PROCEEDINGS OF SCIENCE

Modeling AI Framework for Secure Processing of Confidential Legal Documents

Giang Nguyen,[a,*] **Jozef Židuliak,**[a] **Adam Candrák,**[a] **Hlib Kokin,**[a] **Oleksandr Lytvyn,**[a] **Martin Komák,**[a] **Ján Lang,**[a] **Dominika Pintérová**[b] **and Peter Lukáčka**[b]

[a] *Faculty of Informatics and Information Technologies, Slovak University of Technology Ilkovičova 2, Bratislava 84216, Slovakia*

[b] *Faculty of Law, Comenius University in Bratislava, Šafárikovo námestie 6, Bratislava 810 00, Slovakia*

E-mail: giang.nguyen@stuba.sk, peter.lukacka@flaw.uniba.sk

The growing use of artificial intelligence (AI) systems in legal services necessitates careful consideration of how technological innovation can be aligned with data protection requirements and the preservation of professional confidentiality. This paper proposes a multilayer AI framework for secure, AI-driven processing of confidential legal documents, designed to integrate advanced language models with security controls, data governance mechanisms, and confidentiality safeguards aiming to reconcile technological innovation with the strict professional and legal obligations that define the legal domain. By combining retrieval-augmented generation, federated learning, and privacy-enhancing technologies, the framework aims to enhance efficiency, accuracy, and regulatory conformity in legal and compliance workflows. The paper also highlights the growing importance of distributed machine unlearning in light of the General Data Protection Regulation (GDPR) and the risk of memorization in large language models (LLMs). The work examines practical implementation challenges and trade-offs between model performance and privacy guarantees, emphasizing the importance of interdisciplinary collaboration between legal and IT professionals in both research and deployment. Such cooperation is essential to ensure that technical solutions are aligned with legal norms, ethical principles, and legal professional requirements.

International Symposium on Grids and Clouds (ISGC2026)
15-20, March, 2026
Academia Sinica, Taipei, Taiwan

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0) All rights for text and data mining, AI training, and similar technologies for commercial purposes, are reserved. ISSN 1824-8039 . Published by SISSA Medialab.

PoS(ISGC2026)023

1. Introduction

The legal profession, traditionally characterized by extensive research, documentation, and complex decision-making, is undergoing a profound transformation driven by the integration of artificial intelligence (AI). Technologies such as natural language processing, machine learning, and advanced data analytics are equipping legal professionals with sophisticated tools for legal research, contract and document analysis, and case management. Within the legal profession, AI is increasingly being used to analyze commercial contracts, generate and review their content, identify potential issues related to prospective contracting parties or risky clauses, and extract or modify key contractual provisions. Thus, AI is gradually changing the paradigm of how contracts are drafted and reviewed. It brings revolutionary tools that streamline the entire contract execution process, reduce costs, and minimize legal risks. The use of AI in the contract formation process not only accelerates and streamlines transactions but also increases legal certainty and prevents potential disputes [1]. These developments have significantly enhanced efficiency and accuracy, enabling lawyers to devote more time to strategic and value-added activities. In particular, AI-supported contract analysis facilitates the extraction of key terms and provisions, allowing practitioners to assess risks more rapidly and draft agreements with greater precision. However, the integration of AI into legal practice also raises fundamental ethical and legal concerns. Although AI offers substantial social and professional benefits, it simultaneously introduces risks, including data breaches, bias, lack of transparency, and challenges related to explainability and accountability. These risks are especially significant in the legal domain, where professional activities inherently involve the processing of personal data and other sensitive information.

The protection of personal data has become a central element of our society, particularly as social and professional interactions increasingly shift to the digital sphere. In the legal profession, the handling of personal and confidential information is not incidental but intrinsic to its core function. Lawyers are therefore subject to heightened regulatory obligations arising from both general data protection law—most notably the EU General Data Protection Regulation (GDPR)—and specific professional regulations, including the Act on Advocacy and the rules of the Slovak Bar Association.

Although protection of personal data and the duty of confidentiality [2] are formally distinct legal domains, they are closely interconnected in legal practice. Ensuring robust data protection constitutes a fundamental prerequisite for fulfilling the professional secrecy duty of the lawyer. Any breach of personal data protection can directly undermine this duty and compromise the trust that underpins the lawyer–client relationship. In this context, the growing use of AI systems in legal services requires careful consideration of how technological innovation can be aligned with data protection requirements and the preservation of professional confidentiality.

When using AI in the legal profession, it is also necessary to distinguish between so-called open systems, which operate with freely available and unreviewed sources within an online database, without any restrictions and without being specifically designed or trained for contract drafting, and those systems based on sources or data (e.g., contracts, court decisions, etc.) that have already been selected by an expert for the needs of a specific application and trained with a specific focus on contract drafting—particularly those in the field of commercial law. In our view, it is precisely this aspect—i.e., the creation of a relatively closed system trained for a specific purpose—that can significantly contribute to increasing the efficiency and, in particular, the quality of the outputs

PoS(ISGC2026)023

produced by the system. [28]

As collaboration between legal professionals and AI systems continues to evolve, the central challenge is to harness the efficiency and analytical power of AI while ensuring rigorous protection of confidential and personal data. Addressing this challenge requires not only regulatory compliance but also an appropriate technical and architectural design of AI systems deployed in legal practice. This paper therefore proposes a multilayer AI framework for the secure processing of confidential legal documents, designed to integrate advanced language models with security controls, data governance mechanisms, and confidentiality safeguards aiming to reconcile technological innovation with the strict professional and legal obligations that define the legal domain.

The remainder of the work is organized as follows. Section 2 presents a background on AI in Legal Contract Analysis and Drafting. Section 3 describes our multilayer privacy-preserving architecture. Section 3.2 discusses secure approaches for the modeling and processing of legal documents, such as federated learning and machine unlearning. Section 5 concludes the work and describes future work.

2. AI in Legal Contract Analysis and Drafting

2.1 Traditional AI Approaches

Manual Analysis of Legal Contracts The traditional approach to analyzing legal contracts consists of manually reviewing documents by legal experts. This process is time-consuming and financially costly, especially when handling large volumes of documents. Results from a study [3] comparing the performance of large language models (LLMs), specifically GPT-4, Claude, and PaLM 2, with junior lawyers and outsourced legal service providers in identifying legal issues in contracts show that LLMs achieve comparable output quality while significantly outperforming human experts in terms of processing time and cost efficiency.

Rule-Based Analysis Systems Before the widespread adoption of modern NLP models, legal text analysis was often conducted using rule-based systems built on regular expressions and domain-specific dictionaries. Although these methods enable precise detection of predefined patterns, their main limitation lies in their inability to effectively handle context and their limited capacity to generalize to new or non-standard formulations.

Statistical and Machine Learning Approaches Classical approaches based on methods such as TF-IDF, bag-of-words, simple classifiers or clustering [6] were primarily used for text retrieval and categorization tasks. However, these methods are not capable of effectively capturing semantic relationships [4] in legal documents, which limits their applicability in complex contract analysis.

2.2 Transformer and Domain-Adapted Models

A significant advancement in legal document analysis has been brought about by transformer models adapted to the legal domain, with LegalBERT achieving substantially better results than generic BERT models. The CUAD dataset [5] and the LexGLUE benchmark[7] enable systematic comparison of approaches and confirm that transformer models achieve higher accuracy, particularly in extracting complex contractual clauses and across a wide range of legal NLP tasks. Studies further

PoS(ISGC2026)023

show that combining domain-adaptive and task-adaptive pretraining with fine-tuning on annotated data significantly improves model performance, as also confirmed by language-specific solutions such as Italian Legal-BERT [8].

Language Models and Contextual Instruction-Guided Extraction Large language models (LLMs), such as GPT-4, Claude, or Gemini, enable extraction and classification tasks to be performed through prompt-based information extraction, without the need for explicit model fine-tuning. Through techniques such as zero-shot learning and few-shot learning, it is possible to provide structured instructions and examples directly within the input data [9].

Context window refers to the maximum amount of input text (measured in tokens) that a language model can process at one time. It includes previous inputs, instructions, and already generated responses. The size of the context window significantly affects the model's ability to handle long texts, maintain coherence, and correctly interpret semantic relationships.

Zero-shot, one-shot, and few-shot learning are machine learning paradigms that enable models to perform new tasks with little or no task-specific training data. In zero-shot learning, the model relies solely on general knowledge and semantic representations acquired during pretraining to solve a task without any examples. One-shot learning allows the model to adapt based on a single example, which is particularly valuable in domains where data collection is expensive or impractical. Few-shot learning [10] extends this approach by using a small number of annotated examples, enabling rapid and efficient task adaptation while avoiding the need for large training datasets.

Chain of thought enables language model explicitly generates intermediate reasoning steps leading to the final answer [11]. It allows the model to decompose complex problems into smaller, logically connected steps. The result is increased accuracy in tasks requiring logical, mathematical, or causal reasoning. Current LLM feature extended context windows that allow processing of complete legal documents within a single input prompt without the need for segmentation. Available models such as Gemini 2.5 Flash (1M-token context window), Claude 3.7 Sonnet (200k tokens), and LLaMA 3.3 70B (128k tokens) enable efficient processing of documents spanning dozens of pages while preserving global context.

Analysis of a corpus of Slovak contracts, including related annexes, has shown that the typical document length does not exceed 10–15 pages, which under standard tokenization represents a significantly smaller number of tokens than the limit of even the smallest considered models.

2.3 Retrieval-Augmented Generation (RAG) System for Contract Analysis and Drafting

Retrieval-Augmented Generation (RAG) combines the retrieval of relevant documents with generative models reduces the risk of hallucinations by constraining response generation to explicitly retrieved and verifiable sources, while enabling dynamic integration of external knowledge. For our legal domain, RAG is selected as the preferred architectural approach due to enhanced auditability of outputs, the possibility of traceability of used sources, and improved control over the inference process in legally sensitive tasks where reproducibility and verifiability of results are required [27]. However, the disadvantages of RAG systems include higher implementation complexity, the need for ongoing database maintenance, and increased computational demands.

PoS(ISGC2026)023

2.4 Large Language Models in Legal Professions

Large language models (LLMs), especially open-weight models (OWMs), which publicly release AI parameters [12], offer significant benefits but also pose serious risks. Their openness promotes transparency, reproducibility, innovation, and broader collaboration, yet it simultaneously enables misuse, as anyone can adapt the models for harmful purposes such as misinformation, deepfakes, or malicious code generation without the safeguards typical of API-based systems. Once model weights are public, protective guardrails can be removed, increasing the likelihood of unsafe outputs, potential weaponization for cyberattacks [25], or leakage of sensitive training data—risks that are difficult or impossible to reverse after release. Although open access reduces costs and democratizes research [28], it may also commoditize innovation, allowing large actors to benefit without contributing back, and can create legal liabilities if models were trained on copyrighted or sensitive data.

Table 1: Comparative analysis of LLM performance on legal reasoning tasks (LegalBench [28] [29] [30]).

Model	Provider	Approx. Accuracy	Latency	Cost $/M
GPT-5	OpenAI	84–85%	15 s	1.25 / 10.00 (est.)
Grok-4	xAI	83–84%	24 s	3.00 / 15.00 (est.)
Gemini 2.5 Pro Exp	Google	83.5%	3.5 s	1.25 / 10.00 (est.)
Gemini 2.5 Flash	Google	82.8%	0.4 s	0.15 / 0.60 (est.)
Claude (various)	Anthropic	81–82%	2.5 s	15 / 75 (est.)
Meta-Llama-3.1-70B	Meta	75–76%	-	open/varies
Llama-type (open)	Meta / community	75–80%	-	free / infra cost

LegalBench computes accuracy by evaluating the performance of LLM models separately on each legal task using task-appropriate metrics such as exact match, classification accuracy, or F1 score, depending on the format of the task (e.g. multiple-choice, extraction, or generation). The overall benchmark score is then obtained by averaging the individual task scores across all tasks in the suite to produce an aggregate performance measure. Scores are rounded and may vary depending on instruction configuration and benchmark version. The approximate accuracy denotes the aggregated performance in the LegalBench task suite, as reported by the public leadership boards and evaluation reports (Table 1).

$$\text{Aggregate Score} = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{1}$$

where $N \in \mathbb{N}$ denotes the number of tasks,

$S_i \in [0, 1]$ denotes the normalized score of task i.

From a deployment perspective, LLMs can be run locally on a dedicated infrastructure, giving organizations full control over data, customization, and security, but also placing the burden of maintenance, updates, and compliance entirely on the operator. Alternatively, they can be hosted via cloud services, which provide scalability and operational convenience yet introduce additional concerns regarding data governance, third-party dependence, and cross-border regulatory com-

PoS(ISGC2026)023

pliance. Consequently, open-weight models embody a fundamental tension between advancing research and innovation and maintaining safety, security, and legal control.

3. Privacy-Preserving Architecture

Our platform is based on a multilayer architecture, a term used consistently throughout this paper to refer to the logical separation of data, application/AI, presentation, and cross-cutting security functions, which ensures the following:

- Each user is authenticated and authorized, thereby guaranteeing secure access to sensitive information;
- One of the key features of the system is the generation of contracts based on predefined templates supported by AI models, including a risk assessment module to ensure personal data protection and compliance with applicable legislation;
- The platform provides an environment for feedback and continuous system improvement, enabling ongoing optimization of processes and outcomes;
- In addition, it supports the development of additional intelligent modules for specific subtasks, such as testing and improving the processing of legal documents using advanced natural language processing methods.

3.1 System architecture

Data and Integration Layer manages the collection, storage, and processing of data from internal and external sources, including legal databases, contracts, company registers, and publicly available business information. It ensures data quality, normalization, and consistency, which are essential for accurate contract analysis and risk assessment.

Application Logic and AI Layer contains the system's intelligent modules, including AI-driven contract generation, risk and regulatory compliance assessment (including data protection), legal document analysis, key clause identification, and contract evaluation. It also includes feedback mechanisms that support continuous system improvement and algorithm optimization.

User Interface Layer enables intuitive user interaction through dashboards for data visualization, user account management and authentication, and access to modules such as contract drafting, risk assessment, and business partner evaluation.

Security and Monitoring Layer ensures data protection, access control, auditing, and activity logging. It incorporates privacy-preserving mechanisms and independent system monitoring to minimize the risk of misuse of sensitive information.

PoS(ISGC2026)023

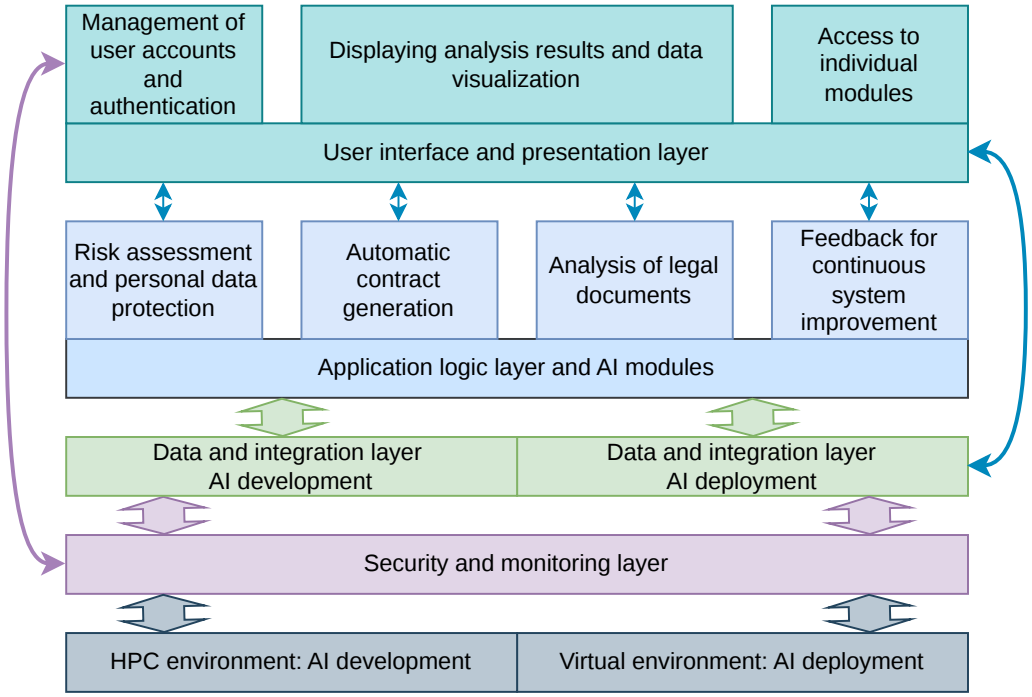

Figure 1: AILE system architecture

Data Flow and Privacy Controls Figure 1 represents a logical multilayer architecture in which confidential information is progressively controlled as it moves between data ingestion, AI processing, and user-facing services.

- At the Data and Integration Layer, legal documents and information obtained from internal and external sources are ingested, validated, normalized, and prepared for downstream processing. Data minimization is applied at this stage so that only attributes required for a particular legal-analysis task are propagated. Where the task permits, direct identifiers are removed or replaced through anonymization and pseudonymization before data are made available to AI components.

- The resulting data are passed to the Application Logic and AI Layer, which performs contract generation, legal-document analysis, clause identification, risk assessment, and related AI-supported operations. For RAG-based processing, relevant document fragments are retrieved from the controlled data layer and supplied as context to the language model rather than indiscriminately exposing the complete underlying corpus [16].

- In distributed model development, federated learning (FL) provides an additional separation: raw legal documents remain within the participating organization, while model parameters or updates are exchanged with the aggregation infrastructure. Depending on the deployment scenario, differential privacy (DP) can be applied to model updates or outputs to reduce information leakage, while homomorphic encryption (HE) or secure multi-party computation

PoS(ISGC2026)023

(SMPC) can protect data or model updates during privacy-sensitive distributed computation. Trusted execution environments (TEEs) provide a complementary mechanism for protecting sensitive processing in isolated execution environments.

- The User Interface and Presentation Layer receives only the results required by an authenticated and authorized user, such as generated contract content, document-analysis results, risk indicators, and visualizations. User corrections and feedback may flow in the opposite direction to the application layer for system improvement; such feedback is subject to the same minimization and access-control requirements because it may itself contain confidential information.

The Security and Monitoring Layer is therefore cross-cutting rather than a separate terminal stage of the pipeline: authentication and authorization, access control, auditing, activity logging, and privacy-preserving mechanisms govern interactions among the data, AI, application, and presentation components. In deployment, encryption in transit and at rest complements the PETs described above, whereas PET selection depends on the processing scenario rather than requiring every PET to be applied simultaneously.

3.2 Secure Legal Document Modeling and Processing

Privacy protection in privacy-preserving machine learning (PPML) can be achieved through the following measures:

- Prevention of identification: PPML techniques prevent anyone, including individuals involved in training or using the model, from identifying individuals in the data. This is achieved through methods such as anonymization or differential privacy.

- Limiting data exposure: PPML minimizes the amount of sensitive data that is made accessible during model training and deployment. This includes using only the necessary data attributes for collaborative machine learning (where models are trained locally on users' devices) or leveraging computation over encrypted data.

- Control over privacy and data security: PPML techniques enable individuals to decide which data is used, for what purpose, and for how long, within a secure and controlled environment.

Threat model. The architecture assumes that legal documents, retrieved RAG context, user feedback, model updates, and model outputs may contain personal, commercially sensitive, or legally confidential information. We consider threats from:

- unauthorized external actors attempting to obtain stored or transmitted information;

- authenticated but unauthorized or over-privileged users attempting to access documents outside their permitted scope;

- honest-but-curious infrastructure or collaborating parties that execute the prescribed protocol but may attempt to infer information from intermediate representations or model updates;

PoS(ISGC2026)023

- privacy attacks against trained models, including membership inference, model inversion, gradient leakage, and extraction of memorized training data.

The security objective is therefore to preserve confidentiality and limit disclosure in storage, communication, distributed training, inference, and presentation while retaining auditability and authorized access. PETs mitigate different parts of this threat surface: minimization and pseudonymization reduce exposed information; DP limits inference about individual records; HE and SMPC protect privacy-sensitive computation and collaboration; TEEs isolate sensitive execution; and federated learning avoids centralized collection of raw documents. These mechanisms are complementary and do not address all conventional security threats alone; authentication, authorization, encryption, logging, auditing, and infrastructure security remain necessary.

4. Federated Learning and Privacy-Enhancing Technologies

Building on the multilayer architecture introduced in Section 3.1, this section examines privacy-preserving mechanisms to protect confidential legal data during the development and distributed processing of AI models. In particular, we focus on federated learning and complementary privacy-enhancing technologies (PETs) that reduce the exposure of raw data and mitigate privacy risks arising from collaborative model training and inference.

4.1 Federated Learning

The transition from traditional machine learning to federated learning (FL) in the context of legal document processing [13] reflects a move away from centralized data aggregation to decentralized, privacy-aware model training [14]. Legal documents often contain highly sensitive and confidential information, such as personal data, contractual terms, and privileged communications, making centralized storage and processing legally and ethically challenging. Federated Learning enables models to be trained locally within law firms, courts, or corporate legal departments, allowing the system to learn from diverse datasets without transferring raw documents to a central server.

By sharing only model updates rather than the underlying legal texts, organizations can collaborate to improve document classification, contract analysis, or prediction of case outcomes while maintaining strict confidentiality requirements (Figure 2). This approach supports compliance with data protection regulations [26], reduces the risk of data breaches, and preserves client privilege, making it particularly suitable for legal environments where data security and trust are paramount.

In the proposed framework, horizontal cross-silo federated learning is considered the primary setting, where participating law firms retain legal documents within their local infrastructure and exchange only model updates with the coordinating server. SecFedAvg is considered the aggregation scheme, with local updates weighted according to the amount of training data available at each participating institution. Since legal datasets across law firms are expected to exhibit non-independent and non-identically distributed (non-IID) characteristics due to differences in client profiles, areas of legal practice, document types, and drafting conventions, the framework does not assume homogeneous local data distributions. This heterogeneity will be addressed through controlled local training and evaluation among participating clients, while more robust aggregation or personalization strategies may be considered where SecFedAvg exhibits convergence or performance degradation.

PoS(ISGC2026)023

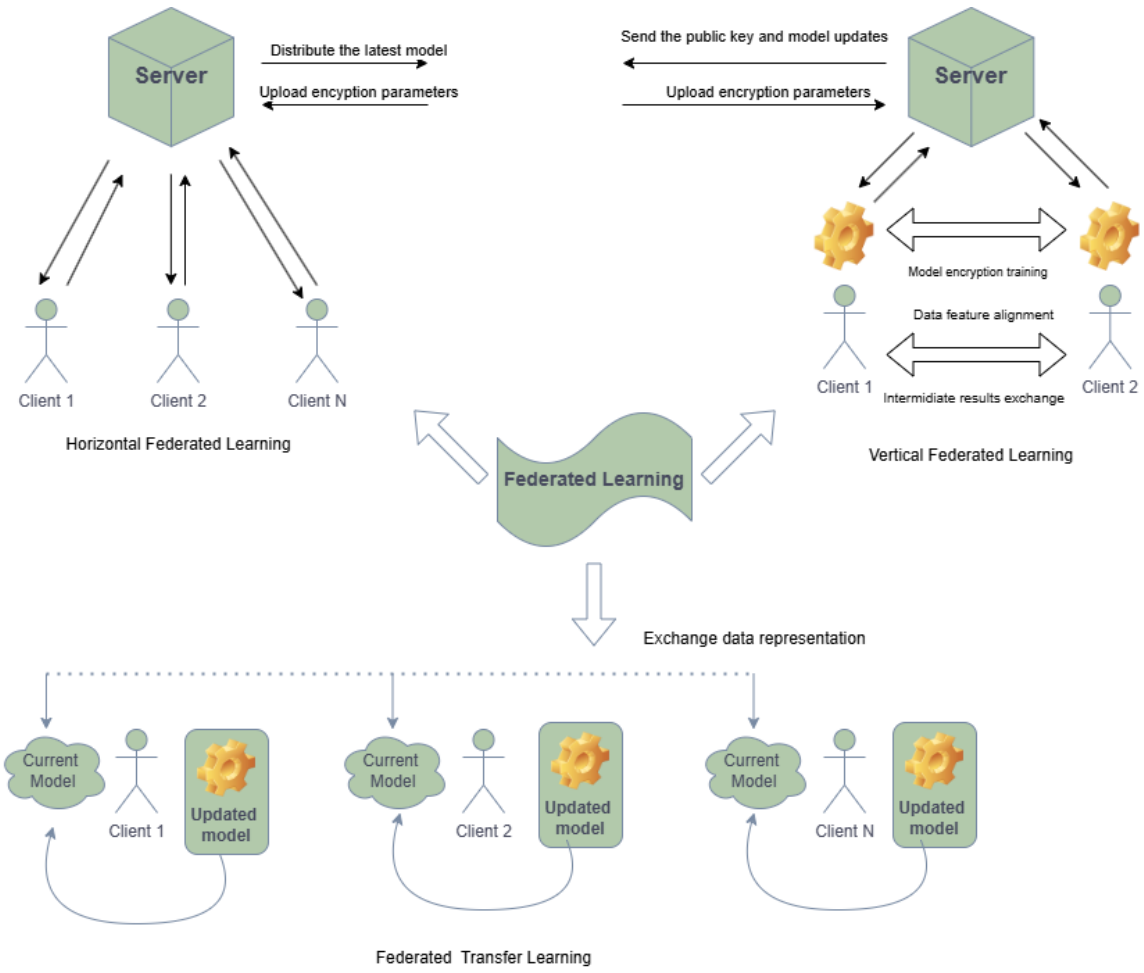

Figure 2: Transition from Machine Learning to Federated Learning

4.2 Privacy-Enhancing Technologies

Privacy-Enhancing Technologies (PETs) [19] approaches are particularly relevant in the PPML and federated learning context for legal document processing:

- Anonymization and Pseudonymization: Techniques that remove or replace personally identifiable information to reduce the risk of re-identification.
- Data Minimization Techniques: Limit data collection and processing to only what is strictly necessary for a specific purpose.
- Differential Privacy [17]: Adds carefully calibrated noise to data or model outputs to prevent the identification of individuals while preserving overall statistical utility.
- Homomorphic Encryption: Allows computations to be performed directly on encrypted data without decrypting it, ensuring data confidentiality during processing.
- Secure Multi-Party Computation [18]: Enables multiple parties to jointly compute a function over their inputs without revealing their private data to one another.

PoS(ISGC2026)023

- Trusted Execution Environments [20]: Use hardware-based secure enclaves to ensure that sensitive data is processed in an isolated and protected environment.

The importance of PETs and federated learning has increased significantly, particularly in connection with the legal document processing and the growing requirements for personal data protection arising from the GDPR as well as under lawyers' confidentiality duty. As modern AI methods operate on large volumes of often sensitive data, PETs represent a key tool to ensure that technological progress remains aligned with the legal and ethical principles of privacy protection.

For federated training, differential privacy is envisaged at the level of shared model updates through gradient clipping followed by calibrated noise addition before or during aggregation. Rather than fixing a universal privacy budget, the framework treats the privacy parameters (ε, δ) as deployment-specific configuration parameters, since the appropriate privacy–utility trade-off depends on dataset size, number of participating institutions, training rounds, and the sensitivity of the processed legal information. Future experimental evaluation will therefore report the resulting privacy budget together with model utility and convergence metrics, allowing the privacy guarantees to be assessed explicitly rather than assuming a predefined value.

4.3 Distributed Machine Unlearning

Machine unlearning is a complex with many unsolved questions, especially in the legal context of the *"right to be forgotten"* in Article 17 of the GDPR establishes a right to erasure, but provides limited practical guidance on how this right should be implemented when personal data has been used to train machine learning models. Given the non-deterministic nature of ML and the complex entanglement of training data within model parameters, the paper analyzes the challenges of removing personal data from trained models and examines how emerging machine unlearning techniques can support GDPR-compliant enforcement of the right to erasure [24].

Training distributed large language models (LLMs) over heterogeneous large-scale corpora introduces significant privacy risks, as overparameterized transformers are known to memorize low-frequency or sensitive training instances that can be later extracted via membership inference, model inversion, or data extraction attacks [21]. In federated or decentralized training settings, these risks are compounded by gradient sharing, cross-silo aggregation, and limited transparency in local updates, which may enable gradient leakage or reconstruction of private data. Furthermore, once sensitive information is embedded in model parameters through distributed optimization, its removal becomes non-trivial due to parameter entanglement and the absence of access to the original training shards. These challenges motivate the integration of PETs (Section 4.2) alongside scalable and verifiable machine unlearning techniques [22] [23] capable of selectively removing the influence of specific data points or clients without full retraining.

5. Conclusion and Future Work

Artificial intelligence offers significant benefits for legal contract analysis and drafting, but its deployment in legal practice must be carefully aligned with strict requirements of confidentiality, data protection, and professional responsibility. This paper proposed a multilayer privacy-preserving architecture that integrates advanced language models with security, auditability, and governance

PoS(ISGC2026)023

mechanisms tailored to the legal domain. By combining retrieval-augmented generation, federated learning, and privacy-enhancing technologies, the framework aims to reconcile technological innovation with regulatory compliance. We also highlighted the growing importance of distributed machine unlearning in light of GDPR Article 17 and the risk of memorization in large language models. Future work will focus on formally verifiable privacy guaranties, certified unlearning methods for distributed LLMs, and RAG validation in dedicated environments for legal professions.

Funding

This work is funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project Artificial Intelligence for Legal Professions No. 09I05-03-V02-00038. The dedicated environment is realized thanks to the support of the FIIT Cloud and the IDC Lab (IDSA) at FIIT STU in Bratislava.

Declaration

The use of AI tools in this work was for grammar correction and English language enhancement.

References

[1] D. Pintérová, "The impact of judicial decisions on the doctrine of commercial contracts in the context of artificial intelligence: challenges and new horizons," *Milestones of Law in the Central European Area 2025 (Volume 2)*, 2025, pp. 980–992. Available at: https://www.flaw.uniba.sk/fileadmin/praf/Veda/Konferencie_a_podujatia/Milniky/2025/Milniky_2025_2.pdf

[2] P. Lukáčka, D. Pintérová, and M. Daňko, "New horizons of contracting: the use of artificial intelligence in the creation of commercial contracts," *Intangible Goods Interdisciplinary – Calls for 2025*, 1st ed. Bratislava, Slovakia: Comenius University in Bratislava, Faculty of Law, 2025, pp. 143–157. Available at: https://doi.org/10.5281/zenodo.18351188

[3] Martin, L., Whitehouse, N., Yiu, S., Catterson, L., and Perera, R. (2024). Better Call GPT: Comparing Large Language Models Against Lawyers. *arXiv preprint arXiv:2401.16212*. Available at: https://arxiv.org/pdf/2401.16212

[4] Elhosary, E., and Moselhi, O. (2025). Evaluating Natural Language Processing Algorithms for Improved Hazard and Operability Analysis. *Geodata and AI*, p.100026. Available at: https://doi.org/10.1016/j.geoai.2025.100026

[5] Hendrycks, D., et al. (2021). CUAD: An Expert-Annotated NLP Dataset for Legal Contract Review. *arXiv preprint arXiv:2103.06268* [cs.CL]. Available at: https://arxiv.org/abs/2103.06268

[6] M. Kretinin and G. Nguyen, "Topic modeling on news articles using latent dirichlet allocation," in *2022 IEEE 26th International Conference on Intelligent Engineering Systems (INES)*, pp. 249–254, 2022. Available at: https://doi.org/10.1109/INES56734.2022.9922609

PoS(ISGC2026)023

[7] Chalkidis, I., et al. (2022). LexGLUE: A Benchmark Dataset for Legal Language Understanding in English. *arXiv preprint arXiv:2110.00976* [cs.CL]. Available at: https://arxiv.org/abs/2110.00976

[8] Licari, D., and Comandé, G. (2024). ITALIAN-LEGAL-BERT Models for Improving Natural Language Processing Tasks in the Italian Legal Domain. *Computer Law & Security Review*, 52, 105908. Available at: https://doi.org/10.1016/j.clsr.2023.105908

[9] Savelka, J., and Ashley, K. D. (2023). The Unreasonable Effectiveness of Large Language Models in Zero-Shot Semantic Annotation of Legal Texts. *Frontiers in Artificial Intelligence*, 6. Available at: https://doi.org/10.3389/frai.2023.1279794.

[10] Wang, Y., Yao, Q., Kwok, J. T., and Ni, L. M. (2020). Generalizing from a Few Examples: A Survey on Few-Shot Learning. *ACM Computing Surveys (CSUR)*, 53(3), 1–34. Available at: https://doi.org/10.1145/3386252

[11] Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., Zhou, D., et al. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems*, 35, 24824–24837. Available at: https://proceedings.neurips.cc/paper_files/paper/2022/file/9d5609613524ecf4f15af0f7b31abca4-Paper-Conference.pdf

[12] Leon, M. (2025). GPT-5 and Open-Weight Large Language Models: Advances in Reasoning, Transparency, and Control. *Information Systems*, 102620. Elsevier. Available at: https://doi.org/10.1016/j.is.2025.102620

[13] E. Quevedo, T. Cerny, A. Rodriguez, P. Rivas, J. Yero, K. Sooksatra, A. Zhakubayev, and D. Taibi, "Legal natural language processing from 2015 to 2022: A comprehensive systematic mapping study of advances and applications," *IEEE Access*, vol. 12, pp. 145286–145317, 2023. Available at: https://doi.org/10.1109/ACCESS.2023.3333946

[14] P. Ezhumalai, V. Sharmila, K. Balasaranya, *et al.*, "SmartLegalCraft: A Novel AI-Powered Platform for Adaptive and Efficient Legal Document Processing," in *Proceedings of the 2025 5th International Conference on Artificial Intelligence and Signal Processing (AISP)*, pp. 1–6, 2025. Available at: https://doi.org/10.1109/AISP68263.2025.11396328

[15] P. Lukáčka and G. Nguyen *Artificial Intelligence and the Legal Profession: Challenges and Risks in the Area of Personal Data Protection and Lawyers' Duty of Confidentiality*. In: *Comenius: Bratislava Legal Forum (BPF 2025)*, Vol. 1, pp. 374–390, 2025. ISBN 978-80-7160-779-3. Available at: https://doi.org/10.46282/bpf.2025.23.

[16] J. Židuliak, M. Babiar, A. Candrák, H. Davydau, V. Gusiev, P. Brandajský, M. Komák, and G. Nguyen, 2026. *LexAI: Intelligent AI System for Creating and Analyzing Commercial Legal Contracts*. In 2026 Information Technology Scientific Research Conference (IIT.SRC). IEEE. 2026. Available at: https://doi.org/10.1109/IIT.SRC67760.2026.11628842

PoS(ISGC2026)023

[17] H. Kokin, O. Lytvyn, and G. Nguyen, "Optimal differential privacy for deep learning model training," *Procedia Computer Science*, vol. 246, pp. 2419–2428, 2024. Available at: https://doi.org/10.1016/j.procs.2024.09.483

[18] O. Lytvyn and G. Nguyen, "Secure federated learning for multi-party network monitoring," *IEEE Access*, vol. 12, pp. 163262–163284, 2024. Available at: https://doi.org/10.1109/ACCESS.2024.3486810

[19] G. Nguyen and O. Lytvyn, *Foundations of Generative AI and Applied Machine Learning*. Spektrum STU Publishing, 2026. ISBN 978-80-227-5556-6. Available at: https://doi.org/10.61544/VNEV8972.

[20] F. Mo, H. Haddadi, K. Katevas, E. Marin, D. Perino, and N. Kourtellis, "PPFL: Privacy-preserving federated learning with trusted execution environments," in *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 94–108, 2021. Available at: https://doi.org/10.1145/3458864.3466628

[21] B. Yan, K. Li, M. Xu, Y. Dong, Y. Zhang, Z. Ren, and X. Cheng, "On Protecting the Data Privacy of Large Language Models (LLMs): A Survey," arXiv preprint arXiv:2403.05156, 2024. Available at: https://arxiv.org/abs/2403.05156.

[22] S. Liu, Y. Yao, J. Jia, S. Casper, N. Baracaldo, P. Hase, Y. Yao, C. Y. Liu, X. Xu, H. Li, K. R. Varshney, M. Bansal, S. Koyejo, and Y. Liu, "Rethinking machine unlearning for large language models," *Nature Machine Intelligence*, vol. 7, no. 2, pp. 181–194, Feb. 2025, doi: 10.1038/s42256-025-00985-0. Available at: https://doi.org/10.1038/s42256-025-00985-0.

[23] B. Liu, M. Ding, S. Shaham, W. Rahayu, F. Farokhi, and Z. Lin, "When machine learning meets privacy: A survey and outlook," *ACM Computing Surveys*, vol. 54, no. 2, pp. 1–36, 2021, Available at: https://doi.org/10.1145/3436755.

[24] B. A. Juliussen, J. P. Rui, and D. Johansen, "Algorithms that forget: Machine unlearning and the right to erasure," *Computer Law & Security Review*, vol. 51, p. 105885, 2023. Available at: https://doi.org/10.1016/j.clsr.2023.105885

[25] V. Khilenko, B. Akhmetov, R. Berdibayev, et al., "Increasing the speed of banking cybersecurity systems based on intelligent data analysis and artificial intelligence algorithms for predicting cyberattacks." *Cybernetics and Systems Analysis*, vol. 59, no. 4, pp. 519–525, 2023. Available at: https://doi.org/10.1007/s10559-023-00587-x

[26] V. Khilenko, "Managing the capital transfer market and optimizing management of the world economic and banking systems under the conditions of globalization. Regulating regulators." *Cybernetics and Systems Analysis*, vol. 59, no. 3, pp. 473–479, 2023. Available at: https://doi.org/10.1007/s10559-023-00582-2

[27] V. Khilenko, O. V. Stepanov, I. Kotuliak, and M. Reis, "Optimization of the selection of software elements in control systems with significantly different-speed processes," *Cybernetics*

PoS(ISGC2026)023

and Systems Analysis, vol. 57, no. 2, pp. 185–189, 2021. Available at: https://doi.org/10.1007/s10559-021-00342-0

[28] N. Guha, J. Nyarko, D. Ho, C. Ré, A. Chilton, A. Chohlas-Wood, A. Peters, B. Waldon, D. Rockmore, D. Zambrano, et al., "Legalbench: A collaboratively built benchmark for measuring legal reasoning in large language models," *Advances in Neural Information Processing Systems*, vol. 36, pp. 44123–44279, 2023. Available at: https://proceedings.neurips.cc/paper_files/paper/2023/file/89e44582fd28ddfea1ea4dcb0ebbf4b0-Paper-Datasets_and_Benchmarks.pdf

[29] LegalEvalHub, "LegalBench Full Leaderboard," 2026. Available: https://www.legalevalhub.ai/leaderboard/legalbench_full

[30] N. Guha et al., "LegalBench Dataset," Hugging Face Repository, 2023. Available: https://huggingface.co/datasets/nguha/legalbench

PoS(ISGC2026)023