

## A Study on the Prospects for Activating Korea's e-Discovery Market and AI Technology Application Strategies

Yongho Kim\*

\*Professor, Dept. of Police Administration, Kyungil University, Gyeongsangbuk-do, Korea

### [Abstract]

In this study analyzes the potential for activating the e-Discovery market under the current legal framework of the Republic of Korea and proposes implementation strategies for innovative e-Discovery solutions utilizing artificial intelligence (AI) technology. The research findings indicate a high likelihood of e-Discovery market activation in Korea, based on the government's commitment to introducing a Korean-style Discovery system, the increasing number of overseas litigations involving domestic companies, and the country's robust IT infrastructure.

In particular, the study confirms that the application of cutting-edge AI technologies such as Technology Assisted Review (TAR), Natural Language Processing (NLP), and generative AI can achieve a 75~90% reduction in time and 30~40% higher accuracy compared to traditional document review methods. This research presents policy implications for the development of the Korean e-Discovery market along with specific utilization strategies for AI-based solutions.

▶ **Key words:** e-Discovery, Electronic Discovery, Artificial Intelligence, Technology Assisted Review, Korean-style Discovery

### [요 약]

본 연구는 대한민국의 현행 법률 체계 하에서 전자증거개시(e-Discovery) 시장의 활성화 가능성을 분석하고, 인공지능(AI) 기술을 활용한 혁신적 e-Discovery 솔루션의 도입 방안을 제시한다. 연구 결과, 정부의 한국형 디스커버리 제도 도입 의지, 국내 기업의 해외 소송 증가, 그리고 우수한 IT 인프라를 바탕으로 한국 e-Discovery 시장의 활성화 가능성이 매우 높은 것으로 분석되었다.

특히 Technology Assisted Review(TAR), 자연어 처리(NLP), 생성형 AI 등 최신 AI 기술의 적용을 통해 전통적인 문서 검토 방식 대비 75~90%의 시간 단축과 30~40% 높은 정확도를 달성할 수 있음을 확인하였다. 본 연구는 한국형 e-Discovery 시장 발전을 위한 정책적 시사점과 함께 AI 기반 솔루션의 구체적 활용 전략을 제시한다.

▶ **주제어:** 전자증거개시, e-Discovery, 인공지능, 기술지원검토, 한국형 디스커버리

## I. Introduction

21세기 디지털 전환 시대를 맞아 기업과 개인이 생성하는 전자정보의 양이 기하급수적으로 증가하고 있다. 국제 데이터 공사(IDC)에 따르면, 전 세계 데이터 생성량은 2025년까지 163 ZB(제타바이트)에 달할 것으로 예상된다 [1]. 이러한 환경 변화는 법정에서 전자증거의 중요성을 크게 부각시켜, 전자증거개시(Electronic Discovery, e-Discovery) 제도의 필요성이 전 세계적으로 확산하고 있다.

미국을 중심으로 한 영미법계 국가들은 이미 2006년 연방민사소송규칙(Federal Rules of Civil Procedure) 개정을 통해 e-Discovery를 공식적으로 제도화하였다[2]. 그 결과 현재 글로벌 e-Discovery 시장은 2024년 기준 약 107억~127억 달러 규모에 달하며, 2030~2035년까지 연평균 7~10%의 지속적인 성장이 예상되고 있다[3].

한편, 우리나라는 2015년 대법원의 '사실심 충실화 마스터플랜' 발표를 계기로 한국형 디스커버리 제도 도입을 본격 추진하기 시작하였으며[4], 2023년에는 중소벤처기업부가 「중소기업기술 보호 지원에 관한 법률」 개정을 통해 제도 도입의 법적 기반을 마련하였다. 이러한 정책적 변화와 함께 국내 기업의 해외 진출 확대에 의한 국제 소송 증가는 한국 e-Discovery 시장의 활성화 필요성은 더욱 부각하고 있다.

본 연구의 목적은 두 가지로 구분된다. 첫째, 대한민국의 현행 법률 체계와 시장 환경을 종합적으로 분석하여 e-Discovery 시장의 활성화 가능성을 평가하는 것이다. 둘째, 인공지능(AI) 기술을 활용한 혁신적 e-Discovery 솔루션의 도입 방안을 제시하여 한국형 e-Discovery 생태계 구축에 이바지하는 것이다.

이러한 연구는 다음과 같은 필요성에 기반한다. 우선, 현재 국내에서 e-Discovery 전문 서비스를 제공하는 기업이 1개 사에 불과한 상황에서 시장 활성화 방안의 체계적 연구가 필요하다. 또한 AI 기술의 급속한 발전으로 e-Discovery 분야에서도 기술적 혁신이 가속화되고 있어 [5], 한국이 이러한 글로벌 트렌드에 선제적으로 대응할 수 있는 전략이 요구된다.

본 연구의 시간적 범위는 2015년 한국형 e-Discovery 제도 도입 논의의 시작부터 2025년 현재까지 설정하였으며, 공간적 범위는 한국을 중심으로 하되 미국, 일본 등 주요 국가와의 비교 분석을 포함하였다. 연구 방법으로는 문헌 연구를 통한 이론적 고찰, 정부 정책 자료 및 시장 보고서 분석, 그리고 국내외 사례 연구를 병행하였다.

## II. Theoretical Background

### 2.1 Concept and Importance of e-Discovery

전자증거개시는 민사소송에서 전자적으로 저장된 정보(Electronically Stored Information, ESI)를 대상으로 하는 Discovery 절차를 의미한다[6]. 이는 전통적인 종이 문서 증거의 증거수집이 디지털 시대에 맞게 확장 발전한 형태로, 이메일, 문서파일, 데이터베이스, 소셜미디어 게시물, 클라우드 저장 데이터 등 모든 형태의 전자정보를 포괄한다[7].

e-Discovery의 중요성은 현대 비즈니스 환경의 디지털 화에서 비롯된다. 기업 의사소통의 대부분이 이메일과 메시지를 통해 이루어지고, 핵심 문서들이 전자 형태로 작성·보관되는 상황에서 전자증거는 종종 사건의 핵심을 결정하는 요소가 된다[8]. 특히 2002년 Zubulake v. UBS Warburg LLC 사건에서 보듯이, 전자증거의 고의적 삭제나 은폐는 소송 결과를 좌우하는 결정적 요인이 될 수 있다[9].

### 2.2 Global e-Discovery Market Trends

글로벌 e-Discovery 시장은 지속적인 성장세를 보인다. Market Research Future의 2024년 보고서에 따르면, e-Discovery 시장 규모는 2024년 127억 달러에서 2035년 275억 달러로 성장할 것으로 예상되며, 이는 연평균 성장률(CAGR) 7.23%에 해당한다[10].

시장 성장의 주요 요인으로는 첫째, 기업 데이터 양의 급증, 둘째, 규제 준수 요구사항 강화, 셋째, 국제 소송의 증가, 넷째, 클라우드 컴퓨팅과 모바일 기술의 확산 등이 꼽힌다[11]. 특히 COVID-19 팬데미 이후 원격근무가 확산되면서 기업 데이터의 분산도가 높아지고, 이에 따라 e-Discovery 서비스 수요가 크게 증가하였다[12].

지역별로는 북미가 전체 시장의 60% 이상을 차지하고 있으며, 아시아 태평양 지역이 가장 빠른 성장률을 보이고 있다[13]. 이는 아시아 기업들의 글로벌 진출 증가와 현지 규제 환경의 변화에 기인한다.

## III. Analysis of the Potential for e-Discovery Market Development in Korea

### 3.1 Legal Foundation and Policy Environment

#### 3.1.1 Limitations of the Current Legal Framework and Directions for Improvement

우리나라의 현행 민사소송 체계는 여전히 증거 문서 중심으로 설계되어 있어 전자증거의 특성을 충분히 반영하지 못하고 있다. 「민사소송 등에서의 전자문서 이용 등에 관한 법률」(민소전자문서법)[14]이 2010년 제정되어 전자문서의 증거력을 인정하고 있으나, 전자증거의 수집, 처리, 제출 등에 관한 세부적인 절차와 기준은 여전히 미비한 상태이다.

특히 현행 문서제출명령제도(민사소송법 제344조)는 증거 현재 형상으로 인해 실질적 문제 해결에 한계를 보이고 있다[15]. 기업 간 분쟁에서 핵심 증거가 상대방에게 집중된 경우, 피해자가 이에 접근하기 어려운 구조적 문제가 존재한다.

이러한 문제를 해결하기 위해 특허법에서는 2016년 개정을 통해 자료제출명령제도를 강화하였다. 특허법 제126조는 법원이 특허권 침해소송에서 침해 증명이나 손해액 산정에 필요한 자료 제출을 요구할 수 있도록 하고, 정당한 이유 없이 거부하면 원고에게 유리한 판단을 할 수 있도록 규정하고 있다[16]. 또한 대상 범위를 '서류'에서 '자료'로 확대하여 데이터, 전자문서 등도 명령 대상에 포함시켰다.

### 3.1.2 Government's Initiative to Introduce a Korean-Style Discovery System

2015년 대법원은 '사실심 충실화 마스터플랜'의 핵심 방안으로 '한국형 디스커버리' 제도 도입을 발표하였다[17]. 이는 증거 편재 현상 해소와 실체적 진실 발견을 통한 사법부 신뢰도 제고를 목적으로 한다.

2019년에는 더불어민주당 조응천 의원과 한국법조인협회가 공동 주최한 정책토론회에서 한국형 증거개시제도 도입의 필요성과 입법 방향이 구체적으로 논의되었다[18]. 이 과정에서 미국식의 광범위한 디스커버리가 아닌 독일의 '전문가 증거조사(뒤셀도르프 절차(Dusseldorfer Verfahren))', 영국의 '자료목록·자료교환' 등을 접목한 한국 실정에 맞는 제도 설계가 제안되었다.

2023년에는 중소벤처기업부가 「중소기업기술 보호 지원에 관한 법률」 개정을 통해 한국형 디스커버리 제도 도입을 위한 법적 근거를 마련하였다[19]. 특허청을 중심으로 특허소송에 우선 도입한 후 민사소송 전반으로 확대하는 단계적 접근 방식이 채택되었다.

### 3.1.3 Key Features of the System Implementation

한국형 디스커버리 제도의 주요 특징은 다음과 같다 [20].

- 저비용 고효율 설계: 미국식 대규모 디스커버리와 달리 비용 부담을 최소화하면서도 실효성을 확보하는 방향으로 설계
- 단계적 확대: 특허소송 → 상사소송 → 민사소송 순으로 점진적 확대
- 하이브리드 모델: 대륙법계의 장점을 살리면서 영미법계의 효과적 요소를 선별적으로 도입
- 기밀정보 보호: 영업비밀 유출 우려를 해소하기 위한 다양한 보호장치 마련

## 3.2 Market Demand and Opportunity Analysis

### 3.2.1 Increase in International Litigation Involving Domestic Companies

우리 기업의 해외 진출이 확대됨에 따라 국제 소송에 노출되는 빈도가 급격히 증가하고 있다. 특히 미국 국제무역위원회(ITC) 소송의 경우, 한국 기업이 연루된 사건이 2015년 이후 연평균 20% 이상 증가하고 있는 것으로 조사되었다[21].

삼성전자, LG전자, SK하이닉스 등 대기업은 물론 중견 기업들도 침해소송, 반독점 소송, 계약 분쟁 등에서 e-Discovery 절차에 직면하는 사례가 늘어나고 있다[22]. 그러나 대부분의 기업이 해외 업체에 의존하고 있어 핵심 기술정보의 유출 위험과 큰 비용 부담에 직면하고 있다.

### 3.2.2 Insufficient Supply of Domestic e-Discovery Services

현재 국내 e-Discovery 서비스 시장은 극도로 제한적인 공급 구조를 보인다. 전문적인 e-Discovery 서비스를 제공하는 국내 기업은 Intellectual Data 단일 업체에 불과한 실정이다[23]. 대형 로펌 역시 체계적인 e-Discovery 역량을 갖춘 곳은 소수에 그치며, 대부분 해외 업체에 의존하거나 자체적으로 제한적인 서비스만을 제공하고 있다.

이러한 공급 부족 현상은 여러 구조적 문제를 야기하고 있다. 첫째, 해외 업체 의존도가 높아짐에 따라 서비스 이용 비용이 증가하고 있다. 둘째, 민감한 기업 정보가 해외로 유출될 가능성이 있어 보안 위험이 가중되며, 셋째, 한국어 문서 처리에 대한 전문성이 부족하여 언어적 장벽이 발생하고 있으며, 마지막으로 국경 간 데이터 이전 과정에서 프로젝트 진행이 지연되는 문제가 빈번하게 발생하고 있다.

### 3.2.3 Estimation of Potential Market Size

국내 e-Discovery 시장은 초기 단계로 공식 시장의 통계가 부재한 상황이다. 대한상사중재원(2019)[24]에 따르

면 연간 440건의 중재가 이루어졌고, 이 중 일부가 e-Discovery 서비스가 필요한 것으로 예상되었다.

2024년 글로벌 시장 규모는 107~127억 달러 가량으로 현재의 국내 기업의 해외 진출 증가추세 등을 고려한다면 국내 시장 규모는 지속적으로 성장할 것으로 예상된다.

### 3.3 Assessment of Technological Readiness

#### 3.3.1 Current State of Domestic IT Infrastructure

한국은 세계 최고 수준의 IT 인프라를 보유하고 있어 e-Discovery 서비스 제공에 매우 유리한 환경을 갖추고 있다[25]. 구체적으로 살펴보면, 통신 인프라 측면에서 5G 통신망 구축률이 세계 1위를 기록하고 있으며, 초고속 인터넷 보급률은 95% 이상에 달한다. 클라우드 서비스 분야에서는 네이버 클라우드 플랫폼, 삼성SDS, LG CNS 등 대형 클라우드 서비스 업체들이 활발히 운영되고 있다. 또한 국내 주요 클라우드 사업자들이 운영하는 데이터 센터는 안정적인 서비스 제공 기반을 구축하고 있으며, K-사이버 시큐리티 기술력을 바탕으로 민감한 법률 데이터의 안전한 처리가 가능한 상황이다.

#### 3.3.2 AI and Big Data Technology Capabilities

한국의 AI 기술 역량은 e-Discovery 분야 진출에 중요한 경쟁 우위 요소로 작용하고 있다. 자연어 처리 분야에서는 네이버의 HyperCLOVA, 카카오브레인의 KoGPT 등 한국어 특화 대규모 언어 모델이 개발되어 운영되고 있다[26]. 머신러닝 기술은 삼성리서치, LG AI 연구원 등을 중심으로 계속해서 발전하고 있으며, 빅데이터 처리 분야에서도 국내 주요 기업들이 대용량 데이터 처리에 대한 풍부한 경험과 기술력을 축적해 왔다.

#### 3.3.3 Digitalization of the Legal Services Market

국내 법률 서비스 시장은 점진적이지만 꾸준한 디지털화 과정을 거치고 있다. 대형 로펌들을 중심으로 문서 관리 시스템과 케이스 관리 솔루션 도입이 확산하고 있으며, 법무 관련 IT 솔루션을 개발하는 리걸테크 스타트업의 수도 증가 추세를 보인다. 정부 차원에서도 법무부와 대법원이 추진하는 사법부 정보화 사업을 통해 디지털 인프라 확충이 이루어지고 있다.

### 3.4 Factors Promoting Market Activation

#### 3.4.1 Policy Factors

정부의 강력한 추진 의지는 시장 활성화의 핵심 동력으로 작용하고 있다. 대법원, 특허청, 중소벤처기업부 등 주요 정

부 기관이 일관된 정책을 추진하고 있으며, 단계적이고 체계적인 법령 개정을 통해 제도적 기반을 마련하고 있다. 또한 R&D 투자, 전문 인력 양성, 인프라 구축 등에 대한 정부 예산이 배정되어 실질적인 지원이 이루어지고 있다.

#### 3.4.2 Economic Factors

경제적 측면에서는 해외 업체 의존도 감소를 통한 서비스 비용 절감 효과가 기대되고 있다. 법무법인, IT 기업, 컨설팅 회사 등에는 새로운 사업 기회가 제공되고 있으며, 중장기적으로는 아시아 시장으로 한국형 e-Discovery 솔루션을 수출할 수 있는 잠재력도 존재한다.

#### 3.4.3 Technological Factors

기술적 측면에서는 기계학습, 자연어 처리 등 핵심 AI 기술이 상용화 단계에 진입했다는 점이 긍정적으로 작용하고 있다. 확장 가능하고 안전한 클라우드 서비스 기반이 확립되었으며, 국제 e-Discovery 표준이 확립됨에 따라 기술 도입에 따른 리스크도 감소하고 있다.

### 3.5 Factors Hindering Market Activation

#### 3.5.1 Institutional Barriers

시장 활성화를 저해하는 제도적 장벽으로는 먼저 법률 전문가들의 e-Discovery에 대한 이해도가 낮고 법조계의 보수적 성향이 강하다는 점을 들 수 있다. 로스쿨이나 사법연수원에서 관련 교육 과정이 부족하여 교육 체계가 제대로 구축되지 못한 상황이다[27]. 또한 국내에서 e-Discovery 관련 판례 축적이 미흡하여 법적 불확실성이 존재하는 것도 중요한 저해 요인으로 작용하고 있다.

#### 3.5.2 Corporate Concerns

기업 측면에서는 핵심 기술정보나 영업비밀의 외부 유출에 대한 정보 보안 우려가 크게 나타나고 있다. 초기 시스템 도입 비용과 운영비용에 대한 부담도 상당하며, e-Discovery 프로세스의 복잡성으로 인한 업무 부담 증가에 대한 우려도 존재한다.

#### 3.5.3 Technological Limitations

기술적 한계로는 법률 전문용어의 복잡성과 한국어의 특성으로 인한 기술적 도전이 존재한다. 기업별로 다른 문서 관리 체계로 인해 데이터 표준화가 어려운 실정이며, 무엇보다 법률과 IT 분야를 모두 이해하는 융합형 전문 인력이 부족한 상황이 시장 발전의 주요 걸림돌로 작용하고 있다.

## IV. AI-Driven e-Discovery Innovation Strategies

### 4.1 Current Status of AI-Based e-Discovery Technologies

#### 4.1.1 Development of Technology Assisted Review (TAR)

Technology Assisted Review는 e-Discovery 분야에서 가장 성숙한 AI 응용 기술로 평가된다[28]. TAR의 발전 과정은 크게 두 세대로 구분하여 살펴볼 수 있다.

초기 TAR 시스템인 TAR 1.0은 Simple Active Learning 방식을 채택하였으며, 변호사가 수동으로 선별한 "시드셋(seed set)"을 기반으로 기계학습 알고리즘을 훈련하는 방식으로 작동하였다[29]. 이러한 TAR 1.0 방식은 2012년 Da Silva Moore v. Publicis Groupe 사건에서 Andrew Peck 판사에 의해 최초로 법정에서 승인되었다는 점에서 의의가 있다[30]. 당시 Peck 판사는 판결에서 "컴퓨터 지원 검토는 대용량 데이터 사건에서 상당한 법무 비용을 절감할 수 있는 도구로 선견하게 고려되어야 한다"라고 판시하며 TAR 사용의 법적 정당성을 확립하였다[31].

2010년대 중반부터는 보다 진보된 형태인 TAR 2.0이 도입되기 시작하였다. Continuous Active Learning 방식을 기반으로 하는 TAR 2.0은 시드셋 없이도 검토자의 실시간 판단을 학습하여 자동으로 중요 문서의 우선순위를 설정하는 방식으로 작동한다[32]. 이는 기존 TAR 1.0 방식보다 더 높은 효율성과 정확도를 보여주며, 현재 대부분의 e-Discovery 플랫폼에서 표준 기술로 채택되고 있다.

TAR 기술의 도입으로 인한 성능 개선 효과는 매우 인상적인 것으로 나타났다[33]. 연구 결과에 따르면, TAR 기술을 활용하면 문서 검토 시간을 75%에서 90%까지 단축할 수 있으며, 관련 문서 탐지율은 95% 이상에 달하는 것으로 확인되었다. 또한 사람의 검토 방식과 비교하여 일관성이 15%에서 20% 향상되었으며, 대형 소송 1건당 평균 100만 달러에서 300만 달러의 비용을 절감할 수 있는 것으로 분석되었다.

#### 4.1.2 Application of Natural Language Processing (NLP) Technology

전통적인 키워드 검색의 한계를 극복하기 위해 개발된 개념 검색(Conceptual Search) 기술은 단순한 키워드 매칭을 넘어 문서의 의미와 맥락을 이해하여 검색을 수행한다[34]. 이 기술은 동의어, 유의어, 개념적 연관성을 인식하는 능력을 갖추고 있어, 기존 키워드 검색 방식 대비

30%에서 40% 높은 재현율을 달성하는 것으로 보고되고 있다[35].

감정 분석(Sentiment Analysis) 기술은 이메일이나 메시지의 감정 상태를 분석하여 갈등 상황이나 문제의 징후를 사전에 감지하는 데 활용된다[36]. 이 기술은 특히 기업 내부 커뮤니케이션에서 분쟁의 조짐이나 협상 과정의 변화를 추적하는 데 유용한 것으로 평가받고 있다.

글로벌 기업의 경우 여러 언어로 작성된 문서를 동시에 처리해야 하는 경우가 많다. 이러한 필요에 대응하여 최신 NLP 기술은 실시간 번역과 다국어 동시 분석 기능을 통해 복잡한 다국어 문서 처리 문제를 해결하고 있다[37].

#### 4.1.3 Innovative Application of Generative AI

2023년 이후 ChatGPT, Claude 등 생성형 AI의 등장으로 e-Discovery 분야에서도 새로운 혁신의 물결이 일어나고 있다[38]. 생성형 AI의 주요 적용 분야는 크게 세 가지로 구분할 수 있다.

첫째, 자동문서 요약 기능은 대량의 계약서, 보고서, 이메일 체인을 자동으로 요약하여 핵심 내용을 추출하는 것을 가능하게 한다[39]. Reed Smith 로펌이 실시한 테스트 결과에 따르면, 생성형 AI를 활용한 문서 요약의 정확도가 기존 방식 대비 25% 향상된 것으로 나타났다[40].

둘째, 법적 쟁점 식별 기능은 문서 내용을 분석하여 잠재적 법적 쟁점을 자동으로 식별하고 관련 법률 조항을 제시한다. 이는 변호사의 초기 사건 분석 시간을 크게 단축하는 효과를 가져온다.

셋째, 자동 보고서 생성 기능은 e-Discovery 프로세스의 단계별 결과를 종합하여 자동으로 분석 보고서를 생성한다. 이를 통해 클라이언트에 대한 보고 업무의 효율성을 크게 향상할 수 있다.

### 4.2 Development Strategies for Korean Language-Specific AI Solutions

#### 4.2.1 Unique Characteristics of Korean Natural Language Processing

한국어는 다른 언어들과 구별되는 고유한 특성이 있어 e-Discovery 시스템 개발 시 특별한 고려가 필요하다[41]. 한국어의 주요 특수성은 크게 두 가지로 설명할 수 있다.

먼저 교착어적 특성을 들 수 있다. 한국어는 어근에 다양한 어미와 조사가 결합하는 교착어로서, 하나의 단어가 수십 가지 형태로 변화할 수 있는 특징을 갖는다. 이러한 특성은 검색어 처리와 문서 분석 과정에서 상당한 복잡성을 증가시키는 요인으로 작용한다.

또한 한국어는 높은 문맥 의존성을 가지고 있어, 같은 단어라도 사용되는 문맥에 따라 서로 다른 의미가 있을 수 있다. 이러한 특성은 특히 법률 용어의 경우 더욱 두드러지게 나타나므로, 한국어 법률 문서 처리 시 각별한 주의가 필요하다.

#### 4.2.2 Development of Korean Legal NLP Models

한국어 법률 NLP 모델을 효과적으로 개발하기 위해서는 먼저 전문 용어 사전을 체계적으로 구축해야 한다[42]. 한국의 법률 전문용어, 판례 표현, 계약서 관용구 등을 수집하여 전문 용어 사전을 구축하는 과정에서는 대법원 판례 데이터베이스, 법제처 국가법령정보센터, 주요 로펌의 계약서 템플릿, 학술논문 및 법률 서적 등을 데이터 출처로 활용할 수 있다.

다음으로 도메인 특화 사전 훈련이 필요하다[43]. HyperCLOVA, KoGPT 등 기존에 개발된 한국어 대규모 언어 모델을 법률 도메인 데이터로 추가 훈련해 법률 특화 성능을 향상해야 한다. 이 과정에서는 일반 언어 모델을 법률 도메인으로 전이하는 Transfer Learning, 법률 문서를 이용한 세밀한 조정을 수행하는 Fine-tuning, 그리고 여러 법률 작업을 동시에 학습하는 Multi-task Learning과 같은 기법들을 활용할 수 있다.

또한 한국어 법률 NLP 모델의 성능을 객관적으로 평가하기 위한 체계적인 성능 평가 체계를 마련해야 한다. 이를 위해서는 벤치마크 데이터셋과 평가 지표를 개발해야 하며, 평가 체계에는 문서 분류 정확도, 핵심 정보 추출 정밀도, 관련성 판단 일치도, 처리 속도 및 확장성 등의 요소들이 포함되어야 한다.

#### 4.2.3 Ensuring Legal Verifiability of AI Models

법정에서 AI의 판단 결과를 증거로 사용하기 위해서는 그 판단 과정을 명확히 설명할 수 있어야 하므로, 설명할 수 있는 AI(Explainable AI, XAI) 기술의 개발이 필수적이다[44]. 이를 위해서는 AI가 어떤 부분에 주목하여 판단했는지를 시각화하는 Attention Mechanism, 판단에 큰 영향을 미친 특성들을 식별하는 Feature Importance, 그리고 의사결정 과정을 단계적으로 설명하는 Decision Tree와 같은 기술들이 개발되어야 한다.

더불어 AI 시스템의 일관성과 신뢰성을 보장하기 위한 품질 보증 체계가 필요하다. 이러한 품질 관리 체계에는 정기적 성능 모니터링, A/B 테스트를 통한 지속적 개선, 그리고 전문가 검증을 통한 품질 확인 절차가 포함되어야 한다.

### 4.3 Practical Application Cases and Results

#### 4.3.1 International Success Cases

세계 최대 e-Discovery 플랫폼 중 하나인 Relativity는 AI 기술을 적극적으로 활용하여 상당한 성과를 달성하였다[45]. Relativity의 AI 기반 플랫폼은 문서 검토 속도를 10배 향상했으며, 관련 문서 탐지율 98%를 달성하였고, 전체 e-Discovery 비용을 60% 절감하는 성과를 거두었다.

Microsoft는 자사의 클라우드 플랫폼에 AI 기반 e-Discovery 기능을 통합하여 Purview eDiscovery 서비스를 제공하고 있다[46]. 2024년 3월 기준으로 Microsoft는 predictive coding 기능을 포함한 고도화된 AI 서비스를 제공하고 있으며, 이 서비스는 Office 365 데이터와의 원활한 연동, 실시간 데이터 분석 및 필터링, 자동화된 보고서 생성 등의 특징을 보인다.

Reed Smith 로펌은 자회사인 Gravity Stack을 통해 AI 기반 케이스 준비 도구를 개발하여 실제 법률 업무에 적용하고 있다[47]. 6개월간의 테스트 동안, 이 도구는 기존 예측 코딩 대비 15% 높은 정확도를 보였으며, 생성형 AI를 결합함으로써 전체적인 성능이 향상되었고, 클라이언트 만족도가 25% 증가하는 결과를 얻었다.

### 4.4 Design of AI-Based Korean e-Discovery System

#### 4.4.1 System Architecture

한국형 e-Discovery 시스템은 대용량 데이터를 효율적으로 처리하기 위해 클라우드 기반의 분산 처리 아키텍처를 채택해야 한다. 이러한 아키텍처의 주요 구성 요소는 데이터 수집 및 저장을 담당하는 단계, AI 분석 처리를 수행하는 단계, 사용자와의 상호작용을 담당하는 사용자 인터페이스 단계, 그리고 보안 및 접근 제어를 관리하는 단계로 구성된다.

시스템 설계 시에는 각 기능을 독립적인 마이크로서비스로 구현하는 마이크로서비스 기반 설계 방식을 채택하여 확장성과 유지보수성을 향상해야 한다. 주요 마이크로서비스로는 데이터 수집 서비스, 전처리 서비스, AI 분석 서비스, 검토 지원 서비스, 보고서 생성 서비스 등이 포함된다.

#### 4.4.2 Core AI Functions

시스템의 핵심 AI 기능으로는 먼저 지능형 문서 분류 기능이 필요하다. 이는 사건 유형별로 문서를 자동으로 카테고리화하고, 중요도에 따라 우선순위를 설정하며, 특권 문서 및 민감정보를 식별하는 역할을 수행한다.

다음으로 내용 기반 검색 및 분석 기능이 구현되어야 한다. 이 기능은 의미 기반 검색을 가능하게 하고, 문서 간 개념적 연관성을 분석하며, 타임라인을 재구성하고 사건의 인과관계를 추적하는 역할을 담당한다.

또한 자동화된 품질 관리 기능도 필수적이다. 이 기능은 검토의 일관성을 지속적으로 모니터링하고, 누락 가능성이 있는 부분에 대해 알림을 제공하며, 품질 점수를 자동으로 산출하여 전체적인 검토 품질을 보장한다.

#### 4.4.3 User Experience Optimization

시스템의 사용자 인터페이스는 변호사 친화적으로 설계되어야 한다. 이를 위해서는 법률 실무진의 실제 업무 플로우를 충실히 반영해야 하며, 직관적인 대시보드와 효과적인 데이터 시각화를 제공해야 하고, 모바일 환경에서도 원활하게 작동할 수 있도록 지원해야 한다.

또한 효율적인 협업을 위한 도구들이 통합되어야 한다. 이러한 협업 기능에는 팀 단위로 문서를 검토하고 태깅할 수 있는 기능, 팀원 간 실시간으로 의견을 공유하고 토론할 수 있는 기능, 그리고 문서의 버전을 관리하고 변경 내용을 추적할 수 있는 기능이 포함되어야 한다.

## V. Conclusion and Policy Implications

### 5.1 Summary of Research Findings

본 연구는 대한민국의 현행 법률 체계하에서 e-Discovery 시장의 활성화 가능성을 종합적으로 분석하고, AI 기술을 활용한 혁신적 솔루션 도입 방안을 제시하였다. 주요 연구 결과는 다음과 같다.

첫째, e-Discovery 시장의 활성화 가능성은 매우 높은 것으로 평가되었다. 정부의 한국형 디스커버리 제도 도입 의지가 명확하고, 국내 기업의 해외 소송이 지속적으로 증가하고 있으며, 우수한 IT 인프라를 갖추고 있다는 점에서 시장 성장의 기반이 충분히 마련되어 있다. 특히 주목할 만한 점은 현재 전문 서비스 업체가 1개 사에 불과한 상황에서 잠재 시장 규모가 2035년까지 2,000~3,000억 원에 달할 것으로 추정된다는 것이다. 이는 시장 선점의 기회가 크다는 것을 시사한다.

둘째, AI 기술의 적용이 e-Discovery 프로세스에 혁신적 효과를 가져올 수 있음을 확인하였다. TAR(Technology Assisted Review), NLP(자연어 처리), 생성형 AI 등 최신 AI 기술을 활용할 경우, 전통적인 문서 검토 방식 대비 75~90%의 시간 단축과 30~40% 높은 정확도를 달성할 수

있는 것으로 나타났다. 특히 한국어 특화 모델 개발을 통해 국내 시장에서의 경쟁 우위를 확보할 수 있을 것으로 판단되며, 이는 향후 아시아 시장으로의 확장 가능성도 내포하고 있다.

## 5.2 Policy Recommendations

### 5.2.1 Legal and Institutional Reform

한국형 e-Discovery 시장의 활성화를 위해서는 먼저 법 제도의 정비가 선행되어야 한다. 현행 민소전자문서법은 전자증거개시 절차를 포괄적으로 규율하기에는 한계가 있으므로, 전자증거개시에 특화된 별도의 한국형 e-Discovery 법률을 제정할 필요가 있다. 이 법률은 전자 증거의 수집·처리·보존 절차, AI 시스템 활용에 대한 법적 근거, 데이터 보안 및 개인정보보호 기준, 국경 간 데이터 이전 규정 등을 명확히 규정해야 한다.

또한 법원과 법조계가 AI 기반 e-Discovery 도구를 안전하고 효과적으로 활용할 수 있도록 구체적인 가이드라인을 제정해야 한다. 이 가이드라인에는 AI 시스템의 검증 기준, 알고리즘의 투명성 요구사항, 품질 보증 절차, 오류 발생 시 대응 방안 등이 포함되어야 하며, 기술의 발전 속도를 고려하여 주기적으로 개정될 수 있는 유연성을 확보해야 한다.

### 5.2.2 Industry Development Policy

e-Discovery 산업의 경쟁력 확보를 위해서는 정부 차원의 적극적인 R&D 투자 확대가 필요하다. 특히 한국어 법률 NLP 기술 개발, 설명할 수 있는 AI 기술 연구, 데이터 보안 및 프라이버시 보호 기술, 국제 표준 연구 및 개발 참여 등 핵심 기술 분야에 대한 집중적인 투자가 요구된다. 이러한 기술 개발은 국내 시장 확보뿐만 아니라 글로벌 시장 진출의 기반이 될 것이다.

아울러 리걸테크 분야의 창업을 촉진하고 중소기업의 기술 개발을 지원하기 위한 종합적인 지원 정책이 마련되어야 한다. 창업 자금 지원 및 세제 혜택, 기술 개발 자금 저리 융자, 정부 조달 시장 진입 우대, 해외 진출 지원 프로그램 등을 통해 생태계 전반의 성장을 도모해야 한다. 특히 초기 스타트업이 시장 진입 장벽을 낮출 수 있도록 규제 샌드박스 제도를 적극 활용하는 것도 고려할 필요가 있다.

### 5.2.3 Establishment of Expert Training System

e-Discovery 산업의 지속 가능한 발전을 위해서는 법학과 공학을 아우르는 융합형 인재 양성이 필수적이다. 이

를 위해 로스쿨에 e-Discovery 관련 교육 과정을 신설하고, 공학계열에서는 법률 정보학 교육을 강화해야 한다. 또한 산학협력을 통한 실무 중심 교육과 해외 연수 및 교류 프로그램을 확대하여 글로벌 수준의 전문가를 양성해야 한다.

더불어 e-Discovery 분야의 전문성을 객관적으로 인정하고 서비스 품질을 보장하기 위한 전문가 인증 제도의 도입이 필요하다. 국가 차원의 전문가 자격 인증 체계를 수립하고, 국제 인증과의 상호 인정 체계를 마련하며, 지속적 교육을 의무화하고, 윤리 강령 및 품질 기준을 확립함으로써 전문가 집단의 역량과 신뢰성을 제고할 수 있을 것이다.

### 5.3 Limitations and Future Research Directions

본 연구는 문헌 연구와 사례 분석에 주로 의존했다는 방법론적 한계를 지니고 있다. 이론적 분석과 해외 사례 검토를 통해 한국 시장의 가능성과 방향성을 제시하였으나, 실제 시스템 구현과 운영 데이터에 기반한 실증적 검증은 이루어지지 못하였다. 따라서 향후 연구에서는 실제 파일럿 시스템을 구축하고 실증 실험을 통해 제안된 모델의 성능과 효과성을 검증하는 작업이 필요하다.

아래 그림과 같이 한국형 e-Discovery 생태계 프레임워크를 바탕으로 구체적이고 깊이 있는 후속 연구가 필요하다.

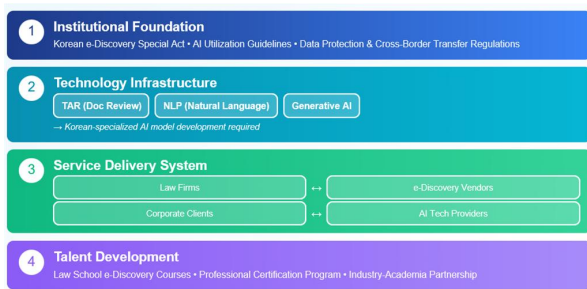


Fig. 1. Korean e-Discovery Ecosystem Framework

또한 본 연구는 주로 민사소송 분야의 e-Discovery를 중심으로 분석하였으나, 형사, 행정, 지식재산권 등 다양한 법률 분야로의 확장 가능성에 관한 추가 연구가 요구된다. 각 법률 분야별 특성과 요구사항이 다르므로, 분야별 맞춤형 솔루션 개발을 위한 심층 연구가 필요할 것이다. 나아가 국제 소송에서 활용을 고려할 때, 국제적 호환성과 상호운용성에 관한 연구도 중요한 과제로 남아 있다.

특히, AI 기술의 급속한 발전을 고려할 때, 지속적인 모니터링과 업데이트가 필요하며, 새로운 기술의 법적·윤리

적 쟁점에 관한 선제적 연구가 중요할 것으로 판단된다. 생성형 AI, 설명할 수 있는 AI, 연합학습 등 신기술의 등장은 e-Discovery 분야에 새로운 기회와 도전을 동시에 제공하고 있으며, 이러한 기술적 변화에 대응하기 위한 법제도적·윤리적 연구가 병행되어야 할 것이다.

## REFERENCES

- [1] IDC, "The Digitization of the World: From Edge to Core", 2018.
- [2] Federal Rules of Civil Procedure, Rule 26, 34, 37 (2006 amendments).
- [3] Market Research Future, "E-Discovery Market Size, Share & Trends Analysis Report", 2024; Verified Market Research, "Global E-Discovery Market Research Report", 2024.
- [4] Supreme Court of Korea, Master Plan for Enhancement of Fact-Finding Trials, 2015, pp. 67-89.
- [5] The Transformative Impact of AI on eDiscovery: Trends and Insights, EDRM, 2024.
- [6] The Sedona Conference, "The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production" (3d ed. 2017), p. 1.
- [7] Fed. R. Civ. P. 34(a)(1)(A); Advisory Committee Notes to 2006 Amendments.
- [8] George L. Paul & Jason R. Baron, "Information Inflation: Can the Legal System Adapt?", 13 Rich. J.L. & Tech. 10, 15-20 (2007).
- [9] Zubulake v. UBS Warburg LLC, 217 F.R.D. 309 (S.D.N.Y. 2003) et seq.
- [10] Market Research Future, "E-Discovery Market Research Report", 2024, pp. 23-45.
- [11] Verified Market Research, "In-Depth Industry Outlook: E-Discovery Market Size, Forecast", 2024.
- [12] Levi Swank, "Remote Work and E-Discovery: New Challenges in the COVID-19 Era", 22 Sedona Conf. J. 45, 50-55 (2021).
- [13] ComplexDiscovery, "2024-2029 eDiscovery Market Size Mashup", 2024.
- [14] Act on the Use of Electronic Documents in Civil Procedure, etc., Article 3 (Effect of Electronic Documents).
- [15] Kim, Jeong-han, "Measures to Improve the System for Efficient Discovery and Collection of Evidence," Judicial Policy Research Institute, 2015, pp. 11-45.
- [16] Patent Act, Article 126 (Order for Submission of Materials); Lee, Na-ra & Shim, Ji-eun, "A Study on Strengthening the Effectiveness of Evidence Collection Systems in Patent Infringement Litigation," Korea Institute of Intellectual Property, 2017.
- [17] Supreme Court of Korea, supra note 4, 2015.
- [18] Office of Representative Jo Eung-cheon, "Proceedings of the 2019 Policy Forum for the Introduction of Korean-style Discovery

- System," 2019.
- [19] Ministry of SMEs and Startups, "Press Release on the Amendment of the Act on the Protection and Support of SME Technology," 2023.
- [20] "Korean-style Discovery Could Trigger Patent Litigation Crisis," Herald Economy, March 12, 2020.
- [21] Joo, Sang-don, "Korea Faces Intensifying 'NPE Patent Offensive,'" IP TREND Annual Report, 2020.
- [22] Han, Kyung-woo, "Korean Companies Intensify Offensive Against Disadvantages in US ITC Litigation," Maeil Business Newspaper, March 15, 2020.
- [23] E-Discovery, Wikipedia, last modified May 17, 2024.
- [24] The Academy of Korean Studies, "Korean Commercial Arbitration Board," (<https://encykorea.aks.ac.kr/Article/E0015084>)Encyclopedia of Korean Culture, 2019.
- [25] National Information Society Agency, 2023 Digital Divide Survey, 2023.
- [26] Naver Clova, "HyperCLOVA X: Hyperscale AI Technical Report," 2023; Kakao Brain, "KoGPT Technical Documentation," 2022.
- [27] None of the domestic law schools in Korea offer regular courses on e-Discovery, and neither the Korean Bar Association nor related organizations operate systematic educational programs on the subject.
- [28] Maura R. Grossman & Gordon V. Cormack, "Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review", 17 Rich. J.L. & Tech. 11 (2011).
- [29] Predictive Coding Technologies and Protocols: Overview and Survey, ComplexDiscovery, 2018.
- [30] *Da Silva Moore v. Publicis Groupe*, 287 F.R.D. 182 (S.D.N.Y. 2012).
- [31] *Id.* at 193.
- [32] Maura R. Grossman & Gordon V. Cormack, "Evaluation of Machine-Learning Protocols for Technology-Assisted Review in Electronic Discovery", 43 U. Balt. L. Rev. 1 (2013).
- [33] Exterro, "Chapter 7B: Predictive Coding (Technology Assisted Review) & Artificial Intelligence", 2024.
- [34] John Tredennick, "The Rise of Analytics in E-Discovery: Practical Applications", 15 Sedona Conf. J. 165, 170-175 (2014).
- [35] Casepoint, "Complete Guide for Predictive Coding and How it Can be Used in eDiscovery", 2022.
- [36] Daniel B. Garrie & Yoav M. Griver, "Shedding Light on Metadata: The Hidden Minefield of Electronic Discovery", 13 Rich. J.L. & Tech. 7 (2007).
- [37] Angela Agrusa & Robert Brownstone, "Cross-Border E-Discovery: Challenges and Solutions in International Litigation", 19 Sedona Conf. J. 45 (2018).
- [38] Reed Smith, "AI explained: AI and e-discovery", Perspectives, 2024.
- [39] Logikcull, "Predictive Coding in eDiscovery - A Complete Guide", 2024.
- [40] Reed Smith, *supra* note 37, 2024.
- [41] Kwon, Yang-sun, "A Study on the E-Discovery System in Criminal Procedure," Journal of Law, Vol. 16, Korean Law Association, 2016, p. 425.
- [42] Korea Legislation Research Institute, "Research on the Development of AI-based Legal Information Processing Technology," 2023.
- [43] Naver AI Lab, "Research on the Development of Legal Domain-Specific Language Models," 2023.
- [44] Craig Ball, "The Black Box Problem: Making AI Decisions Transparent in Legal Settings", 21 Sedona Conf. J. 89 (2020).
- [45] Relativity, "State of the Industry Report: E-Discovery Trends and Predictions 2024", 2024.
- [46] Microsoft Learn, "Predictive coding in eDiscovery (Premium) - Quick start", 2024.
- [47] Reed Smith, *supra* note 37, 2024.

## Authors



Yongho Kim received the B.S. degree in Law from Dong-a University in 2007, M.S. and Ph.D. degrees in Information Security from Pukyong National University in 2011 and 2019, In addition, he worked in

cybercrime investigation and digital forensics for 24 years at the Korea National Police Agency. Dr. Kim joined Kyungil University's Professor of Police Science Department in 2022. He is interested in criminal investigations, digital forensics, and cybercrime investigations.