International Journal of Computing Sciences Research (ISSN print: 2546-0552; ISSN online: 2546-115X)
Vol. 10, pp. 4292-4330
doi: 10.25147/ijcsr.v10i0.826
https://stepacademic.net

Long Paper

Evolution of Artificial Intelligence within LegalTech-A Literature Review

Maarten Looijenga
University of Twente, The Netherlands
maartenlooijenga@live.nl

Faiza Allah Bukhsh
University of Twente, The Netherlands
f.a.bukhsh@utwente.nl
ORCID: 0000-0001-5978-2754

Jeewanie Jayasinghe Arachchige
Vrije University, The Netherlands
j.jayasinghearachchige@vu.nl
ORCID: 0000-0001-8619-6523
(corresponding author)

Frances Borkent
University of Twente, The Netherlands
f.borkent@utwente.nl

Maya Daneva
University of Twente, The Netherlands
m.daneva@utwente.nl

Date received: January 13, 2026
Date received in revised form: May 11, 2026
Date accepted: June 9, 2026

Recommended citation:

> Looijenga, M., Bukhsh Allah, F., Jayasinghe Arachchige, J., Borkent, F., & Daneva, M. (2026). Evolution of Artificial Intelligence within LegalTech - A literature review. *International Journal of Computing Sciences Research*, *10*, 4292-4330. https://doi.org/10.25147/ijcsr.v10i0.826

Purpose – This research aims to provide an overview of the current state of research regarding the possible future use of AI in the legal domain and its impact.

Method – This research follows the guidelines proposed by Kitchenham and Charters (2007) for a systematic literature review.
Results – This study presents the current state of research on the use of artificial intelligence in LegalTech. This research shows that AI applications in the legal domain have existed for many years, but interest has grown sharply since around 2016. Across the literature, AI is used in a wide range of legal areas, and most papers focus on a single domain. A significant amount of research examines the justice systems. This study reveals that two major AI technologies are dominating the legal domain, namely, machine learning and natural language processing.

Conclusion – This review examines 91 studies and provides an overview of how AI is currently used in the legal domain. Most studies focus on a single application domain. The reported benefits of using AI include improved efficiency and accuracy, reduced costs, and the ability to process large datasets without fatigue or human bias. However, the review also identifies major challenges such as liability, privacy risks, system complexity, and high implementation costs.

Recommendations – AI can be used to enhance the legal process and activities, but cannot yet replace human expertise.

Research Implications – Results show that current research highly focuses on specific legal domains, for example, lawyers, law firms, and judicial decision-making. Future studies can be open to other areas like legal education and public sector legal services.

Practical Implications – The findings demonstrate that legal practitioners and law firms can benefit by adopting AI to enhance efficiency, accuracy, and decision-making.

Keywords – artificial intelligence, LegalTech, machine learning, natural language processing, literature review

INTRODUCTION

AI has experienced an exponential rise in recent years. In 2016, the commercial and private use of AI was marginal; adaptation by organisations doubled between 2017 and 2022 (Chui et al., 2022). A breakout was achieved by introducing a publicly available generative AI tool called "ChatGPT" in November 2022. This tool could chat with users and generate text based on users' text prompts (Deng, 2022; Milmo, 2023). Through this tool, a massive interest in AI arose, and it gathered a user base of 100 million within two months after launch (Milmo, 2023). One year later, McKinsey declared 2023 as the breakout year of generative AI, based on the exponential growth of AI tools (Chui et al., 2023). One-third of their respondents stated that their organisations use generative AI in at least one business process (Chui et al., 2023).

Discussions about AI adaptation in the legal industry increased significantly following the exponential growth. Implementing AI in this sector is met with both enthusiasm for its possibilities and a heightened awareness of the potential risks. The integration of AI, while linked to an increased chance of data threats, leakage of company information (Agrawal et al., 2022; Beebeejaun & Gunputh, 2023; Hilt, 2017), and elevated risk of cyberattacks (Beebeejaun & Gunputh, 2023), also promises to make legal workers more efficient (Alarie et al., 2018; Francia et al., 2022; Baser & Saini, 2023; Csányi et al., 2022; Gravett, 2020; Sil et al., 2019; Kronblad et al., 2023; Thornton 2021; Vaciago 2019; Soukupová, 2021) and deepen and broaden their areas of expertise (Alarie et al., 2016; Gravett, 2020).

While the usage of AI in the legal domain is considered to have much potential, adoption is in its early stages, with only a few implementations made (Thomson Reuters Institute, 2023). Ready-to-use AI solutions are still scarce (Zhang et al., 2020). This research aims to provide an overview of the current state of research regarding the possible future use of AI in the legal domain and its impact. In order to achieve this goal, the study has formulated research questions. The primary research question is: "What is the current state of research on the use of artificial intelligence in the field of LegalTech?" A literature review has been conducted to gain a deep understanding of the current state and the scope of research. With the literature review, current available knowledge is summarised, with the aim of guiding future research in the use of AI in the legal domain.

This paper is structured as follows: The next section provides an overview of the research design and methodology, detailing how the study was conducted. This section also presents the research questions that guided this study. The following sections present the results for each research question in dedicated sections. Finally, the paper concludes with a discussion chapter summarising the findings and their implications. implications.

RESEARCH DESIGN AND METHODOLOGY

The following section outlines the methodology used for our study. We conducted a systematic literature review following the guidelines proposed by Kitchenham and Charters (2007). We selected Kitchenham's approach as it allows not only systematic literature reviews but also reviews with broad-reach questions. And it is well-suited for identifying the research gaps and providing a structured understanding of the field.

Initially, we presented our research questions, followed by our search strategy, which includes details about the databases we searched, the search query used, and the criteria for inclusion and exclusion. Finally, we discussed the procedure for the data extraction process.

Research Questions

This literature review aims to provide a thorough summary of the current state of research on the use of artificial intelligence in LegalTech. To achieve the stated goal in the introduction, the following five sub-questions have been formulated:

1. What is the current publication trend of research on the use of artificial intelligence in the field of LegalTech?
2. In which legal domains is AI most widely adopted?
3. What are the primary applications of AI in LegalTech?
4. Which AI technologies are used in LegalTech applications?
5. What are the reported benefits, challenges, and limitations of AI adoption in LegalTech?

Search Strategy

This section discusses the SLR's search strategy. We selected Scopus as our digital library, which includes publications from peer-reviewed journals and conference proceedings. After determining the search query and criteria, we conducted additional searches in other digital libraries, namely Web of Science, ScienceDirect, ACM, IEEE Xplore, JSTOR, and Legal Intelligence. This extra search was executed to include peer-reviewed literature that may not have been covered in Scopus.

Keywords & Search Query

The process of constructing the search string involved multiple iterations, with the aim of creating a comprehensive string that excluded irrelevant results while including all pertinent material.

The most common keywords used in research about using AI in the legal sector are identified. The keyword "legaltech" is often used, defined by Salmerón-Manzano (2021, page numbers) as the technology that supports activities undertaken by legal professionals. Legaltech generally encompasses the application of technology to provide legal services (Salmerón-Manzano 2021). In this paper, the definition of legaltech by Salmerón-Manzano (2021) is followed. Other variations of this keyword used in literature to denote legaltech are legal technology and legal tech (as separate words) (Salmerón-Manzano 2021). Additionally, the keyword "lawtech" is frequently used. According to Salmerón-Manzano (2021), there is no clear distinction between these two terms in the literature, but they noted a preference for using "legaltech" in the USA and "lawtech" in the UK. To comprehensively search relevant literature on the topic, both "legaltech" and "lawtech" will be used in our search queries.

These terms of legal technology were combined with terms for Artificial Intelligence. Both the fully written word artificial intelligence and the abbreviation " AI were used, which led to the following initial search string:

> ("legal tech*" OR "legaltech*" OR "law tech*" OR "lawtech*") AND ("artificial intelligence" OR ai)

After conducting some trial searches to refine the search query, it was observed that combining the terms "legaltech" and "lawtech" with AI yielded limited results (Alarie et al., 2018; Dale, 2019; Hunter, 2020; Surden, 2022). It was noted that many relevant papers did not explicitly use "legaltech" or "legal technology" in their titles or abstracts. In some cases, the concept of legaltech was mentioned but not the term itself (Alarie et al., 2018; Dale, 2019; Hunter, 2020; Surden, 2022). These papers often used distinctive terms like "legal profession" and "legal practice". Therefore, these terms are added to the search query.

It was also noted that AI-related technologies, like Machine Learning (ML) (Delgado et al., 2022; Metsker et al., 2021) and Natural Language Processing (NLP) (Dale, 2019; Moreno-Schneider et al., 2020; Nazarenko & Wyner, 2017), were used without explicitly mentioning AI in the abstract, keywords, or title. Alongside AI, NLP and ML were added as supplementary terms in the search query to cover these papers. These terms were used in both their total and abbreviated forms. The final search string, which is used for data extraction, is as follows:

> ("legal tech*" OR "legaltech*" OR "law tech*" OR "lawtech*" OR "legal profession" OR "legal practise") AND ("artificial intelligence" OR ai OR "natural language processing" OR nlp OR "machine learning" OR ml)

Wildcard characters have been used in this search string. to make it easier to search for particular subjects.

Inclusion Criteria

Inclusion criteria are defined to only take into account studies that are relevant to the review's objectives. The inclusion criteria are:

1. The paper is directly relevant to the scope of our review, focusing exclusively on the application of artificial intelligence to enhance legal services.
2. The paper does discuss legal technology and artificial intelligence as its main topic.
3. The paper is published in a peer-reviewed journal or conference proceedings.
4. The paper is written in English.
5. The paper is available for download.

The inclusion criteria have been applied via automatic filtering options within the various search engines. Then, all remaining literature is manually analysed to apply the

first two inclusion criteria. The selection process only looks at a paper's title, keywords, and abstract, ignoring the remainder of its text.

Data Execution Process

This section provides a detailed overview of our process of gathering and excluding studies. Figure 1 gives a detailed overview of our execution of the data extraction process. The final study was carried out in October 2023. In the initial step, which was conducted on Scopus and other libraries, 380 papers were found. Before, multiple rounds of testing on Scopus were performed to identify the most effective search query and inclusion criteria. This process allowed us to fine-tune search parameters to ensure the results aligned with our expectations. Following this, we applied the inclusion criteria, excluding 38 non-English papers and 78 non-peer-reviewed papers, resulting in 264 papers.

Papers from various libraries were combined, duplicate papers were removed, and any results without a full-text version available were removed. During this process, it was discovered that four papers were not written entirely in English, so these papers were also excluded, leaving a total of 148 papers. The remaining papers have been subject to a qualitative content analysis. This analysis checked all titles, abstracts, and keywords to see if the papers are directly relevant to the scope of the SLR and if their main topic discusses LegalTech and AI. After this last selection, 91 papers were retained. Two significant categories of papers have been excluded from further review, although they were notably present in the subset of papers selected for the review. First, we excluded papers that addressed the education of law students. Although these papers examined how IT might change the way legal professionals practice, they were excluded from our study because they were deemed not directly relevant. Papers examining the ethics of using AI systems in legal practice have also been excluded. These papers do not offer any new information regarding the use of artificial intelligence in legal technology.

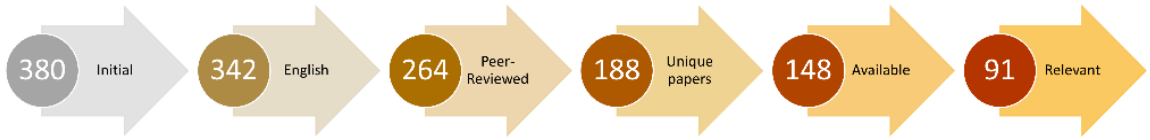

Figure 1. Data Execution Process

RESULTS

In this section, we present the results of the SLR. Each subsection focuses on addressing one of the research questions.

Current Publication Trend

This section explains the publication trend in research on artificial intelligence in LegalTech. It answers the research question: "What is the current publication trend of research on the use of artificial intelligence in the field of LegalTech?"

Figure 2 illustrates the distribution of papers according to their publication year. A remarkable rise in publications can be identified beginning in 2016. This rise highlights a growing interest in implementing artificial intelligence in the legal sector. Before 2016, researchers did not often mention artificial intelligence in research on legal technology. However, Figure 2 demonstrates a notable early interest in applying artificial intelligence within the legal sector.

Early Publications

This paragraph discusses some early publications. Not all early publications have been listed, but a selection has been made that was deemed interesting and relevant to the review. Leith (1988) authored the oldest paper identified in this literature review. Published in 1988, Leith (1988) wrote about the use of AI and computing within the legal sector. Leith believed computers could achieve higher accuracy and replace human work when legal tasks involved tracing consequences within laws. However, they noted there was no solid proof that computers could detect contradictions in law in every case.

In 1992, Ashley (1992) discussed using Case-Based Reasoning (CBR) in the legal sector. CBR employs methods for generalising cases to support indexing and relevance assessment. It evidences two basic inference methods: constraining search by tracing a solution from a past case or evaluating a case by comparing it to past cases. Artificial intelligence is mentioned to automate this generalisation process. Ashley already pushed in 1992 to actively pursue CBR research and development to solve problems efficiently and improve legal document assembly programs and data retrieval systems.

Around 1999, knowledge management was rising due to electronic systems. Oskamp et al. (1999) expected that the rise of knowledge management would have much to offer to the legal sector and lead to new methods enabling lawyers to detect their need for information and knowledge.

In 2002, Roberge et al. (2002) discussed the emergence of electronic databases and raised important questions regarding access to previously inaccessible information, foreseeing a transformative shift in legal practice. They anticipated that discovering, packaging, and disseminating information would elevate standards within the legal sector. The increased accessibility of information began to influence the established norms of the profession. Before the rise of electronic databases, information was primarily gained through experience, which was not uniformly accessible to all lawyers. Veteran attorneys had to rely on their expertise within the legal system to craft case strategies. Cases that

previously demanded expensive data collection and analysis would now become more accessible.

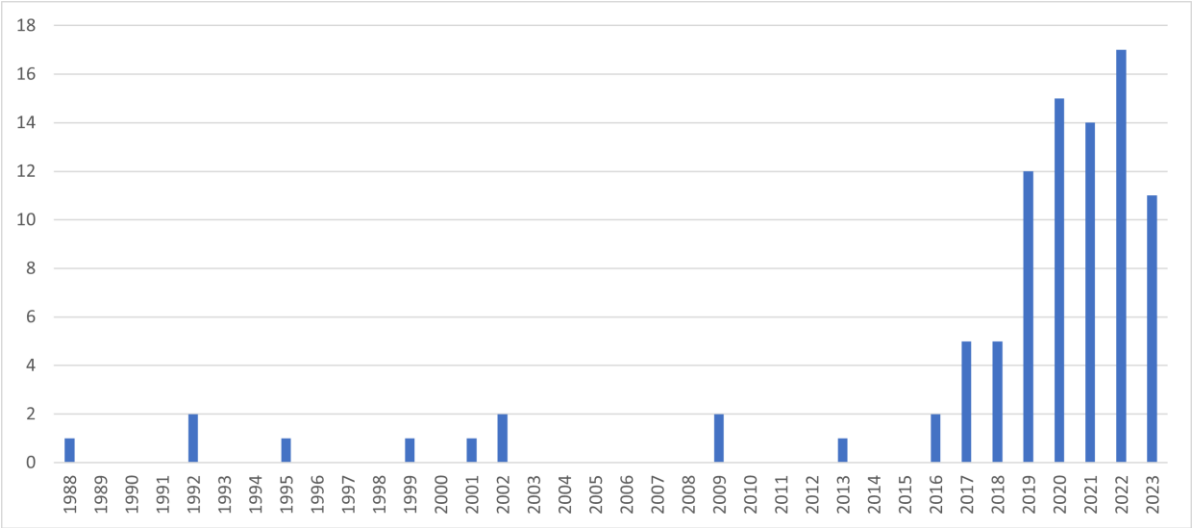

Figure 2. Number of publications per year

Publication Types

Remarkable is the higher number of journal articles compared to the number of conference proceedings. Our collection of 91 papers is 55% to 45%, divided between journal articles (50) and conference proceedings (41). The distribution is shown in Figure 3. Conference proceedings contain research presented at academic conferences, while journal articles contain peer-reviewed research published in academic journals. Conference proceedings often offer more current insights that might be in the early stages of development. At the same time, journal articles undergo extensive peer review, resulting in longer, more detailed publications that take longer to publish but often present more thoroughly validated research outcomes.

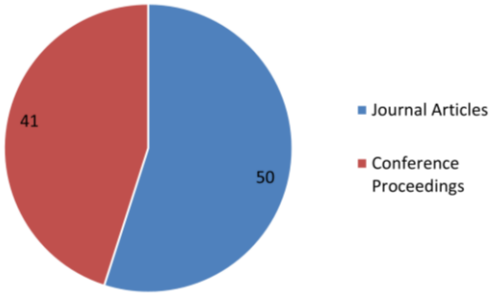

Figure 3. Types of publications of research papers

The distribution and higher number of journal articles compared to conference proceedings are remarkable. Given the recent rise in publications, more conference proceedings are expected. However, the more significant number of journal articles

suggests a deliberate focus on peer-reviewed knowledge within the field, indicating an emphasis on prioritising research quality over speed of publication.

Research Classification

After analysing the papers' full texts, we define two main categories. For all papers with fundamental concepts and definitions, we will call them "descriptive papers", explaining possible benefits that may arise from the use of legal tech in the legal industry. These articles elaborate on the potential advantages, risks, and benefits of artificial intelligence for legal corporations. Furthermore, some papers explore the possibility of a disruptive change in the legal industry caused by the rise of artificial intelligence, estimating notable adaptations. This particular category of papers plays an essential part in our research as it facilitates the identification of potential advantages, obstacles, and constraints related to implementing artificial intelligence in the legal domain.

Papers in the second category are papers with a validation cycle, evaluations, or case studies referred to as "non-descriptive" papers since they involve empirical investigations to determine how well artificial intelligence can assist in legal processes. These papers include developing and testing classifiers and conducting experiments to investigate the integration of artificial intelligence into legal workflows empirically.

Legal Domains where AI is Adopted

This section aims to answer the research question: "*In which legal domains is AI most widely adopted?*" All the papers have been analysed and classified into different legal domains to address this question. Six different categories have been identified: Lawyers & Law Firms, Justice & Court, Clients, Government & Lawmakers, and Law Enforcement. Additionally, a general "not specified" category is used to categorise papers that do not explicitly discuss a specific legal domain. The various domains and the corresponding number of papers explicitly applicable to each domain are detailed in Figure 4, while the exact studies per legal domain are shown in Table 8 in Appendix 1.

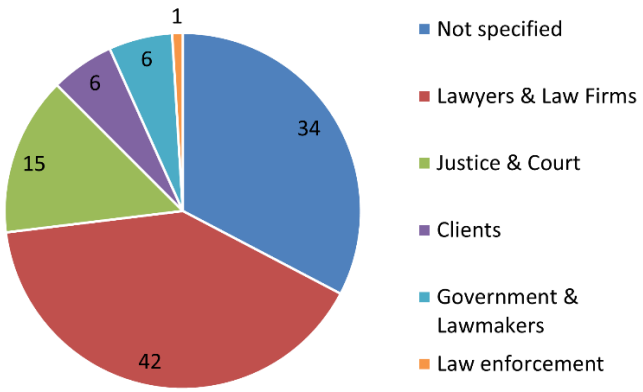

Figure 4. Presence of legal domains in research papers

The tables reveal a significant representation of lawyers and law firms in papers addressing AI in LegalTech. The majority of AI implementations and discussions are concentrated within this legal domain. Apart from focusing on lawyers and law firms, many papers address the legal world without specifying a particular legal domain. A smaller but significant proportion of papers delve into the application of AI in the justice and court sector. These papers suggest that AI could assist judges in their work or even replace them through automated dispute resolution solutions. Some papers explore the potential benefits of AI for clients, government, or law enforcement, although the scientific community has not yet studied these domains extensively.

Primary Applications of AI in LegalTech

The legal industry is not immune to change due to AI (Alarie et al., 2016). AI might replace many of what lawyers do, and AI systems may substitute a few tasks (Armour & Sako, 2020). However, law is a complex domain (Oskamp & Lauritsen, 2002); the sector is hesitant to adapt (Agrawal et al., 2022), and some legal tasks will remain beyond the capabilities of AI for the foreseeable future (Armour & Sako, 2020). Possible applications for which AI can be used in the foreseeable future are discussed in this section. We outline several applications that are discussed in the literature. These applications are identified and grouped into categories. With this categorisation, we aim to answer the research question: "*What are the primary applications of AI in LegalTech?*". Eight categories have been identified: eDiscovery, Litigation & Prediction analysis, Legal search, Document generation, Contract analysis, Due diligence reviews, and Dispute resolution.

Additionally, a general "Others" category is used to group all applications that have not been explored in multiple publications. The various applications and the corresponding number of papers discussing these applications are listed in Figure 5. The exact studies per application are listed in Table 9 in Appendix 2. This section explains all identified categories.

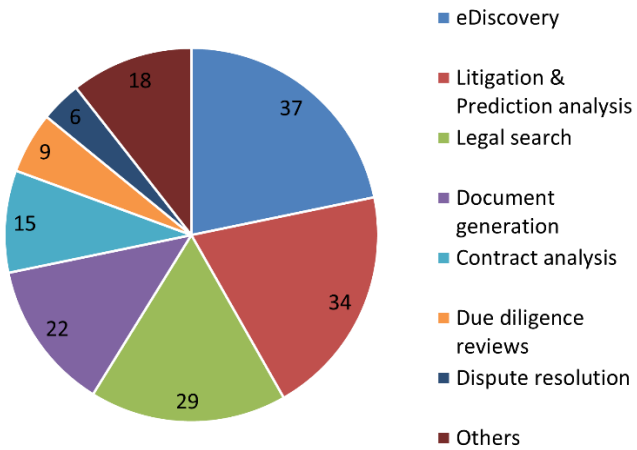

Figure 5. Number of AI applications in research papers

eDiscovery

The volume of potentially relevant data in a large lawsuit has grown exponentially, creating demand for tools that can help identify potentially relevant documents more quickly and cheaply (Armour et al., 2022). Using AI, large amounts of data can be analysed in a much better way (Vardhan et al., 2021). It can process large amounts of data quickly (Khabibullina et al., 2019). eDiscovery is used on large datasets. AI can automate this process in a fast manner (Baser & Saini, 2023). It reduces what may take months of laborious screening to days, if not in some cases even hours (Burt, 2021). The technology helps legal professionals to manage, classify, categorise and relate documents more smartly (Baser & Saini, 2023). One example of a legal AI eDiscovery system given in literature uses Human Aided Computer Cognition, which uses a combination of user modelling and document assessment and measurement to calculate relevance and a mechanism for iteration (Hogan et al., 2009). The resulting system exhibits exceptional performance, exceeding full manual search and traditional Computer-Aided Human Cognition (Hogan et al., 2009).

Litigation & Prediction Analysis

With litigation and prediction analysis, the outcome of litigation and court cases can be predicted. Machine learning algorithms have been used to predict legal case outcomes by finding patterns in similar cases (Baser & Saini, 2023; Sivaranjani et al., 2019). This technology can be used by lawyers to adequately plan their litigation strategies and expedite settlement negotiations (Baser & Saini, 2023). Some tests in this field have already obtained an accuracy rate of 60 to 80 per cent (Francia et al., 2022; Scherer, 2019; Sivaranjani et al., 2019). However, there are some concerns about the explainability of how these models make decisions and the potentially biased data used to train such models (Francia et al., 2022). The technology aims not only to stop at the prediction of litigation outcomes but also to predict lawyers' chances of winning cases (Sivaranjani et al., 2019).

McKeown et al. (2020) proposed using an intelligent decision support system to automate legal judgment processes, in which AI is used to provide legal professionals with the knowledge and support they need to carry out their jobs (McKeown et al., 2020). This study aimed to design and develop an intelligent legal system that streamlines the current legal process. A case study was performed with the designed system, which "provided an abundance of knowledge and understanding of the subject matter, AI and Law" (McKeown et al., 2020, p.32).

The most recent literature on Machine Learning- and Deep Learning-based techniques for rule decision prediction used in predicting judicial decisions has been reviewed in a publication by Francia et al. (2022). They concluded that the application of artificial

intelligence techniques in the legal domain, especially in predicting judicial decisions, has been increasing in recent years. When applying the techniques for prediction, Francia found that these techniques in different legal domains reach an accuracy of 60% (Francia et al., 2022).

Legal search

Dale (2019, p. 212) defines legal search as "the process of finding information that is needed to support legal decision-making". In practice, this generally means searching through both statute (as created by the legislature) and case law (as developed by the courts) to find what is relevant for some specific matter at hand." AI can be used to utilise search tools. These tools assist lawyers in gaining insights from a large number of legal documents (Baser & Saini, 2023).

Many papers describing legal search also describe eDiscovery, as can be noted from Table 9 in Appendix 2. Both implementations have many similarities, as both applications are used for searching for information. However, these two applications are different. Legal search is used to search for information in public databases and paid information resources. eDiscovery is used to search through local sources for evidence in litigation. An example of these sources is company emails.

According to Burt (2021), AI is used to return more meaningful and insightful results. Machine learning algorithms learn from users' search strategies to show more effective solutions to search queries (Armour et al., 2022). The algorithms can be used to analyse search results, highlight words according to relevance, and identify links between cases (Armour et al., 2022). Using AI for search purposes means that AI can focus on the mechanical aspects of lawyering so that lawyers can spend more time with clients (Hilt 2017).

Document generation

Document generation, also called document automation or document assembly in the literature, is the creation of legal documents tailored to specific criteria (Dale, 2019; Lauritsen, 1992). Where in legaltech AI is used for contract analysis, does this application go further to fully generate documents, mostly with minimal human input? Document generation can be achieved by using template mechanisms or question-and-answer dialogue, such as a chatbot (Dale, 2019).

There are multiple reasons to automate document creation. These document generators can save costs, increase lawyers' efficiency, open up access to law and justice, and make legal services more accessible and affordable (Armour and Sako 2020; Soukupová 2021).

The literature debates whether AI can completely replicate lawyers' tasks by drafting documents. Armour (2020) argues that lawyers' tasks should be automated, not jobs. It is not necessary to replicate what lawyers do, as long as what lawyers deliver can be done more cheaply.

Ivaschenko et al. (2023) present a software solution for text understanding and generation of normative legal documents. His solution is based on ontologies. They proved that AI can generate new standard regulatory legal documents (Ivaschenko et al., 2023).

Contract analysis

An everyday activity for lawyers is to review contracts, make comments and changes, and advise their clients on whether to sign or negotiate for better terms (Dale, 2019). AI-based solutions for reviewing contracts provide great help to legal workers in doing routine tasks of reviewing and highlighting documents so that they can concentrate on more complex and highly valuable tasks (Baser & Saini, 2023). AI tools can identify important information in contracts for particular analysis and monitoring, such as termination dates, choice of law provisions, and clauses that have high value in complex undertakings (Burt, 2021).

Large amounts of data can be analysed in a much better manner (Vardhan et al., 2021). Lawyers would manually review large numbers of documents in short time frames, which could lead to error-prone decisions (Donnelly & Roegiest, 2020). AI systems would reduce the time spent on each file by 90%, according to Zhang et al. (2020). However, for contract analysis, it is impossible to trust AI completely to solve these legal issues (Khabibullina et al., 2019). Lawyers remain necessary, but their expertise helps interpret results from AI support systems and interface these with human clients (Armour and Sako 2020).

Due diligence reviews

Due diligence is the process of uncovering background information for legal practitioners (Baser & Saini, 2023). These reviews are often performed for corporate transactions for mergers and acquisitions (Burt, 2021). It is a specific overall process that includes aspects of earlier discussed applications, legal research, electronic discovery, and contract review (Sivaranjani et al., 2019), which are used to highlight and classify essential clauses (Chandra et al., 2020).

Due diligence reviews are essential for lawyers to advise clients on their actions (Sivaranjani et al., 2019). Traditionally, these reviews were time-consuming, with lawyers manually reviewing large numbers of documents in short time frames, leading to potential errors (Donnelly & Roegiest, 2020). However, the introduction of AI-powered tools has significantly sped up this process, enhancing its efficiency and reducing the risk

of mistakes (Baser & Saini, 2023). AI helps in finding important parts in contracts, legal cases, or government investigations that could have significant financial consequences. Human reviews may sometimes miss these crucial provisions, which could lead to substantial financial risks (Burt 2021).

Dispute resolution

Many disputes never reach court due to perceived barriers preventing citizens from pursuing legal action (Vladimirovich & Sergeevich, 2022). Online or alternative dispute resolution could help alleviate the workload of judges. Online dispute resolution, a growing legal technology endorsed and adopted by courts for alternative dispute resolution (ADR) through online negotiation, mediation, or arbitration, is also benefiting from AI development (Gingras & Morrison, 2021). Participants in such systems may benefit from shorter decision-making deadlines, reduced fees, and access to a higher standard of decision-makers, contributing to a more efficient and accessible dispute resolution process (Livson et al., 2021; Vladimirovich & Sergeevich, 2022).

Automated dispute resolution is hoped to improve the equality of legal decisions. It is believed that computers are less biased than humans, improving equality by law, as computers are usually not deliberately trained to discriminate based on colour or gender (da Luz Scherf et al., 2019; Delacroix, 2018). However, humans have a greater emotional and situational awareness, while electronic systems do not have empathy with humans in vulnerable cases like the loss of a job (Delacroix 2018). The complete automation of the legal proceedings process, even in the case of its potential technological feasibility, is said to be inopportune today (Miroshnichenko & Proscurina, 2021). Having automated judges and lawyers is unlikely, but the possibility is not unreal (Sil et al., 2019).

Other

Applications of artificial intelligence within legaltech that haven't been extensively explored by multiple authors are categorised as "Other". This category includes diverse applications such as dynamic allocation of legal cases to legal staff (Armour et al., 2022), the development of a lawyer bot (Armour & Sako, 2020; Xu et al., 2022), preparing invoices for clients, predicting the time duration of a case (Armour and Sako 2020), digital judging (D'Rosario, 2017), electronic filing of lawsuits (McKeown et al., 2020), criminal intent detection (Metsker et al., 2019), analysing the genuinely of expert testimony (Armour and Sako 2020), tasks that an attorney must personally handle (Mok et al., 2021), legal data analysis support (Munthuli et al., 2023), intelligent forms that apply the law (Munthuli et al., 2023), recommendations to customers (Munthuli et al., 2023), evaluating specific situations like divorce (Munthuli et al., 2023), work product retrieval (Niman & William, 2013), deception analysis (Barnett & Treleaven, 2018) and legal analytics (Shelar & Moharir, 2018).

Commonly Used AI Technologies

This section answers the question: *"Which AI technologies are commonly used in LegalTech applications?"* To answer this question, all sample papers are analysed to identify the technologies used. After that, all identified technologies are categorised into different categories. As mentioned in Section 3.1, all papers have been categorised into two main categories: "descriptive" and "non-descriptive". Descriptive papers explain possible benefits that may arise from the use of legaltech in the legal industry. These articles elaborate on the potential advantages, risks, and benefits of artificial intelligence for legal corporations. Furthermore, some papers explore the possibility of a disruptive change in the legal industry caused by the rise of artificial intelligence, estimating notable adaptations. Non-descriptive papers involve empirical investigations to determine how well artificial intelligence can assist legal processes. These papers include developing and testing classifiers and conducting experiments to investigate the integration of artificial intelligence into legal workflows empirically.

AI is a vast field with many branches. The papers identified two categories of AI fields overrepresented in adoption or discussion: machine learning (ML) and natural language processing (NLP). While machine learning can sometimes be used to complete NLP tasks, the technologies are discussed separately in this review.

Machine Learning (ML)

Machine learning explores the study and construction of algorithms. These algorithms can learn from and make predictions on data (Kohavi & Provost, 1998). The algorithms are distinguished from each other by how they learn to find patterns and produce an outcome that is desired by the user. There is no universal algorithm for detecting valuable patterns. The choice of algorithms depends on the specific problem the developer is solving and the available data (Mahesh, 2020). The most widely adopted machine learning methods are supervised learning and unsupervised learning, while reinforcement learning is sometimes used (Ongsulee, 2017).

Supervised Learning

Supervised learning is an algorithm that trains the model based on a labelled dataset. The input is associated with the corresponding output or target. Experts manually label a subset of the learning data, and the algorithm tries to predict the most probable output for the unlabelled data. Supervised learning is often used for classification and regression. The supervised machine learning algorithms and the studies that used these algorithms can be found in Table 1.

For example, Metsker et al. (2019) compared the application of supervised machine learning methods to analyse judgment open data. They compared Support Vector Machines (SVM), Logistic Regression (LR), Decision Trees (DT), Random Forest (RF), K-

Nearest Neighbours (KNN), Naïve Bayes (NB) and Gradient Boosting (GB), of which Naïve Bayes performed the best with 93% ROC coverage. Keeling et al. (2019) compared a Convolutional Neural Network (CNN) with SVM, LR and RF on text classification in legal document reviews with documents of varying lengths, with no algorithm that performed best on all used datasets. Csányi et al. (2022) built a multi-label classifier to classify legal documents by their subject to find similar legal cases. They evaluated five different algorithms: SVM, NB, LR, KNN and RF, with SVM performing best.

Unsupervised learning

Unsupervised learning is an algorithm that will train the model based on an unlabelled dataset. The model must discover patterns and structures in data without explicit guidance. Therefore, unsupervised learning is often less effective than supervised learning. Unsupervised learning is often used for clustering data and association. The unsupervised machine learning algorithms and the studies that used these algorithms can be found in Table 2.

Kuleshov et al. (2020) developed an unsupervised machine learning algorithm based on Latent Dirichlet Allocation (LDA) and Latent Semantic Indexing (LSI) to validate legal documents in their texts, including checking for the presence of mandatory structural elements. Vianna and Moura (2022) explored three different topic modelling approaches to organise and summarise complex legal documents, including the unsupervised machine learning approach CombinedTM (CTM).

Reinforced Learning

Reinforced learning is an algorithm involving an agent that learns to make decisions by interacting with an environment. The model learns when users are interacting with the model. Due to this interaction, agents receive feedback through rewards or penalties. Reinforcement learning is often used for game-playing AI and robotics. Examples of reinforcement learning algorithms include AlphaZero and Q-learning. Reinforcement learning algorithms have not been found in the selected studies.

Deep Learning

Deep learning is a sub-branch of machine learning. A deep learning algorithm involves training a neural network with multiple hidden layers (Ongsulee 2017). A deep learning algorithm "learns" gradually to understand and make decisions independently. It aims to mimic the way a human brain processes. It learns from information by using interconnected layers of nodes, called neurons or units, to identify patterns and features within complex datasets (Chowdhary 2020). Deep neural networks have been demonstrated to outperform other machine learning algorithms (Krizhevsky et al., 2012). They are often applied to fields like computer vision, automatic speech recognition, and natural language processing (Ongsulee 2017).

Table 1. Supervised machine learning algorithms present in research papers

Supervised algorithm	Descriptive papers	Non-Descriptive papers
Support Vector Machines (SVM)	de Siles (2021); Francia et al. (2022); Sivaranjani et al. (2019)	Al-Qurishi et al. (2022); Chalkidis et al. (2022); Csányi et al. (2022); Humphreys et al. (2020); Keeling et al. (2019); Metsker et al. (2019); Munthuli et al. (2023); Orosz et al. (2022)
Logistic Regression (LR)	de Siles (2021); Francia et al. (2022); Sivaranjani et al. (2019)	Csányi et al. (2022); Keeling et al. (2019); Metsker et al. (2019)
Decision Trees (DT)	de Siles (2021); Francia et al. (2022); Sivaranjani et al. (2019)	Keeling et al. (2019); Metsker et al. (2019)
Random Forest (RF)	de Siles (2021); Francia et al. (2022); Sivaranjani et al. (2019)	Csányi et al. (2022); Keeling et al. (2019); Metsker et al. (2019)
K-Nearest Neighbours (KNN)	de Siles (2021); Francia et al. (2022); Sivaranjani et al. (2019)	Csányi et al. (2022); Metsker et al. (2019)
Naive Bayes (NB)	de Siles (2021); Francia et al. (2022); Sivaranjani et al. (2019)	Braun et al. (2019); Callister (2020); Csányi et al. (2022); Metsker et al. (2019); Waltl et al. (2017)
Gradient Boosting (GB)	Francia et al. (2022)	Metsker et al. (2019)
Conditional Random Fields (CRF)		Leitner et al. (2019); Moreno-Schneider et al. (2020)
General	Armour et al. (2021, 2022); Delgado et al. (2022); Hunter 2020	Hogan et al. (2009); Iatrou (2022)
Other	de Siles (2021); Francia et al. (2022)	

The deep learning algorithms and the studies that used these algorithms can be found in Table 3. For example, Pande and Alam (2020) used deep learning algorithms like Recurrent Neural Network (RNN), Convolutional Neural Network (CNN) and Capsule Neural Network (CapsNet) models to analyse the semantics of employment-related judicial cases. With this semantic analysis, they tried to use these algorithms to predict the outcome of judicial cases. Jayasinghe et al. (2022) developed domain-specific auxiliary models to predict the winning party of legal cases based on a Recurrent Neural Network (RNN) using document encoders consisting of a single layer of either a Gated Recurrent Unit (GRU) or a Long Short-Term Memory (LSTM). Wang et al. (2019) developed a

Chinese legal system for consultation, law searching, and document analysis. They used a bidirectional LSTM architecture to obtain a fine-grained candidate set to use in their model.

Table 2. Unsupervised machine learning algorithms present in research papers

Unsupervised algorithm	**Descriptive papers**	**Non-Descriptive papers**
k-means	de Siles (2021)	
Latent Dirichlet Allocation (LDA)	Francia et al. (2022); Silva et al. (2022)	Kuleshov et al. (2020); Pande and Alam (2020)
Latent Semantic Indexing (LSI)		Kuleshov et al. (2020)
Correlated Topic Model (CTM)		Vianna & Moura (2022)
Global Vectors for Word Representation (GloVe)	Francia et al. (2022)	
Other		de Siles (2021)

Table 3. Deep learning algorithms present in research papers

Deep Learning algorithm	**Descriptive papers**	**Non-Descriptive papers**
Artificial Neural Network (ANN)	Francia et al. (2022)	D'Rosario (2017)
Convolutional Neural Network (CNN)	Francia et al. (2022); Silva et al. (2022)	Keeling et al. (2019); Pande and Alam (2020); Vardhan et al. (2021)
Recurrent Neural Network (RNN)	Francia et al. (2022)	Jayasinghe et al. (2022); Pande and Alam (2020); Wang et al. (2019)
Capsule Networks (CapsNet)		Pande and Alam (2020)
Gated Recurrent Unit (GRU)	Francia et al. (2022)	Jayasinghe et al. (2022)
Multilayer Perceptron (MLP)	Francia et al. (2022)	Vardhan et al. (2021)
Long Short-Term Memory (LSTM)	Francia et al. (2022); Silva et al. (2022)	Jayasinghe et al. (2022)
Bidirectional Long Short-Term Memory (BiLSTM)	Francia et al. (2022)	Leitner et al. (2019); Moreno Schneider et al. (2020); Wang et al. (2019)
Markov Logic Network (MLN)	Francia et al., 2022	
General	Gravett (2020); Rhim and Park (2019); Sivaranjani et al. (2019)	Donnelly and Roegiest (2020)

Natural Language Processing (NLP)

Natural Language Processing (NLP) is a branch of artificial intelligence that focuses on enabling computers to understand, interpret, and generate human language (Chowdhary, 2020). As society becomes increasingly digitised, many documents are written in natural

language. These documents' sheer volume makes it progressively challenging for humans to extract relevant knowledge from them, particularly within specified time constraints (Chowdhary, 2020). The aim of automating NLP is to efficiently and accurately handle automatic analysis and representation of human languages (Chowdhary, 2020).

Language Models

A language model forms the basis of many NLP applications. These models provide the ability to process, understand, and generate human language. Through these models, programs can be designed to interact with humans in a way they understand. Language models are used to understand the context and semantics of human text, but can also be used to predict or generate text. In the last few years, using LLMs has boosted the development of Natural Language Processing (NLP) (Delobelle et al., 2020). These models are created using transformer technologies. Pre-trained language models are effective in improving many NLP tasks, including sentence-level tasks, paraphrasing, and token-level tasks (Delobelle et al., 2020). Different families of Large Language Models have been identified in the reviewed literature, which can be found in Table 4.

Table 4. Families of large language models identified in the literature

Technique	**Descriptive papers**	**Non-Descriptive papers**
BERT	Francia et al. (2022); Silva et al. (2022)	Al-Qurishi et al. (2022); Chalkidis et al. (2020, 2022); Darji et al. (2023); Jayasinghe et al. (2022); Munthuli et al. (2023); Vianna and Moura (2022); Wang et al. (2019); Zhao et al. (2021)
GPT		Chalkidis et al. (2022)
Transformer encoder		Jayasinghe et al. (2022)
YARGY		Ivaschenko et al. (2023)

One of the most prominent pre-trained language models used in the legal domain is BERT. BERT is designed "to pre-train deep bidirectional representations from unlabelled text by jointly conditioning on both left and right context in all layers" (Devlin et al., 2018, p. 1). BERT can be fine-tuned for use in NLP in the legal sector. Chalkidis et al. (2020) saw that BERT achieved impressive performance in several NLP tasks. When using BERT, people can choose and use pre-trained models that are publicly available.

However, Chalkidis et al. (2020) state that BERT has been reported to be underperforming in specific domains, including the legal domain. According to Al-Qurishi et al. (2022), BERT has much potential for use in a narrow and specific domain, such as legal. Vocabulary in some fields significantly deviates from the natural language used for everyday purposes. Text written for legal or everyday purposes can differ significantly between formal formulations and text semantics. Some authors even describe the linguistic content used for legal matters as almost a language of its own (Al-Qurishi et al., 2022). Domain-specific adjustments in the training process could resolve this issue.

Table 5. Types of BERT identified in the literature

Technique	Descriptive papers	Non-Descriptive papers
BERT	Francia et al. (2022); Silva et al. (2022)	Al-Qurishi et al. (2022); Chalkidis et al. (2020, 2022); Darji et al. (2023); Jayasinghe et al. (2022); Munthuli et al. (2023); Vianna and Moura (2022); Wang et al. (2019); Zhao et al. (2021)
AraLegal-BERT		Al-Qurishi et al. (2022)
AraBERTv2-Large		Al-Qurishi et al. (2022)
ARBERT		Al-Qurishi et al. (2022)
BART		Chalkidis et al. (2022)
BERTopic		Vianna and Moura (2022)
BERTimbau		Vianna and Moura (2022)
BERTikal		Vianna and Moura (2022)
BigBird		Chalkidis et al. (2022)
BiMPM		Wang et al. (2019)
CamemBERT	Francia et al. (2022)	
CaseLaw-BERT		Chalkidis et al. (2022)
DeBERTa		Chalkidis et al. (2022)
DistilBERT		Jayasinghe et al. (2022)
FlauBERT	Francia et al. (2022)	
ftBERT	Francia et al. (2022)	
LegalBERT	Francia et al. (2022)	Chalkidis et al. (2020, 2022)
LegalBERT-SMALL		Chalkidis et al. (2020)
Longformer		Chalkidis et al. (2022)
mBERT		Al-Qurishi et al. (2022)
MiniLM		Darji et al. (2023)
ooBERT	Francia et al. (2022)	
RoBERT	Francia et al. (2022)	
RoBERTa		Chalkidis et al. (2022); Darji et al. (2023)
XLM-RoBERTa		Munthuli et al. (2023)
Sentence-BERT		Jayasinghe et al. (2022)

BERT has been mentioned in the literature as the technology for implementing NLP in the legal domain. All usage of BERT and creation of models that are adjustments to other BERT models are listed in Table 5. Chalkidis et al. (2020) evaluated different strategies when applying BERT to the legal domain. They found that further pre-training BERT or pre-training BERT from scratch on specific corpora performs better than using BERT out of the box for domain-specific tasks (Chalkidis et al., 2020). They also found that smaller BERT-based models can compete with larger, computationally heavier ones in specialised domains.

One language where a monolingual BERT model was successfully fitted to the legal domain is Arabic. Al-Qurishi et al. (2022) introduced AraLegalBERT, in which BERT was trained from scratch to utilise the potential of BERT for a narrow and specific domain. The model is fine-tuned and evaluated against three generic-linguistic Arabic BERT models in three natural language understanding tasks.

From the research of Al-Qurishi et al. (2022), we can conclude that their developed model achieved better accuracy than the original over legal text. When trained with domain-relevant examples, a base version of AraLegalBERT could reach comparable accuracy levels to larger versions of the general and original BERT model. Therefore, Al-Qurishi concluded that lighter, smaller domain-specific models could match the output of computationally more demanding models while reaching findings faster and expending far fewer resources (Al-Qurishi et al., 2022).

Another language model, Generative Pre-trained Transformers (GPT), is a well-known type of language model. It is often used for generative AI and was developed by OpenAI in 2018. GPT is based on the transformer architecture and is pre-trained on large data sets of unlabelled text using artificial neural networks (ANN). OpenAI's most recent foundation model, GPT-4, was released on March 14, 2023.

Named Entity Recognition

One NLP technique is Named Entity Recognition (NER). NER is a technology that identifies named entities in a body of text, most often including an assignment to a set of semantic categories, such as person, location, or organisation (Leitner et al., 2019). NER is a technique frequently used or mentioned in research on legaltech. Descriptive NER studies are presented in the works of Francia et al. (2022) and Silva et al. (2022) only. In contrast, non-descriptive contributions are presented by Al-Qurishi et al. (2022), Darji et al. (2023), Donnelly and Roegiest (2020), Ivaschenko et al. (2023), Leitner et al. (2019), Moreno-Schneider et al. (2020), and Vardhan et al. (2021).

Despite NER's development for languages and domains other than English and the legal domain, the legal domain has not been exhaustively addressed yet (Leitner et al., 2019). NER produces somewhat generalised results. For example, all persons, organisations, or dates in the text can be returned, but if you want to isolate a specific person, organisation, or date, then you need to apply additional logic (Ivaschenko et al., 2023).

Benefits, Challenges, and Limitations of AI Adoption in LegalTech

This section mainly answers research question 5: *"What are the reported benefits, challenges, and limitations of AI adoptions in legaltech?",* and discusses the most important observations of the overall literature review.

As seen from the research on the topic and the rise in papers in the last year, as shown in Figure 2, AI is an exciting opportunity or threat. Research question 5 reflects this issue. Many of the listed benefits and limitations have implications for law firms in the legal domain. In the section "Legal Domain where AI is adopted", it is already shown that a significant part of all reviewed papers is set in the context of law firms. When identifying listed benefits and limitations in the use of AI in the legal domain, many papers mention law firms and lawyers when listing these characteristics.

Benefits

AI can be used to complement the work of lawyers. AI is expected not to replace workers but only to assist them (Agrawal et al., 2022; Gravett, 2020; Hilt, 2017). The existing literature highlights multiple benefits associated with AI integration in legaltech. Foremost, the most mentioned benefit is the substantial improvement in efficiency (Alarie et al., 2018; Baser and Saini, 2023; Csányi et al., 2022; Francia et al., 2022; Gravett, 2020; Kronblad et al., 2023; Sil et al., 2019; Soukupová, 2021; Thornton, 2021; Vaciago, 2019). Alongside this, the literature underlines several general benefits, including improved accuracy (Agrawal et al., 2022; Csányi et al., 2022; Francia et al., 2022; Thornton, 2021), reduced costs (Agrawal et al., 2022; Francia et al., 2022; Gingras and Morrison, 2021; Gravett, 2020; Zhang et al., 2020), time-saving (Zhang et al., 2020), increased revenue (Baser and Saini, 2023; Gingras and Morrison, 2021; Gravett, 2020), and improved quality of work (Kronblad et al., 2023; Sil et al., 2019). AI has the potential to reduce the disparity between large and small law firms.

According to Alarie et al. (2018), smaller firms can now take on tasks that were previously reserved for larger firms, as AI reduces the reliance on extensive labour power, which traditionally only large firms could afford. This technological shift enables smaller firms to operate with a similar level of efficiency and capability as their larger counterparts (Gravett, 2020). Also, the need for lawyers to work exclusively in large law firms may diminish (Alarie et al., 2018). An aspect of AI that helps increase work efficiency is that, contrary to humans, computers are not subject to certain human factors or weaknesses (da Luz Scherf et al., 2019). They can work around the clock without sleep or caffeine (Gravett, 2020). Computers do not have poor health or get fatigued (Khabibullina et al., 2019).

Besides, AI can help analyse large volumes of data more effectively (Vardhan et al., 2021). Computers can process large amounts of data and information in a short time (Khabibullina et al., 2019). AI systems would reduce the time spent on each file by 90%, but some minutes of human validation for every analysed document are still needed (Zhang et al., 2020). Another positive characteristic of using AI algorithms, when compared to human work, is the absence of bias. While humans are naturally somewhat biased, computers are less biased than humans (Delacroix, 2018). The fundamental aim is that AI algorithms cannot be biased based on colour, gender, or other visible

characteristics and can make everyone truly equal (da Luz Scherf et al., 2019; de Siles, 2021), unless they are intentionally or unintentionally specifically trained otherwise.

A possible effect of using AI in the legal sector is the opportunity to change the work that lawyers can perform. Due to increased efficiency, lawyers can perform additional work for existing clients or perhaps help more clients (Alarie et al., 2018). Lawyers may be able to broaden rather than narrow their areas of specialisation (Alarie et al., 2018; Gravett, 2020). AI also minimises the routine component of a lawyer's work, allowing them to concentrate on finding non-standard solutions within their specialisation (Miroshnichenko & Proscurina, 2021). Therefore, by using AI, law firms will have opportunities, for instance, to introduce new products, services, and business models, become active in new legal areas, and position themselves differently in the market (Kronblad et al., 2023; Soukupová, 2021).

Challenges & Limitations

One of the most commonly cited challenges in implementing AI within the legal sector concerns liability and the potential risk of privacy infringements (Chandra et al., 2020). Law firms face difficulties complying with existing regulations when exploring new digital opportunities (Kronblad et al., 2023). Adopting new technologies requires the development of new legal regulations. As previously mentioned, the EU has announced a new digital package, which, in part, still needs to be put in place (Vladimirovich and Sergeevich, 2022). Additionally, firms are worried about the potential leakage of sensitive information (Agrawal et al., 2022; Hilt, 2017). The integration of AI is often linked to an increased chance of data threats, leakage of important company information, and mismanagement of essential information (Agrawal et al., 2022; Beebeejaun and Gunputh, 2023). Moreover, using AI is frequently linked to an elevated risk of cyberattacks (Beebeejaun and Gunputh, 2023).

The current capabilities of AI cause some challenges. According to multiple researchers, current AI algorithms cannot replace human judgment entirely (Alarie et al., 2016, 2018; Gravett, 2020). Lawyers must assess clients' sentiments, motivations, and reactions in real time, ensuring effective communication and determining whether additional questions or advice are necessary (Armour et al., 2021). Besides, humans have greater emotional and situational awareness (Delacroix, 2018). AI can only aid lawyers in executing routine tasks, such as drafting basic contracts, seeking court precedents, analysing information, and filing court documents (Khabibullina et al., 2019).

Another limitation of AI is its complexity. Often termed a "Black Box", AI algorithms' decision-making processes are challenging for most people to understand (Soukupová 2021; Zhang et al., 2020). This lack of transparency makes it difficult for lawyers to entirely trust AI for solving legal issues (Khabibullina et al., 2019). Moreover, this complexity has led to unrealistic expectations about what AI can achieve and its capabilities, creating impossible beliefs about its potential (Soukupová 2021).

When making an IT system for a legal firm, it must be noted that law is a complex domain (Oskamp and Lauritsen, 2002). Additionally, legal documents often include sensitive personal data that require additional care and precautions to comply with the GDPR (Council of the European Union, 2016). Ready-to-use and trustworthy AI solutions are still scarce and must be developed (Zhang et al., 2020). To build valuable applications for the legal sector that appeal to lawyers, insight into what lawyers do and want is necessary (Oskamp & Lauritsen, 2002). Training AI algorithms, especially when existing data is labelled manually, generally requires human professional expertise, which is costly (Armour et al., 2021). Besides, some of the stability and structure of the law are transferred to data scientists and developers (Callister, 2020).

Money can also be a significant limitation regarding the cost of implementing and maintaining an AI system, and how money is earned in a law firm. When implementing an AI system, the design and training require significant investments (Alarie et al., 2018; Hilt, 2017). Due to the high implementation costs and the fact that AI systems need a sufficiently large and relevant dataset, not all AI systems will be cost-effective in all cases where they may be technically feasible (Armour et al., 2021). When looking at the way of earning within law firms, they tend to operate based on billable hours (Dale, 2019; Gravett, 2020). Some researchers suggest that this principle is about to change. They believe that computers will reduce the workload for law firm employees (Gravett, 2020).

Research has also shown that the way law firms are structured sometimes also forms a limitation for adopting AI systems. Law firms tend to be structured through employee ownership (partners). This structure complicates the fundraising for innovative projects (Armour et al., 2022; Armour & Sako, 2020). Besides, it is currently shown that exploration of the possible use of AI is often encouraged by the individual motivation of partners instead of their structural reward systems (Kronblad et al., 2023). Heavy investments are hard to justify, especially when they require a possible change in the way of working, which can be met with certain reservations (Oskamp & Lauritsen, 2002). Because of these reasons, large projects like the implementation of AI are often postponed (Oskamp and Lauritsen 2002).

DISCUSSION

This review is organised into five distinct categories, corresponding to the predefined research sub-questions: publication trends, legal domains, AI applications, AI technologies and benefits, challenges, & limitations. The key observations are summarised and structured according to this categorisation.

Publication Trend

Tracing the historical development of AI in LegalTech, the first paper on this topic was published in 1988. From then until 2016, occasional publications marked the field's early

stages. However, a significant surge in interest and research can be observed from 2016 onwards, indicating a growing interest in AI in LegalTech. The collection of research papers (91) is almost evenly split between articles published in journals (50) and conference proceedings (41), reflecting a trend towards peer-reviewed research in this relatively new field. A classification system has been applied, categorising papers as descriptive or non-descriptive.

When examining the publication trend, it becomes clear that papers addressing the application of AI in the legal domain have been available for a considerable time. However, a substantial rise was detected around 2016. Since then, many papers have been released about the topic, indicating a significant interest in AI within the legal domain, with a considerable volume of scientific research already conducted. Furthermore, it is noteworthy that a substantial portion of this research undergoes rigorous peer review, indicating a commitment to thoroughly investigating this topic.

Legal Domains

The review examined in which legal domains AI is most widely adopted. While many papers offer a general perspective on the legal field without specifying a particular area of expertise, most papers focus on a single legal domain, concentrating on lawyers and law firms. Another noteworthy category includes papers discussing the field of justice, where the focus is on the augmentation or replacement of judgment and dispute resolution within the court system. Additionally, some papers explore how AI can support clients, government & lawmakers, and law enforcement.

It is interesting to see that many papers are focusing on the domain of lawyers and law firms. This observation leads to the conclusion that current research efforts by scientists are mainly concentrated on developing and implementing AI solutions within this specific legal field.

AI Applications

Several applications of AI within LegalTech have been identified. Of all the papers, the most discussed application is eDiscovery, in which AI categorises documents as relevant or irrelevant. Litigation & Prediction analysis is often mentioned as a promising application to predict the outcome of litigation and court decisions. Legal search involves the analysis of legal documents by AI in large databases to get better results and more of what the searchers are looking for. Document generation uses AI to create documents, while contract analysis encompasses examining contracts and identifying crucial information that may need additions, alterations, or deletions within the contractual framework.

Additionally, specific applications are occasionally referenced in the papers. One such application is due diligence reviews, employed for examining a wide array of corporate documents, particularly for assessment in the context of acquisitions or mergers. Predictive billing is another application used to optimise the allocation of fee earners and provide informed estimates of the costs associated with legal services. Moreover, the potential for dispute resolution through E-justice is acknowledged, where online, digital, or alternative dispute resolution mechanisms are leveraged to automate and optimise decision-making procedures. This section concludes that specific fields within AI applications are interesting for further research. The identification of promising applications suggests some opportunities for future investigation. This insight provides developers and researchers with diverse, prospective options and underscores the high potential for future work in AI applications.

AI Technologies

All papers have been analysed to identify AI technologies used, leading to the identification of two prominently discussed and adopted categories: machine learning and natural language processing. When looking at machine learning, supervised machine learning techniques are primarily represented. Also, in the field of machine learning with LegalTech, deep learning is a widespread technique. A clear trend emerges in natural language processing, with the BERT language model standing out as the most extensively employed model. BERT manifests in various iterations tailored to specific purposes within the LegalTech domain.

Threats to Validity

This research has some limitations. Firstly, the review process is conducted by a single researcher, involving the manual review of all papers without external validation. The selection procedure, including the judgment-based selection of papers, was also manually reviewed without external validation. Additionally, paper selection is primarily based on titles and keywords, potentially introducing biases in the inclusion criteria. Lastly, there was no subsequent search for additional papers after the initial selection procedure, and any mentioned papers not included in the sample during the initial phase were not added.

CONCLUSION

This literature review provides a structural understanding of the current state of research on AI utilisation in LegalTech. The review has resulted in a qualitative investigation of 91 selected papers about the use of AI in LegalTech. This study answered the primary research question: *"What is the current state of research on the use of artificial intelligence in the field of LegalTech?"* with five different perspectives. This research shows a growing interest in the use of AI in LegalTech. The findings of this research can serve as a foundational source for future investigations into the use of AI within the legal domain.

Many benefits, challenges, and limitations are mentioned and discussed in the reviewed papers. The most important are mentioned in the corresponding sector. We see multiple benefits for the use of AI in legal technology. AI can complement the work of lawyers and is considered a potential substantial improvement in efficiency. Alongside improving efficiency, AI could benefit from improved accuracy, reduced costs, time savings, increased revenue, and improved quality of work.

Furthermore, AI can decrease the gap between large and small law firms. Moreover, a notable benefit of AI algorithms is that, unlike humans, they do not experience fatigue, health issues, or fluctuations in productivity. Besides, AI can help analyse large volumes of data much better. AI algorithms cannot be biased by colour, gender, or other visible characteristics. A possible effect of using AI in the legal sector is the opportunity to change the work that lawyers can perform, allowing lawyers to broaden or further specialise their work. There are also some challenges and limitations to consider when discussing the use of AI within legal technology. One of the most commonly cited challenges in implementing AI within the legal sector is liability and the potential risk of privacy infringement.

Besides the challenges connected to privacy, litigation, and risk, additional challenges and limitations must be considered. Firstly, AI is considered very complex, making it challenging for most people to understand its functioning. Additionally, the field of law is very complex, making it difficult for IT specialists to develop AI programs for this specific sector. The cost of implementing and maintaining an AI system and how money is earned in a law firm are also considered limitations. The use of billable hours and the partner structure that most law firms have further add to these limitations. Moreover, the current capabilities of AI mean that lawyers cannot be replaced, so human input and validation are still required.

FUNDING

No funding was received for conducting this study. The authors have no financial or proprietary interests in any material discussed in this article.

DECLARATIONS

Conflict of Interest

The researcher declares no conflict of interest in this study.

Informed Consent

Not applicable as the data used in this research is publicly available.

Ethics Approval

Not applicable as no human data or human participation are involved.

REFERENCES

Agrawal, S., Sahu, A., & Kumar, G. (2022). A conceptual framework for the implementation of Industry 4.0 in legal informatics. *Sustainable Computing: Informatics and Systems, 33*, 100650. https://doi.org/10.1016/j.suscom.2021.100650

Al-Qurishi, M., AlQaseemi, S., & Soussi, R. (2022). AraLegal-BERT: A pretrained language model for Arabic legal text. In *Proceedings of the Natural Legal Language Processing Workshop 2022* (pp. 338–344). Association for Computational Linguistics.

Alarie, B., Niblett, A., & Yoon, A. H. (2016). Law in the future. *University of Toronto Law Journal, 66*, 423–428. https://doi.org/10.3138/UTLJ.4005

Alarie, B., Niblett, A., & Yoon, A. H. (2018). How artificial intelligence will affect the practice of law. *University of Toronto Law Journal, 68*, 106–124. https://doi.org/10.3138/utlj.2017-0052

Armour, J., & Sako, M. (2020). AI-enabled business models in legal services: From traditional law firms to next-generation law companies? *Journal of Professions and Organization, 7*(1), 27–46. https://doi.org/10.1093/jpo/joaa001

Armour, J., Parnham, R., & Sako, M. (2021). Unlocking the potential of AI for English law. *International Journal of the Legal Profession, 28*(1), 65–83. https://doi.org/10.1080/09695958.2020.1857765

Armour, J., Parnham, R., & Sako, M. (2022). Augmented lawyering. *University of Illinois Law Review*, 71–138.

Ashley, K. D. (1992). Case-based reasoning and its implications for legal expert systems. *Artificial Intelligence and Law, 1*(1–2), 113–208. https://doi.org/10.1007/BF00114920

Ayodele, M., Allmendinger, R., & Papamichail, K. N. (2020). Heuristic search in legaltech: Dynamic allocation of legal cases to legal staff. In *Lecture Notes in Computer Science* (pp. 326–338). Springer. https://doi.org/10.1007/978-3-030-64580-9_28

Barnett, J., & Treleaven, P. (2018). Algorithmic dispute resolution—the automation of professional dispute resolution using AI and blockchain technologies. *Computer Journal, 61*(3), 399–408. https://doi.org/10.1093/comjnl/bxx103

Baser, P., & Saini, J. R. (2023). AI-based intelligent solution in legal profession. In *ICT Systems and Sustainability* (pp. 75–84). Springer. https://doi.org/10.1007/978-981-19-5221-0_8

Beebeejaun, A., & Gunputh, R. P. (2023). A study of the influence of artificial intelligence and its challenges: The impact on employees of the legal sector of Mauritius. *Global Business Review*. https://doi.org/10.1177/09721509231193803

Braun, D., Scepankova, E., Holl, P., & Matthes, F. (2019). Consumer protection in the digital era: The potential of customer-centered legaltech. In *Lecture Notes in*

Informatics (pp. 407–420). Gesellschaft für Informatik. https://doi.org/10.18420/inf2019_58

Burt, J. A. (2021). The revolutionary impact of artificial intelligence on the future of the legal profession. *Kutafin Law Review, 8*(3), 390–402. https://doi.org/10.17803/2313-5395.2021.3.17.390-402

Callister, P. D. (2020). Law, artificial intelligence, and natural language processing: A funny thing happened on the way to my search results. *Law Library Journal, 112*(2), 161–212.

Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020, November). Legal-BERT: The muppets straight out of law school. In *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 2898–2904). Association for Computational Linguistics.

Chalkidis, I., Jana, A., Hartung, D., Bommarito, M., Androutsopoulos, I., Katz, D., & Aletras, N. (2022, May). LexGLUE: A benchmark dataset for legal language understanding in English. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (pp. 4310–4330). Association for Computational Linguistics.

Chandra, G., Gupta, R., & Agarwal, N. (2020). Role of artificial intelligence in transforming the justice delivery system in COVID-19 pandemic. *International Journal on Emerging Technologies, 11*(2), 344–350.

Chen, S., Wang, J., & Zhang, Q. (2023). Informetric analysis of research on application of artificial intelligence in legal practice. In *Proceedings of ICIITCEE 2023* (pp. 406–408). IEEE. https://doi.org/10.1109/IITCEE57236.2023.10091079

Chowdhary, K. R. (2020). *Fundamentals of artificial intelligence*. Springer. https://doi.org/10.1007/978-81-322-3972-7

Chui, M., Bryce, H., Hellen, M., Alex, S., & Alex, S. (2022). The state of AI in 2022—and a half decade in review. *McKinsey & Company*.

Chui, M., Yee, L., Hall, B., & Singla, A. (2023). The state of AI in 2023: Generative AI's breakout year. *McKinsey & Company*.

Council of the European Union. (2016). *General Data Protection Regulation (GDPR)*. https://eur-lex.europa.eu/eli/reg/2016/679/oj

Csányi, G. M., Vági, R., Nagy, D., Üveges, I., Vadász, J. P., Megyeri, A. & Orosz, T. (2022). Building a production-ready multi-label classifier for legal documents with digital-twin-distiller. *Applied Sciences, 12*(3), 1470. https://doi.org/10.3390/app12031470

Dale, R. (2019). Law and word order: NLP in legal tech. *Natural Language Engineering, 25*(2), 211–217. https://doi.org/10.1017/S1351324918000475

Darji, H., Mitrović, J., & Granitzer, M. (2023). Exploring semantic similarity between German legal texts and referred laws. In *Communications in Computer and Information Science* (pp. 37–50). Springer. https://doi.org/10.1007/978-3-031-35924-8_3

de Siles, E. L. (2021). AI, on the law of the elephant: Toward understanding artificial intelligence. *Buffalo Law Review, 69*, 1389–1469.

Delacroix, S. (2018). Computer systems fit for the legal profession? *Legal Ethics, 21*(2), 119–135. https://doi.org/10.1080/1460728X.2018.1551702

Delgado, F., Barocas, S., & Levy, K. (2022). An uncommon task: Participatory design in legal AI. *Proceedings of the ACM on Human-Computer Interaction, 6*(CSCW2), Article 352. https://doi.org/10.1145/3512898

Delobelle, P., Winters, T., & Berendt, B. (2020). RobBERT: A Dutch RoBERTa-based language model. In *Findings of the Association for Computational Linguistics: EMNLP 2020* (pp. 3255–3265). Association for Computational Linguistics.

Deng, Y. J., & Lin, Z. (2022). The benefits and challenges of ChatGPT: An overview. *FCIS, 2*(2), 81–83. https://doi.org/10.54097/fcis.v2i2.4465

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. https://arxiv.org/abs/1810.04805

Donnelly, J., & Roegiest, A. (2020). The utility of context when extracting entities from legal documents. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management* (pp. 2397–2404). Association for Computing Machinery. https://doi.org/10.1145/3340531.3412746

D'Rosario, M. (2017). Fair use defences during copyright litigation: Is the success of a fair use defence strategy predictable? *International Journal of Strategic Decision Sciences, 8*(2), 31–51. https://doi.org/10.4018/IJSDS.2017040103

Francia, O. A., del Prado, M. N., & Alatrista-Salas, H. (2022). Survey of text mining techniques applied to judicial decisions prediction. *Applied Sciences, 12*(20), 10200. https://doi.org/10.3390/app122010200

Gingras, D., & Morrison, J. (2021). Artificial intelligence and family ODR. *Family Court Review, 59*(2), 227–231. https://doi.org/10.1111/fcre.12569

Gordon, G., Rieder, B., & Sileno, G. (2022). On mapping values in AI governance. *Computer Law & Security Review, 46*, Article 105712. https://doi.org/10.1016/j.clsr.2022.105712

Gowder, P. (2018). Transformative legal technology and the rule of law. *University of Toronto Law Journal, 68*(Suppl. 1), 82–105. https://doi.org/10.3138/utlj.2017-0047

Gravett, W. (2020). Is the dawn of the robot lawyer upon us? The fourth industrial revolution and the future of lawyers. *Potchefstroom Electronic Law Journal, 23*, Article 6794, 1–37. https://doi.org/10.17159/1727-3781/2020/V23I0A6794

Hilhorst, R., & Van Engers, T. M. (2009). E-dossier at the Dutch Council of State: Design, implementation and lessons learned. In *CEUR Workshop Proceedings* (pp. 13–32). CEUR-WS.org.

Hilt, K. (2017). What does the future hold for the law librarian in the advent of artificial intelligence? *Canadian Journal of Information and Library Science, 41*(3–4), 211–227.

Hogan, C., Bauer, R., & Brassil, D. (2009). Human-aided computer cognition for e-discovery. In *Proceedings of the International Conference on Artificial Intelligence and Law* (pp. 194–201). Association for Computing Machinery. https://doi.org/10.1145/1568234.1568256

Humphreys, L., Boella, G., Caro, L. D., et al. (2020). Populating legal ontologies using semantic role labeling. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)* (pp. 2157–2166). European Language Resources Association.

Hunter, D. (2020). The death of the legal profession and the future of law. *University of New South Wales Law Journal, 43*(4), 1199–1225.

Iatrou, E. (2022). A normative model of explanation for binary classification legal AI and its implementation on causal explanations of answer set programming. In *CEUR Workshop Proceedings*. CEUR-WS.org.

Ivaschenko, A., Golovnin, O., Syusin, I., et al. (2023). Ontology-based text understanding and text generation for legal technology applications. In *Lecture Notes in Networks and Systems* (pp. 1080–1089). Springer. https://doi.org/10.1007/978-3-031-37963-5_75

Jayasinghe, S., Rambukkanage, L., Silva, A., De Silva, N., & Perera, A. S. (2022, November). Legal case winning party prediction with domain-specific auxiliary models. In *Proceedings of the 34th Conference on Computational Linguistics and Speech Processing (ROCLING 2022)* (pp. 205–213). https://aclanthology.org/2022.rocling-1.26/

Keeling, R., Chhatwal, R., Huber-Fliflet, N., Zhang, J., Wei, F., Zhao, H., Shi, Y., & Qin, H. (2019). Empirical comparisons of CNN with other learning algorithms for text classification in legal document review. In *Proceedings of the 2019 IEEE International Conference on Big Data (Big Data 2019)* (pp. 2038–2042). IEEE. https://doi.org/10.1109/BigData47090.2019.9006248

Khabibullina, A. S., Seleckaya, S. B., & Shpagonov, A. N. (2019). The problems of robotization of the legal profession. *Revista General de Derecho, 8*(6), 397–405.

Kitchenham, B. A., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Technical Report). Keele University & University of Durham.

Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine Learning, 2*, 271–274. https://doi.org/10.1023/A:1017181826899

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (Vol. 25, pp. 1097–1105). Curran Associates.

Kronblad, C., Pregmark, J. E., & Berggren, R. (2023). Difficulties to digitalize: Ambidexterity challenges in law firms. *Journal of Service Theory and Practice, 33*(2), 217–236. https://doi.org/10.1108/JSTP-05-2022-0120

Kuleshov, S., Zaytseva, A., & Nenausnikov, K. (2020). Legal tech: Documents' validation method based on the associative-ontological approach. In *Proceedings of the International Conference on Speech and Computer* (pp. 244–254). Springer.

Lauritsen, M. (1992). Technology report: Building legal practice systems with today's commercial authoring tools. *Artificial Intelligence and Law, 1*(2–3), 87–102. https://doi.org/10.1007/BF00118480

Leenes, R. E. (2001). Burden of proof in dialogue games and Dutch civil procedure. In *Proceedings of the International Conference on Artificial Intelligence and Law* (pp. 109–118). Association for Computing Machinery. https://doi.org/10.1145/383535.383549

Leith, P. (1988). The application of AI to law. *AI & Society, 2*, 31–46. https://doi.org/10.1007/BF01891441

Leitner, E., Rehm, G., & Moreno-Schneider, J. (2019). Fine-grained named entity recognition in legal documents. In *Proceedings of the International Conference on Semantic Systems* (pp. 272–287). Springer.

Livson, M., Eshtokin, S., Vasyukov, V., Yudina, E., Baybarin, A., & Pivneva, S. (2021). Impact of digitalization on legal regulation: Formation of new legal practices. *Journal of Law and Sustainable Development,* 9(2), Article e0749. https://doi.org/10.37497/REVCAMPOJUR.V9I2.749

da Luz Scherf, E., Silva, M., & Silva, J. (2019). In tech we trust? Some general remarks on law in the technological era from a third world perspective. *Journal of Juridical Opinion, 17*, 107–123. https://doi.org/10.12662/2447-6641oj.v17i25.p107-123.2019

Mahesh, B. (2020). Machine learning algorithms: A review. *International Journal of Science and Research,* 9(1), 381–386.

Mania, K. (2023). Legal technology: Assessment of the legal tech industry's potential. *Journal of the Knowledge Economy, 14*, 595–619. https://doi.org/10.1007/s13132-022-00924-z

McKeown, T., Mustafina, J., Magizov, R., & Gataullina, C. (2020). AI in law practices. In *Proceedings of the International Conference on Developments in eSystems Engineering (DeSE)* (pp. 27–32). IEEE. https://doi.org/10.1109/DeSE51703.2020.9450780

McLachlan, S., Kyrimi, E., Dube, K., Fenton, N., & Webley, L. C. (2023). LawMaps: Enabling legal AI development through visualisation of the implicit structure of legislation and lawyerly process. *Artificial Intelligence and Law, 31*, 169–194. https://doi.org/10.1007/s10506-021-09298-0

Metsker, O., Trofimov, E., Sikorsky, S., & Kovalchuk, S. (2019). Text and data mining techniques in judgment open data analysis for administrative practice control. In *Communications in Computer and Information Science* (pp. 169–180). Springer. https://doi.org/10.1007/978-3-030-13283-5_13

Metsker, O., Trofimov, E., & Kopanitsa, G. (2021). Application of machine learning metrics for dynamic e-justice processes. In *Proceedings of the Conference of Open Innovation Association (FRUCT)*. IEEE. https://doi.org/10.23919/FRUCT50888.2021.9347598

Milmo, D. (2023, February 2). ChatGPT reaches 100 million users two months after launch. *The Guardian*. https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app

Miroshnichenko, O. I., & Proscurina, D. S. (2021). The role of artificial intelligence in the professional legal sphere: Development tool or existential threat? In *Lecture Notes in Networks and Systems* (pp. 928–937). Springer. https://doi.org/10.1007/978-3-030-69415-9_103

Mok, J. R., Mok, W. Y., & Mok, R. V. (2021). Sentence classification for contract law cases: A natural language processing approach. In *Proceedings of the 18th International Conference on Artificial Intelligence and Law (ICAIL 2021)* (pp. 260–261). Association for Computing Machinery. https://doi.org/10.1145/3462757.3466074

Moreno-Schneider, J., Rehm, G., Montiel-Ponsoda, E., Rodríguez-Doncel, V., Revenko, A., Karampatakis, S., Khvalchik, M., Sageder, C., Gracia, J., & Maganza, F. (2020). Orchestrating NLP services for the legal domain. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC 2020)* (pp.

2332–2340). European Language Resources Association. https://aclanthology.org/2020.lrec-1.284/

Munthuli, A., Socatiyanurak, V., Sangchocanonta, S., Kovudhikulrungsri, L., Saksakulkunakorn, N., Chairuangsri, P., & Tantibundhit, C. (2023). Transformers for multi-intent classification and slot filling of supreme court decisions related to sexual violence law. *IEEE Access, 11*, 76448-76467. https://doi.org/10.1109/ACCESS.2023.3296261

Nazarenko, A., & Wyner, A. (2017). Legal NLP introduction. *TAL: Traitement Automatique des Langues, 58*, 7–19.

Niman, J., & William, S. (2013). Leveraging technology to become a better lawyer. In *ACM International Conference Proceeding Series*. Association for Computing Machinery. https://doi.org/10.1145/2466816.2466832

Nissan, E. (2017). Digital technologies and artificial intelligence's present and foreseeable impact on lawyering, judging, policing and law enforcement. *AI & Society, 32*, 441–464. https://doi.org/10.1007/s00146-015-0596-5

Ogurlu, Y., Shagieva, R. V., & Bersanov, A. S. (2022). A debate on artificial intelligence in the area of law and the legal professions. In *Lecture Notes in Networks and Systems* (pp. 365–374). Springer. https://doi.org/10.1007/978-3-030-93155-1_41

Ongsulee, P. (2017). Artificial intelligence, machine learning and deep learning. In *Proceedings of the 15th International Conference on ICT and Knowledge Engineering (ICT&KE)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICTKE.2017.8259629

Orosz, T., Vági, R., Csányi, G. M., Nagy, D., Üveges, I., Vadász, J. P., & Megyeri, A. (2022). Evaluating human versus machine learning performance in a LegalTech problem. *Applied Sciences, 12*(1), Article 297. https://doi.org/10.3390/app12010297

Oskamp, A., & Lauritsen, M. (2002). AI in law practice? So far, not much. *Artificial Intelligence and Law, 10*, 227–236. https://doi.org/10.1023/A:1025402013007

Oskamp, A., Tragter, M. W., & Lodder, A. R. (1999). Mutual benefits for AI & law and knowledge management. In *Proceedings of the 7th International Conference on Artificial Intelligence and Law* (pp. 126–127). Association for Computing Machinery. https://doi.org/10.1145/323706.323725

Pande, R., & Alam, S. (2020). Predicting the outcome of judicial cases using semantic analysis. In *Proceedings of the IEEE Symposium Series on Computational Intelligence* (pp. 1757–1761). IEEE. https://doi.org/10.1109/SSCI47803.2020.9308506

Porto, F. D. (2023). Algorithmic disclosure rules. *Artificial Intelligence and Law, 31*, 13–51. https://doi.org/10.1007/s10506-021-09302-7

Rhim, Y. Y., & Park, K. (2019). The applicability of artificial intelligence in international law. *Journal of East Asia and International Law, 12*, 7–30. https://doi.org/10.14330/jeail.2019.12.1.01

Robaldo, L., Batsakis, S., Calegari, R., Calimeri, F., Fujita, M., Governatori, G., Morelli, M. C., Pacenza, F., Pisano, G., Satoh, K., Tachmazidis, I., & Zangari, J. (2024). Compliance checking on first-order knowledge with conflicting and compensatory norms: A comparison among currently available technologies. *Artificial Intelligence and Law, 32*(2), 505–555. https://doi.org/10.1007/s10506-023-09360-z

Roberge, L., Long, S., Hassett, P., & Burnham, D. (2002). Technology and the changing practice of law: An entrée to previously inaccessible information via TRAC. *Artificial Intelligence and Law*, *10*(4), 261-282. https://doi.org/10.1023/A:1025433709192

Salmerón-Manzano, E. (2021). Legaltech and lawtech: Global perspectives, challenges, and opportunities. *Laws, 10*(2), Article 24.

Sanctis, F. M. D. (2021). Artificial intelligence and innovation in Brazilian justice. *International Annals of Criminology, 59*, 1–10. https://doi.org/10.1017/cri.2021.4

Scherer, M. (2019). Artificial intelligence and legal decision-making: The wide open? *Journal of International Arbitration, 36*(5).

Shelar, A., & Moharir, M. (2018). A comparative study to determine a suitable legal knowledge representation format. In *Proceedings of the 3rd International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT 2018)* (pp. 514–519). IEEE. https://doi.org/10.1109/ICEECCOT43722.2018.9001363

Sil, R., Roy, A., Bhushan, B., & Mazumdar, A. K. (2019, October). Artificial intelligence and machine learning–based legal application: The state of the art and future research trends. In *Proceedings of the 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS 2019)* (pp. 57–62). IEEE. https://doi.org/10.1109/ICCCIS48478.2019.8974479

Silva, H., António, N., & Bação, F. (2022). A rapid semi-automated literature review on legal precedents retrieval. In *Proceedings of the EPIA Conference on Artificial Intelligence* (pp. 53–65). Springer. https://doi.org/10.1007/978-3-031-16474-3_5

Sivaranjani, N., Jayabharathy, J., & Safa, M. (2019). A broad view of automation in legal prediction technology. In *Proceedings of the 3rd International Conference on Electronics, Communication and Aerospace Technology (ICECA 2019)* (pp. 180–185). IEEE. https://doi.org/10.1109/ICECA.2019.8822019

Soukupová, J. (2021). AI-based legal technology: A critical assessment of the current use of artificial intelligence in legal practice. *Masaryk University Journal of Law and Technology, 15*, 279–300. https://doi.org/10.5817/MUJLT2021-2-6

Sun, C., Zhang, Y., Liu, X., & Wu, F. (2020, July). Legal intelligence: Algorithmic, data, and social challenges. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 2464–2467). Association for Computing Machinery. https://doi.org/10.1145/3397271.3401466

Surden, H. (2022). Values embedded in legal artificial intelligence. *IEEE Technology and Society Magazine, 41*(1), 66–74. https://doi.org/10.1109/MTS.2022.3147542

Thomson Reuters Institute. (2023). *ChatGPT and generative AI within law firms*. Thomson Reuters. https://www.thomsonreuters.com/en-us/posts/wp-content/uploads/sites/20/2023/04/2023-Chat-GPT-Generative-AI-in-Law-Firms.pdf

Thornton, M. (2021). Legal professionalism in a context of uberisation. *International Journal of the Legal Profession, 28*, 243–263. https://doi.org/10.1080/09695958.2021.1901715

Vaciago, G. (2019). Opportunities and challenges in legal tech services in the Italian and European framework. In *Frontiers in Artificial Intelligence and Applications* (Vol. 317, pp. 280–289). IOS Press. https://doi.org/10.3233/FAIA190029

Vardhan, H., Surana, N., & Tripathy, B. K. (2021). Named-entity recognition for legal documents. In *Advances in Intelligent Systems and Computing* (pp. 469–479). Springer. https://doi.org/10.1007/978-981-15-3383-9_43

Vianna, D., & Moura, E. S. D. (2022). Organizing Portuguese legal documents through topic discovery. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 3388–3392). Association for Computing Machinery. https://doi.org/10.1145/3477495.3536329

Vladimirovich, M. A., & Sergeevich, E. K. (2022). Alternative dispute resolution in digital government. *Revista Brasileira de Alternative Dispute Resolution, 4*, 119–146. https://doi.org/10.52028/rbadr.v4i7.8

Waltl, B., Bonczek, G., Scepankova, E., Landthaler, J., & Matthes, F. (2017, July). Predicting the outcome of appeal decisions in Germany's tax law. In *Lecture Notes in Computer Science* (pp. 89–99). Springer. https://doi.org/10.1007/978-3-319-64322-9_8

Wang, Z., Wang, B., Duan, X., Wu, D., Wang, S., Hu, G., & Liu, T. (2019, November). iFlyLegal: A Chinese legal system for consultation, law searching, and document analysis. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics.

Xu, N., Wang, K. J., & Lin, C. Y. (2022). Technology acceptance model for lawyer robots with AI: A quantitative survey. *International Journal of Social Robotics, 14*, 1043–1055. https://doi.org/10.1007/s12369-021-00850-1

Yang, E., Lewis, D. D., & Frieder, O. (2021). On minimizing cost in legal document review workflows. In *Proceedings of the ACM Symposium on Document Engineering (DocEng 2021)*. Association for Computing Machinery. https://doi.org/10.1145/3469096.3469872

Zhang, Z., Hummel, J. T., Nandhakumar, J., & Waardenburg, L. (2020). Addressing the key challenges of developing machine learning AI systems for knowledge-intensive work. *MIS Quarterly Executive, 19*, 221–238.

Zhao, H., Ye, S., & Yang, J. (2021). An empirical study on transfer learning for privilege review. In *Proceedings of the IEEE International Conference on Big Data* (pp. 2729–2733). IEEE. https://doi.org/10.1109/BigData52589.2021.9672008

Appendix 1 Table with Papers on Legal Domain

Table 8. Legal domains in research papers

Legal Domain	Papers
Not specified	Al-Qurishi et al. (2022), Ashley (1992), Chalkidis et al. (2020), Chalkidis et al. (2022), Chen et al. (2023), Csányi et al. (2022), de Siles (2021), Delacroix (2018), Delgado et al. (2022), Francia et al. (2022), Hogan et al. (2009), Iatrou (2022), Ivaschenko et al. (2023), Keeling et al. (2019), Khabibullina et al. (2019), Leenes (2001), Leith (1988), Leitner et al. (2019), Mania (2023), Miroshnichenko and Proscurina (2021), Moreno-Schneider et al. (2020), Nazarenko and Wyner (2017), Ogurlu et al. (2022), Oskamp and Lauritsen (2002), Oskamp et al. (1999), Robaldo et al. (2024), Sil et al. (2019), Sivaranjani et al. (2019), Thornton (2021), Vaciago (2019), Vianna and Moura (2022), Wang et al. (2019), Yang et al. (2021), Zhao et al. (2021)
Lawyers & Law firms	Agrawal et al. (2022), Alarie et al. (2016), Alarie et al. (2018), Armour et al. (2021), Armour et al. (2022), Armour and Sako (2020), Ayodele et al. (2020), Baser and Saini (2023), Beebeejaun and Gunputh (2023), Burt (2021), Callister (2020), da Luz Scherf et al. (2019), Dale (2019), Darji et al. (2023), Delgado et al. (2022), Donnelly and Roegiest (2020), D'Rosario (2017), Gingras and Morrison (2021), Gordon et al. (2022), Gravett (2020), Hilt (2017), Hunter (2020), Jayasinghe et al. (2022), Kronblad et al. (2023), Kuleshov et al. (2020), Livson et al. (2021), McKeown et al. (2020), McLachlan et al. (2023), Mok et al. (2021), Nissan (2017), Orosz et al. (2022), Pande and Alam (2020), Rhim and Park (2019), Roberge et al. (2002), Scherer (2019), Shelar and Moharir (2018), Soukupová (2021), Sun et al. (2020), Vardhan et al. (2021), Waltl et al. (2017), Xu et al. (2022), Zhang et al. (2020)
Justice & Court	Agrawal et al. (2022), Barnett and Treleaven (2018), Chandra et al. (2020), da Luz Scherf et al. (2019), Gowder (2018), Hilhorst and Engers (2009), Livson et al. (2021), Mok et al. (2021), Moreno-Schneider et al. (2020), Nissan (2017), Sanctis (2021), Silva et al. (2022), Soukupová (2021), Sun et al. (2020), Vladimirovich and Sergeevich (2022)
Clients	Barnett and Treleaven (2018), Braun et al. (2019), Gowder (2018), Lauritsen (1992), Munthuli et al. (2023), Sun et al. (2020)
Government & Law makers	Agrawal et al. (2022), Livson et al. (2021), Metsker et al. (2021), Metsker et al. (2019), Porto (2023), Surden (2022)
Others	Nissan (2017)

Appendix 2 Table with Papers on AI Applications.

Table 9. AI applications in research papers

AI Application	Papers
Litigation & Prediction analysis	Alarie et al. (2016), Alarie et al. (2018); Armour et al. (2021), Armour et al. (2022); Armour and Sako (2020); Barnett and Treleaven (2018); Baser and Saini (2023); Beebeejaun and Gunputh (2023); Burt (2021); Callister (2020); Chalkidis et al. (2022); Delacroix (2018); Delgado et al. (2022); D'Rosario (2017); Francia et al. (2022); Gordon et al. (2022); Gravett (2020); Hunter (2020); Jayasinghe et al. (2022); Livson et al. (2021); Metsker et al. (2021); Miroshnichenko and Proscurina (2021); Mok et al. (2021); Munthuli et al. (2023); Nazarenko and Wyner (2017); Nissan (2017); Orosz et al. (2022); Pande and Alam (2020); Rhim and Park (2019); Scherer (2019); Sil et al. (2019); Sivaranjani et al. (2019); Soukupová (2021); Sun et al. (2020); Waltl et al. (2017)
eDiscovery	Alarie et al. (2018); Al-Qurishi et al. (2022); Armour et al. (2021), Armour et al. (2022); Barnett and Treleaven (2018); Baser and Saini (2023); Braun et al. (2019); Burt (2021); Callister (2020); Chalkidis et al. (2022); Chandra et al. (2020); Chen et al. (2023); Csányi et al. (2022); da Luz Scherf et al. (2019); Dale (2019); Darji et al. (2023); Delgado et al. (2022); Gravett (2020); Hogan et al. (2009); Humphreys et al. (2020); Hunter (2020); Keeling et al. (2019); Khabibullina et al. (2019); Leitner et al. (2019); Metsker et al. (2019); Moreno-Schneider et al. (2020); Nazarenko and Wyner (2017); Nissan (2017); Ogurlu et al. (2022); Orosz et al. (2022); Sil et al. (2019); Thornton (2021); Vardhan et al. (2021); Vianna and Moura (2022); Wang et al. (2019); Yang et al. (2021); Zhang et al. (2020); Zhao et al. (2021)
Legal search	Agrawal et al. (2022); Alarie et al. (2018); Armour et al. (2021), Armour et al. (2022); Armour and Sako (2020); Baser and Saini (2023); Beebeejaun and Gunputh (2023); Burt (2021); Callister (2020); Chalkidis et al. (2020); Chandra et al. (2020); Chen et al. (2023); Csányi et al. (2022); da Luz Scherf et al. (2019); Dale (2019); Gravett (2020); Hilt (2017); Humphreys et al. (2020); Khabibullina et al. (2019); Nazarenko and Wyner (2017); Ogurlu et al. (2022); Orosz et al. (2022); Oskamp and Lauritsen (2002); Sil et al. (2019); Soukupová (2021); Sun et al. (2020); Vardhan et al. (2021); Wang et al. (2019); Zhang et al. (2020)
Document generation	Alarie et al. (2016); Alarie et al. (2018); Baser and Saini (2023); Beebeejaun and Gunputh (2023); Callister (2020); Chandra et al. (2020); Dale (2019); Gingras and Morrison (2021); Hilhorst and Engers (2009); Hogan et al. (2009); Ivaschenko et al. (2023); Khabibullina et al. (2019); Kuleshov et al. (2020); Lauritsen (1992); Moreno-Schneider et al. (2020); Nazarenko and Wyner (2017); Oskamp and Lauritsen (2002); Sanctis (2021); Sil et al. (2019); Sivaranjani et al. (2019); Soukupová (2021); Sun et al. (2020)

Contract analysis	Agrawal et al. (2022); Armour et al. (2021), Armour et al. (2022); Armour and Sako (2020); Baser and Saini (2023); Beebeejaun and Gunputh (2023); Burt (2021); Chandra et al. (2020); Dale (2019); Gingras and Morrison (2021); Gravett (2020); Hunter (2020); Khabibullina et al. (2019); Miroshnichenko and Proscurina (2021); Ogurlu et al. (2022)
Due diligence reviews	Armour et al. (2021); Armour et al. (2022); Armour and Sako (2020); Baser and Saini (2023); Burt (2021); Chandra et al. (2020); Donnelly and Roegiest (2020); Sivaranjani et al. (2019); Soukupová (2021)
Dispute resolution	Armour et al. (2021); Armour et al. (2022); Delacroix (2018); Munthuli et al. (2023); Orosz et al. (2022); Vladimirovich and Sergeevich (2022)
Not specified	de Siles (2021); Iatrou (2022); Leith (1988); Porto (2023); Surden (2022)
Others	Armour et al. (2022); Armour and Sako (2020); Ashley (1992); Barnett and Treleaven (2018); Beebeejaun and Gunputh (2023); D'Rosario (2017); Hunter (2020); Leenes (2001); McKeown et al. (2020); Metsker et al. (2021); Metsker et al. (2019); Mok et al. (2021); Munthuli et al. (2023); Niman and William (2013); Nissan (2017); Rhim and Park (2019); Shelar and Moharir (2018); Xu et al. (2022)

Author's Biography

ir.drs. Maarten Sjoerd Looijenga is a Master's graduate in Business Information Technology from the University of Twente. His research focuses on natural language processing, machine learning, and transformer-based language models for the legal domain. His master's thesis explored the application of transformer models to Dutch legal texts. He also supports the implementation of AI solutions in professional practice to improve efficiency. Additionally, he is a qualified teacher with a strong interest in computer science and AI education.

Dr. Faiza Allah Bukhsh is an Associate Professor in the Data Management and Biometrics group at the University of Twente. Her research focuses on machine learning, process mining, event-driven intelligence, and data privacy. She has collaborated with healthcare providers and industry partners, including KPN, Vodafone, Philips, and ING. She also leads initiatives supporting women in computing and has contributed extensively to research in privacy and data-driven technologies. Dr. Bukhsh earned her PhD from Tilburg University and completed postdoctoral research in cybersecurity at the University of Twente.

Dr. Jeewanie Jayasinghe Arachchige earned her PhD in Information Systems and Management from Tilburg University in 2013 and holds a Master's degree in Information Technology from Keele University. Her research focuses on conceptual modelling, ontologies, service-oriented design, and process mining. She is currently a lecturer in the Department of Computer Science at Vrije Universiteit Amsterdam and previously worked

at the University of Twente. She has also served on program committees for several international conferences.

Frances Borkent holds an LL.M. in Law and Technology from Utrecht University and currently works as a legal advisor specializing in privacy and data management. Her professional and academic interests lie at the intersection of law and technology, with a particular focus on data protection, encryption, and the legal governance of emerging technologies.

Dr. Maya Daneva is an Associate Professor in the Services and Cybersecurity group at the University of Twente. She leads industry–university research collaborations and teaches in the Computer Science and Business & IT programs. Her expertise includes Requirements Engineering, Software Engineering, and Empirical Methods. Her current research focuses on digital transformation, crowdsourced software engineering, and large-scale Agile and DevOps practices, with an emphasis on security, accessibility, and resilience.