement science”, *Frontiers in Computer Science*, Vol.5(2023).

인터넷 자료

지식재산처, “인공지능 기반의 상표디자인 이미지 검색시스템 개통”, 지식재산처, <<https://www.kipo.go.kr/ko/kpoBultnDetail.do?menuCd=SCD0200618&ntatcSeq=18743&sysCd=SCD02&aprchId=BUT0000029>>, 작성일: 2021. 2. 15.

지식재산처, “인공지능으로 특허 심사 더 빠르고 정확하게”, 지식재산처, <<https://www.kipo.go.kr/ko/kpoBultnDetail.do?menuCd=SCD0200618&ntatcSeq=19668&sysCd=SCD02/1000&aprchId=BUT0000029>>, 작성일: 2023. 2. 2.

연구보고서

김문현 외 2인, “설명가능한 인공지능(XAI)의 공공분야 활용을 위한 정책 방향”, 한국행정연구원, 2024.

지식재산처, “인공지능을 활용한 디자인의 법적 쟁점 연구”, 지식재산처, 2025.

최형규, “설명가능한 인공지능(Explainable AI: XAI) 연구 동향과 시사점: 학습이 완료된 딥러닝 모델에 대한 설명을 중심으로”, 소프트웨어정책연구소, 2021.

National Institute of Standards and Technology, “Artificial Intelligence Risk Management Framework (AI RMF 1.0)”, NIST AI 100-1, National Institute of Standards and Technology, 2023.

기타 자료

과학기술정보통신부·한국정보통신기술협회, “2024 신뢰할 수 있는 인공지능 개발 안내서: 일반 분야”, 과학기술정보통신부·한국정보통신기술협회, 2024.

지식재산처, “디자인심사기준”, 지식재산처예규 제4호(2025. 11. 28. 시행), 지식재산처, 2025.

European Union, “Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence”, Official Journal of the European Union, 2024.

Leilani H. Gilpin et al., “Explaining Explanations: An Overview of Interpretability of Machine Learning”, 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 2018.

Marco Tulio Ribeiro et al., “Why Should I Trust You?: Explaining the Predictions of Any Classifier”, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.

Zhu Wang et al., “DesignCLIP: Multimodal Learning with CLIP for Design Patent Understanding”, Findings of the Association for Computational Linguistics: EMNLP 2025, 2025.