HANDBOOK FOR POLICY MAKERS

AI AND JUSTICE

SHUBHAM RAJ

BlueRose ONE .com
Stories Matter
NewDelhi • London

BLUEROSE PUBLISHERS
India | U.K.

For permissions requests or inquiries regarding this publication, please contact:

BLUEROSE PUBLISHERS
www.BlueRoseONE.com
info@bluerosepublishers.com
+91 8882 898 898
+4407342408967

ISBN: 978-93-7542-313-3

Cover Design: Shivam Gaur
Typesetting: Sagar

First Edition: March 2026

Author's Note

Through this book I have tried to explain, in great depth, how we can leverage the developments of Artificial Intelligence in law and ensure that the process of Justice delivery is made more efficient. The intended audience of this book spans both the domains of Law and Software Engineering. Consider this as a legal book for the technocrats and a technical book for the legal professionals.

I hope that this innocuous attempt resonates with people from all walks of life, and the ultimate goal of ensuring justice to the citizens of this great nation is fulfilled- '*Satyameva Jayate*'.

Preface

If it was not for my revered teachers who taught me how to read and write, who am I?

If it was not for my revered Mother and Father who taught me how to live, who am I?

If it was not for my lovely sisters and brothers who taught me how to share and have fun, who am I?

If it was not for the universe which made it possible for me to have this experience, who am I?

I thank everyone who has ever been part of my life, as you have helped me in ways that are ineffable.

As I sit down to pen this preface, I am filled with excitement and gratitude. The convergence of law and technology has always been my passion, and this book is a testament to that passion. Throughout my journey as a Law student at Campus Law Centre, Delhi University, I have witnessed the profound impact that AI can have on the legal profession, and I am eager to share my insights with you.

In this book, I aim to provide a comprehensive exploration of how AI is reshaping the legal landscape across time zones. May this book come in handy as the years progress and help us all in reshaping our relationship with Artificial Intelligence (especially catering to the field of Law).

Author,
Shubham Raj

Dedicated to,

This book is dedicated to my mother, Rashmi Sinha; my father, Mithilesh Kr. Sinha; my beautiful sisters, Aishwarya Sinha and Dr. Pragati Prakash; and my wise brothers, Rahul Kishore and Rohit Kishore, without whose support and guidance this would not have been possible.

Contents

Chapter 1

Human in the loop

If someone asked me to define the capability and limitations of Artificial Intelligence especially with respect to the subject of Law, I would start by emphasising on the concept of 'Human in the loop'. The very first question that anyone who is curious about utilising the capabilities of AI in this field would have is: Are the outputs generated by Legal AI tools trustworthy? and the answer to that question is 'No', the answers are not to be trusted blindly. The answers generated by any software or application may carry some inaccuracies when we are using Generative AI technology.

"Generative AI is the field of AI which deals with generation of new material (e.g. generating text, generating images, generating videos etc.) ."

The concept of 'human in the loop' indicates that each and every output that is generated by a system is verified by a human at the other end of the spectrum before it is finally used for professional exchange. The reason why this is quintessential is because softwares/applications cannot guarantee 'accountability'. Whereas, the 'human in the loop' methodology ensures that the accountability rests on a human being.

In short, this methodology ensures that every AI generated output has to first go through a human being and then find its way in the actual workflows. To be more precise, this is not a methodology per se but an essential step that needs to be followed whenever we are dealing with AI generated content. Now coming back to the original question (*Are the outputs trustworthy?: Yes, when there is a human in the loop who can verify the outputs*).

Now the next logical question would be: How is this concept relevant to understand Legal AI ? This concept of 'Human in the loop' addresses the core issue of '*trustworthiness*' because of which Courts/Offices/Corporations around the globe are hesitant in adapting this technology. When we are certain that the results that are visible to us are trustworthy, then this significantly reduces the time spent on verifying the outputs.

While Legal AI softwares/applications are extremely efficient in analysing large volumes of legal data and identifying patterns, they inherently lack the ability to apply human judgment, empathy, and moral reasoning (elements which form the backbone of any justice delivery system). Therefore, any legal AI software/application is to be treated only as an assistant and it should never be treated as legal advice (unless verified by an actual Lawyer).

Let us understand this concept using a simple example.

In order to make justice widely accessible, a certain High Court decided to introduce a new software which could translate the judgements in local languages, thereby making the judgements easily accessible to large segments of the society. Let's call this software 'Y'. Henceforth, all recent judgements of the court are translated in 16 languages (hypothetical assumption regarding the number of languages). Now can we directly trust the output

generated by 'Y' and make these judgements available to the larger audience? No

In order to do that we will have to ensure that an intelligent human being on the other end of the spectrum has reviewed the output. Speaking from my experience as a Legal AI Engineer, these outputs generally go through multiple review cycles, before it reaches the lead reviewer. Only if this step is mandatorily followed can we say that the outputs generated by the software 'Y' are trustworthy (*practical application of human in the loop*).

Wait! So is the software 'Y' solving a problem or creating new ones? If you look closely, it has done the most difficult and mundane task of translation. The human translators are now focusing on reviewing the outputs and making pertinent miniscule changes. When we talk about Generative AI based technologies, 85-90 % accuracy is considered a very robust number. So the software 'Y' is automating 85-90 % of the work and also saving the amount of time it would take a translator to complete the task manually. At the same time, it also has the potential of creating newer dynamic roles, thereby augmenting the socio-economic well being of the nation at large.

In essence, the concept of *Human in the Loop* forms the backbone of responsible Legal AI adoption. While Artificial Intelligence can significantly enhance efficiency by automating repetitive tasks and processes, it cannot replace human judgment, accountability, or ethical reasoning. Trust in AI-generated outputs emerges only when a human being verifies and takes responsibility for them. Therefore, in the legal ecosystem, AI must remain an assistive tool augmenting human capabilities while ensuring that the ultimate control and responsibility continue to rest with humans.

In the last few paragraphs of this book, I would lay emphasis on the cardinal reason why this topic ('human in the loop') was chosen as the first chapter of a legal-tech book.

- I've recently interacted with various Judges and Lawyers of different high-courts who are tech-friendly and far sighted in their vision but all of them were having some sort of apprehension regarding the integration of Artificial Intelligence based technologies in the Courtrooms of India. All of their apprehensions were tied to either 'accountability' or 'trustworthiness of AI based outputs'. The topic 'human in the loop' directly hits the bull's eye and provides an answer to their genuine concern. The involvement of a human ensures that the accountability rests on a specific individual and the outputs go through a series of review cycles thereby ensuring trustworthiness of outputs.

- Trustworthiness is also ensured through 'provenance', which is the ability of any software to cite the actual source of the generated output. This means that the answers that are being displayed on the screen are either tied to an internal set of documents (uploaded on the software/tool) or linked to a reliable external database, which can be verified at the click of a button.

I hope this chapter removes the stigma surrounding the question: Whether the legal fraternity should embrace AI or not?

Now that the question regarding accountability and trustworthiness is addressed, let us look at the real world applications of AI in Law (what is happening and what are the possibilities?).

Chapter 2

Application of AI in Law

Before we start talking about the applications of this technology in the legal domain, it is important to be aware about 'automation'. Automation is a process by which a 'routine repetitive work' can be delegated to a software/application, which is done by infusing logic through programming languages.

Programmes can now be written in a variety of ways, but the easiest methodology which can even be used by a lawyer is '*natural language programming*' or '*vibe coding*'. You just have to give your AI assistant (which could be either a general purpose LLM such as chat gpt or a specialised LLM for coding such as codex) instructions in plain english language, which will then be converted into lines of code. These codes reflect the logic that we would want our software/application to have. When these logics are applied, the end result is automation.

"A Large Language Model (LLM) is an AI model trained on large volumes of textual data, which enables it to understand language patterns and generate coherent and meaningful text-based outputs. Examples of LLMs include ChatGPT, Claude, and LLaMA, among others. When AI systems are designed to process and generate multiple forms of data such as text, images, audio, video, and code they are referred to as Large Multimodal Model (LMM). In the case of LMMs, the training data is not limited to

text alone but may also include images, audio recordings, videos, and code, allowing the model to operate across multiple modalities."

At this stage, it becomes important to shed light on two crucial aspects. First, the possible use cases of Artificial Intelligence in the field of law, supported by real-world examples of adoption. Second, an equally important discussion on 'what AI is *not* capable of doing in the legal domain', based on the prevailing principles and limitations.

What AI can do for Law?

To understand its practical utility, it is useful to classify its applications into two broad categories. The first relates to *Legal Assistant capabilities*, where AI functions as a support system for lawyers and legal professionals. The second concerns *Design based augmentation*, particularly in the context of intelligent case management, file/folder management and workflow optimisation.

Legal Assistant Capabilities

1. Legal Research

As a Legal AI Assistant, a software or application can perform a variety of tasks. First and foremost, let us discuss *Legal Research*. Traditionally, legal research has involved navigating through vast volumes of statutes, judicial precedents, commentaries, and regulatory materials, a process that is both time-consuming and labour-intensive.

AI-powered legal research tools are capable of scanning large repositories of legal documents, identifying relevant statutory provisions and case laws, and presenting the results in a structured and accessible manner. This has the potential to

significantly reduce the time spent on preliminary research and allow legal professionals to reduce their cognitive load and focus on higher-value tasks.

However, the primary hindrance to its real-world implementation revolves around one critical factor: **accuracy**. How can end users ascertain that the results displayed before them are valid and reliable? The answer lies in **source preview**.

Modern legal AI tools address this concern by ensuring that every generated output is accompanied by a verifiable source preview. These sources may take the form of live links to internal documents present within the application's database or external links pointing to reliable legal repositories. Simply put, the best legal research tools currently available in the market provide both **internal and external source preview capabilities**.

Internal source preview is utilised when legal research is confined to an internal database or user uploaded documents within the application. External source preview, on the other hand, refers to the ability of the software to pull relevant and credible links from across the webspace, enabling independent validation of the generated output.

This approach where answers are not merely generated but are also supported by clearly identifiable sources ensures that the outputs are not random or speculative. Instead, they become verifiable, explainable, and practically useful. At this stage, we once again return to the foundational principle of *'Human in the Loop'*, where a legal professional personally reviews and validates the obtained results before relying upon them.

To illustrate this in practice, consider a scenario where a Legal AI research tool is tasked to identify judicial precedents on *vicarious liability*. The output generated by the system does not merely summarise the legal position, it also embeds direct links to

authoritative external databases such as **SCC Online** and **Westlaw**, enabling the end user to access the full text of the cited judgments with a single click. In addition to retrieving case law, advanced research tools are also capable of providing **citation history and treatment analysis**. Citation history and treatment analysis refers to the process of examining how a judicial decision has been subsequently interpreted or reconsidered by later courts. Citation history traces *where and how frequently a judgment has been cited*, while treatment analysis deals with the *precedential value of such citations*, it explains the manner in which decision has been followed, relied upon or overruled thereby assisting legal professionals in assessing its present precedential value, **Shepherd citation** is a practical example of this feature (all this is done while providing active links for instant verification).

This allows the researcher to instantly ascertain whether a particular judgment used in their final outputs are accurate or not (at every stage 'human in the loop' is the final authority and the role of any feature is confined to enhancing efficiency only).

Furthermore, the AI-generated research may also reference authoritative texts and commentaries, including well-recognised works such as **Blackburn on Contract**, thereby anchoring the analysis in established legal scholarship. This layered approach combining primary judgments, citation validation, and doctrinal commentary ensures that the research output is not merely informative but legally dependable.

At every stage, however, the responsibility of interpretation and reliance continues to rest with a legal professional, thereby reinforcing the foundational principle of *Human in the Loop*. Following this structured approach, the end user now has access to high quality legal research backed by source preview, citation history and treatment analysis (each aiding the end goal of obtaining accurate results by providing instantly clickable links

for cross-verification). This is how the legal-tech community has systematically addressed the pertinent issue of 'accuracy' for any task involving Legal Research.

This is perhaps the first time in history where every legal professional regardless of their economical and educational constraints can access the highest quality of legal research which has the capability of enhancing the overall eruditeness of this noble profession. Smaller law firms, chambers, and individual practitioners often lacked the means to support R&D intensive efforts and were consequently limited to traditional, manual research methodologies. The emergence of Legal AI tools has fundamentally altered this dynamic. Today, even a resource constrained organisation can access advanced legal research capabilities that allows them to conduct research with a level of depth previously reserved only for established market players. In this sense, AI has led to the *democratisation of legal knowledge.*

2. Drafting

Anyone who is in the legal profession knows the importance of drafting and AI based applications/softwares can certainly augment the drafting capabilities of the end users. Before we dive deep into this topic one thing needs to be understood clearly that the idea behind discussing these use cases is not to bypass the process of manual drafting, which is one of the most essential skills any legal practitioner possesses. The methods which are discussed in this topic should only be used when a professional level of manual drafting acumen has been attained by the practitioner. Many expert studies have pointed out that using Gen AI based technologies for drafting can to a large extent reduce the cognitive and writing skills, which forms the bedrock of legal education. Professors reading this segment should to the best of their abilities ensure that their students understand the

underlying principle as to why manual drafting is the best practice to follow (since banning something rarely works, its use can only be discouraged).

Now let us look at the types of drafting that can be done using AI based tools.

A. Instruction based drafting: In this methodology you simply write down the specific instruction in plain english language for generating your draft. Although this could be written in the form of simple instructions but it works best if the following elements are included within the instructions:

- Role (eg. Act as an Advocate with 20+ years of experience in drafting)
- Type of draft (eg. Writ petition)
- Context or body (the main story)
- Tonality (eg. professional)
- Intended audience (eg. Judges, law students)
- Constraints (eg. jurisdiction shall be India)

An example prompt to understand this:

Act as an Advocate with over 20 years of experience in constitutional and administrative law drafting in India. Draft a **Writ Petition under Article 226 of the Constitution of India**, seeking the issuance of a **writ of mandamus** against a Government College and the concerned University authorities. The petitioner is a student enrolled in a Government College affiliated with a public university. Due to administrative lapses and procedural irregularities attributable to the respondent authorities, the petitioner was unjustly deprived of the opportunity to appear for or be fairly evaluated in a scheduled

examination. Despite multiple representations and formal requests submitted by the petitioner, the authorities have failed to take corrective action or provide an opportunity for re-examination. The inaction of the respondents has resulted in serious academic prejudice and has adversely impacted the petitioner's right to education. The draft should maintain a formal, professional, and respectful tone, appropriate for constitutional litigation before a High Court. The draft is intended for consideration by Hon'ble Judges of a High Court and may also serve as a reference for law students studying constitutional remedies.

Constraints:

- Jurisdiction shall be limited to India.
- The draft must adhere to the conventional structure of a writ petition, including a brief statement of facts, grounds for relief, and a clear prayer clause.

As a thumb rule we must always be as elaborate and structured with our instructions as possible. When these instructions/commands are written in a structured manner, they are referred to as Prompts.

When these elements are clearly articulated within the instructions, the quality and relevance of the generated draft improves significantly. The role definition helps the system calibrate the level of sophistication and domain-specific language, while the type of draft and jurisdictional constraint ensure procedural and legal alignment. Providing adequate context allows the AI to maintain narrative coherence, and specifying tonality and intended audience ensures that the draft adheres to professional expectations.

It is important to note that the output generated through instruction-based drafting should always be treated as a first

draft. While such drafts may be structurally sound and linguistically coherent, they are inherently limited by the quality of the instructions provided and the data on which the system has been trained. Therefore, the legal practitioner must critically review, refine, and, where necessary, rework the draft to ensure accuracy, factual correctness, and compliance with applicable procedural laws. This reinforces the 'human in the loop' principle and establishes the fact that AI-assisted drafting is a productivity enhancer rather than a replacement for legal expertise.

B. Template based drafting: It operates on a slightly different principle where instead of generating a draft entirely from instructions, the AI system relies on *pre-existing templates* that have been curated, standardised, and approved over time. These templates may include commonly used legal documents such as petitions, contracts, notices, affidavits, and agreements, which are structured in accordance with established legal conventions and procedural requirements.

In this methodology, the role of the AI is primarily to customise and populate the template based on the specific facts provided by the end user. Variables such as party names, dates, jurisdiction, relief sought, and factual matrices are dynamically inserted into predefined sections of the document. This ensures consistency in structure while allowing flexibility in substance. Template-based drafting is particularly useful in high-volume or repetitive drafting scenarios, where uniformity and compliance with standard formats are critical. However, as with instruction-based drafting, the final responsibility of reviewing, modifying, and approving the document continues to rest with the legal practitioner, thereby reinforcing the principle of Human in the Loop.

Lets see an example to understand its use case better:

The end user has uploaded a pre-approved '*Licence Agreement template*' into a Legal AI application. The template already contains the standard clauses relating to grant of licence, term, consideration, confidentiality, termination, and governing law. The objective is not to generate a fresh contract, but to populate specific variables within the existing template.

Prompt:

Act as an experienced legal drafting assistant and follow the following instructions :

You are provided with an attached **Licence Agreement template**. Populate the relevant variables in the template using the following information:

- Licensor: ABC Technologies Private Limited
- Licensee: XYZ Solutions LLP
- Licensed Subject Matter: Non-exclusive licence to use proprietary software
- Licence Fee: INR 5,00,000 per annum
- Licence Term: Three years
- Territory: India
- Governing Law: Laws of India
- Effective Date: 1st April 2025

Do not modify the structure or standard clauses of the template. Only populate the relevant fields based on the information provided.

The primary problem that template-based drafting seeks to address is inefficiency arising from repetitive and standardised legal work. In the legal profession, a significant amount of time is spent on recreating documents that follow an almost identical structure, with variations limited to factual details such as party

names, dates, consideration, jurisdiction, and subject matter. This repetitive exercise not only consumes valuable professional time but also increases the risk of clerical errors and inconsistencies, particularly in high-volume drafting environments such as corporate transactions, compliance documentation, and routine litigation filings.

By automating the population of predefined variables within legally vetted templates, AI-based drafting tools enable legal professionals to redirect their attention towards substantive legal analysis, strategic decision-making, and client advisory. This approach aligns with the well-established legal maxim *justitia dilata est justitia negata*, which emphasises that delay at any stage of the justice delivery system ultimately undermines the cause of justice itself.

Although the maxim is most frequently invoked in the context of judicial delays, its underlying principle of timely and efficient legal processes is equally relevant to modern legal practice. Inefficiencies at the drafting and preparatory stages often cascade into procedural delays, thereby burdening both litigants and courts. Template-based drafting, when used responsibly and under human supervision, contributes to mitigating such delays by streamlining routine legal work while preserving professional accountability, accuracy, and legal rigour.

3. Due Diligence

Due diligence is the process of carefully examining the information contained within a document to ensure its accuracy, reliability, and legal validity before it is relied upon. Its purpose is to ensure that decisions are made on the basis of verified facts and legal realities, rather than assumptions.

To carry out due diligence on any document or group of documents, there must first be a reliable database, commonly referred to as a base document (a trusted reference database

containing verified and authoritative information). The documents being examined, known as target documents (the documents on which due diligence is performed), are then checked against the information contained within the base document. This base document must be a reliable set of documents containing verifiable data points with clear provenance capabilities. The moment an end user wishes to verify any information relied upon during the due diligence process, they should be able to do so instantaneously. The references used for verification are hyperlinked either to an internal database of pre-verified documents or to external reliable databases that are trustworthy and routinely cited during court proceedings, with any inconsistencies or deviations in the target document being clearly flagged.

Target document → Base document (internal + external database) ☒ Due diligence output

There are two possible variations of using diligence in a legal AI tool:

A. Local Diligence:

Local diligence refers to the process of verifying legal references, citations, and factual assertions against an internally curated and authenticated database that is either maintained by the organisation or uploaded by the end user. This database typically consists of pre-verified judgments, statutes, contracts, regulatory documents, and other authoritative materials that are considered reliable for professional and courtroom use. The objective of local diligence is to ensure that any legal authority cited within a draft, research output, or advisory note can be instantly traced back to a trusted internal source, thereby ensuring accuracy, provenance, and consistency. Since the verification occurs within a controlled ecosystem, local diligence is particularly useful in

litigation and advisory workflows where reliance on internally approved materials is preferred, and where rapid validation of cited authorities enhances both efficiency and credibility.

B. External Diligence:

External diligence refers to the process of verifying legal references, citations, and factual assertions by linking them to independent and authoritative external databases that are widely relied upon by courts and legal professionals. Unlike local diligence, which operates within an internal and controlled repository, external diligence enables real-time validation against publicly recognised legal sources. These sources include established legal research platforms such as SCC Online, Westlaw, and LexisNexis, as well as official government databases and websites. Government portals hosting statutes, rules, notifications, circulars, and regulatory filings play a crucial role in enhancing the reliability quotient of legal verification. References drawn directly from official legislative and regulatory websites provide an additional layer of authenticity, as they reflect the most current and authoritative position of the law. By integrating legal AI tools with both private legal databases and government-maintained databases, external diligence ensures that legal outputs are not only comprehensive but also grounded in sources that carry institutional and judicial credibility. This multi-source verification framework significantly strengthens accuracy, transparency, and courtroom reliability, while continuing to uphold the principle that final legal judgment rests with the human practitioner.

Let's understand this using a simple courtroom example:

Imagine a morning in a High Court where a Public Interest Litigation (PIL) concerning air pollution is listed for hearing. The petition does not merely raise a technical grievance. It speaks of

the everyday reality of citizens breathing polluted air and places its legal foundation squarely on Article 21 of the Constitution of India, arguing that the right to life must necessarily include the right to live in a clean and healthy environment. To strengthen this claim, the petitioner relies upon a series of judgments, statutory provisions, environmental regulations, and government notifications.

As the judges begin examining the matter, they are assisted by the Court's proprietary legal software, a system designed around the seamless integration of both local and external diligence. The role of the software is not to decide anything, but to quietly perform a task that would otherwise consume a significant amount of judicial time. Every judgment cited in the petition is automatically verified through local diligence against an internally maintained and regularly updated database of decisions of the High Court and the Supreme Court. This ensures that the precedents relied upon are accurate, current, and traceable to authenticated court records.

At the same time, the software performs external diligence by cross-verifying references to environmental statutes, rules, and regulatory frameworks through live integrations with official government websites that host updated legislations, delegated regulations, and policy notifications issued by various ministries. With each citation, whether judicial or statutory, the system provides an instant link to the original and authoritative source, allowing the judges to independently examine the material without delay or doubt.

What is important to understand here is that the judges do not delegate their judgment to the software. They read the statutes, examine the precedents, and apply their own experience and constitutional wisdom to the facts before them. *The technology merely removes the friction of verification*. It ensures that no time is lost

in locating documents and no error creeps in due to reliance on outdated or unauthenticated material. In this quiet but meaningful manner, the Human in the Loop principle comes alive. Technology works in the background, while human judgment remains firmly at the centre. The result is a process that feels more fluid, more focused, and ultimately more aligned with the larger goal of delivering timely and accurate justice.

The concepts discussed in the context of due diligence can also be effectively applied in the area of *compliance*. In compliance, the objective remains largely the same: to ensure that what is stated in a document is accurate, current, and legally correct. In simple terms, compliance involves checking whether the statements made in a document truly align with the laws, rules, notifications, and regulatory requirements that apply to an organisation. If a document asserts that a particular requirement has been complied with, that assertion must be traceable to an existing legal or regulatory provision that is presently in force.

AI-based tools assist in this process by comparing the text of compliance documents with internal policies, standard operating procedures, and prior regulatory filings, while simultaneously validating them against authoritative external sources such as official government portals and regulatory websites. This helps ensure that the document is not relying on outdated regulations or incorrect legal positions.

As a broader principle applicable across all legal AI use cases, it is important to recognise that complete automation is neither feasible nor desirable in the legal domain. The law operates in shades of interpretation, context, and human judgment. Consequently, many experts consider a 30 to 40 percent level of automation to be both realistic and optimal, as it significantly improves efficiency while ensuring that accountability and decision-making remain firmly in human hands.

4. Contract Redlining

The term "contract redlining" originates from a very literal practice that existed long before the advent of digital tools. Traditionally, when one party received a draft contract from a counterparty, the document would be printed and reviewed manually. Clauses or terms that were not acceptable to the receiving organisation were marked using a red coloured pen to indicate proposed additions, deletions, or changes. These markings served as a clear visual representation of the points on which the receiving party did not agree. Once the document was redlined, it would be sent back to the counterparty for their consideration. The counterparty would then review the suggested additions, deletions, or changes, accept certain modifications, reject others, and in some cases propose further revisions. This exchange of drafts would continue over multiple iterations until both parties were satisfied with the language of the contract. Only after this iterative process is concluded would the final version of the agreement be agreed upon and executed.

While the medium has evolved from physical documents and red pens to digital files and tracked changes, the underlying objective of redlining remains unaltered. Modern legal AI tools are helping organisations in their Contract Redlining process by using automation. The process is simple: You upload the 1st draft received by the Counter party in your legal AI tool and the tool easily redlines the Contract by comparing this Contract with your *pre-defined non-negotiable instances* (set of instructions which are approved for business use by the organisation). It is imperative to understand the possible operations that could be performed by this tool: *retaining* the language if it is in lines with the organisation's agreed upon principles, *deleting* the language which is not acceptable, *amending* the language if it deviates from the agreed upon positions or *adding* the language which is not present.

Each of these operations is designed to assist the legal professional by reducing manual effort, while ensuring that the final authority over the contract continues to rest with human judgment.

To understand how AI-assisted contract redlining works in practice, consider a simple example involving a Non-Disclosure Agreement (NDA). An organisation may define a limited set of non-negotiable instructions that reflect its standard legal and commercial positions. These instructions act as the reference point against which any incoming draft from a counterparty is reviewed.

Sample Instruction Set for an NDA:

- The confidentiality obligation shall apply mutually to both parties.
- The duration of confidentiality obligations shall not be less than three years from the date of disclosure.
- Liability for breach of confidentiality shall be capped at an amount not exceeding the fees payable under the agreement for the preceding twelve months, and must not be unlimited.
- Governing law and jurisdiction shall be India, with exclusive jurisdiction vested in Indian courts.

Now consider a draft NDA received from a counterparty that contains the following deviations:

- The confidentiality clause imposes obligations only on one party, while the counterparty is excluded from similar duties.
- The term of confidentiality is limited to twelve months.
- The liability clause imposes unlimited liability on the receiving party for any breach.

- The draft contract is silent on governing law and jurisdiction.

When this draft is uploaded into a contract redlining tool, the system compares the clauses in the counterparty's NDA against the organisation's approved instruction set. Based on this comparison, the tool performs a series of targeted redlining actions.

First, the confidentiality clause is flagged and amended to ensure that the obligation applies mutually to both parties, in line with the approved instruction. Second, the duration of the confidentiality obligation is identified as insufficient, and the clause is amended to reflect a minimum period of three years. Third, the liability clause is highlighted as non-compliant due to the presence of unlimited liability, and a suggested amendment is introduced to cap liability at an amount equivalent to the fees payable for the preceding twelve months, in accordance with the organisation's approved standard. Finally, since the governing law and jurisdiction clause is missing, the tool proposes the insertion of the organisation's approved clause specifying Indian law and exclusive Indian jurisdiction.

Throughout this process, the role of the contract redlining tool is limited to identifying deviations and proposing changes based on predefined instructions. The legal professional reviews each suggested amendment, assesses its commercial and legal implications, and decides whether to accept, modify, or reject the proposed changes. In this manner, contract redlining tools significantly reduce manual effort while ensuring that the final contract continues to reflect informed human judgment and organisational intent.

Version Control

Another challenge inherent in the traditional manual redlining process was the inevitable multiplication of documents arising from successive rounds of negotiation. Each exchange between the parties typically resulted in a new version of the contract, often saved and circulated as separate files. Over time, this led to the accumulation of numerous drafts, each containing incremental changes, comments, and tracked revisions. Managing these multiple versions not only increased the administrative burden on legal teams but also heightened the risk of confusion, duplication of effort, and inadvertent reliance on an incorrect or outdated draft.

AI-assisted contract redlining tools address this challenge by maintaining a single, structured record of all changes and iterations within a unified system. By consolidating negotiation history and highlighting accepted and pending changes, such tools reduce document clutter, minimise errors, and bring greater clarity and control to the contract negotiation process, while continuing to preserve human oversight at every stage.

5. Translation enabled application

Traditionally, legal translation has been a time-consuming and specialised task. Translating statutes, judgments, contracts, or regulatory documents required skilled human translators who not only understood the language but were also well-versed in legal terminology. While this ensured quality, it also meant that translation at scale was often impractical. As a result, access to legal information remained largely confined to those proficient in English or the language in which the document was originally drafted.

AI-based translation tools have significantly altered this landscape. Modern language models are capable of translating

complex legal text across multiple languages while preserving contextual meaning. Models such as **Claude Sonnet** have demonstrated strong capabilities in handling long-form documents, maintaining coherence, and respecting nuanced instructions. This makes them particularly suitable for legal translation, where fidelity to structure and meaning is critical. Such models can process judgments, pleadings, or statutory provisions and generate translations that are not merely literal but context-aware.

However, it is crucial to reiterate that translation in law is not a standalone mechanical exercise. A single mistranslated word can alter the meaning of an entire provision or misrepresent a judicial finding. This is where the Human in the Loop framework becomes indispensable. AI tools can perform the initial translation with remarkable efficiency, but the translated output must always be reviewed by a legally trained human translator or practitioner before being relied upon for official or professional use. In practice, this significantly reduces effort while retaining accountability.

That said, it is important to approach automation in translation with realism. In the legal domain, **100% automation is neither achievable nor desirable**. Language, especially legal language, carries nuance, intent, and interpretative depth that cannot be fully entrusted to machines. Most experts agree that **automating approximately 30–40% of the process** strikes the right balance in case of Legal automation. At this level, AI handles the most time-consuming and repetitive portions of the task, while human experts focus on review, refinement, and validation.

6. **Multi-Jurisdiction Statute Comparison:** An increasingly valuable capability within advanced Legal AI platforms is the ability to compare statutory provisions across multiple jurisdictions in a structured manner. Certain tools in the market

now enable end users to map differences in definitions, reporting obligations, and penalty frameworks across countries or states. This functionality is particularly critical for organisations and law firms engaged in cross-border financial transactions, where regulatory divergence can materially affect risk exposure, structuring decisions, and compliance strategy. Some tools even go the extra mile and send instant notifications to their users as and when a change in a particular law is made in their database.

7. **Speech to text based applications:** Another emerging interface layer within Legal AI involves speech to text based applications, where users can prompt, instruct, and use legal AI tools entirely through voice commands. Such systems have the potential to improve accessibility, enhance productivity during meetings or courtroom preparation, and reduce friction in drafting or research tasks. However, their effectiveness remains heavily dependent on the clarity and unambiguity of the input audio. Variations in accent, background noise, overlapping speech, or imprecise articulation can materially affect output accuracy, which in a legal context may lead to substantive misunderstandings. Accordingly, while voice enabled Legal AI represents a promising advancement, its reliability is contingent upon high quality and unambiguous sound input (therefore it must be used with caution).

8. Legal Analytics & Visualization

Legal analytics, in its simplest sense, is the segment of law which deals with identifying data points and patterns that already exist within legal documents. Judgments, contracts, pleadings, statutes, and regulatory materials are no longer viewed merely as text, but also as structured sources of *data* from which *information* can be churned. Legal analytics helps legal professionals understand

large volumes of material by bringing structure to what would otherwise remain scattered across multiple documents.

" Data is unprocessed raw input, while information is data that has been processed in a way that makes it useful for understanding and decision-making."

Visualisation is what makes this exercise practically useful. Once the information is identified, it can be presented in the form of charts, graphs, bars, or simple dashboards. This significantly reduces the cognitive load of the end user. Instead of manually scanning through dozens of documents, a legal professional can grasp key insights at a glance, while still retaining the ability to drill down into the underlying material whenever required.

A useful way to understand this is through contract analysis. Imagine an organisation that has uploaded hundreds of contracts into its legal AI system, spanning multiple jurisdictions. A legal analytics dashboard can visually display how these contracts are distributed across different jurisdictions, along with the frequency of certain clauses or governing law provisions. At a glance, the legal team can see which jurisdictions account for the highest volume of contracts and where most negotiation effort is being spent. This information can directly support better resource allocation decisions, such as identifying the need for specialised regional expertise or prioritising review efforts in high-volume jurisdictions.

An important feature of such dashboards is multi-preview capability. The visuals shown are not abstract summaries. Each graph or chart is actively linked to the specific contracts and clauses from which the data has been drawn. A single visual may be based on multiple uploaded agreements, yet every data point remains traceable to its original source through clickable references. This ensures that the legal team can instantly verify

the information and move seamlessly from a high-level view to the exact contractual language when required.

This approach strikes a balance between efficiency and trustworthiness. Visualisation simplifies understanding, while source-linked citation preserves accuracy and accountability. As with all legal AI applications, the Human in the Loop principle remains intact. The system organises and presents information, but interpretation, strategic decision-making, and final reliance continues to rest with the legal professional.

Chapter 3

Design-Based Augmentation

The discussion so far has focused on the role of Artificial Intelligence as a Legal Assistant, where technology functions as a support system for legal professionals. In this capacity, AI has demonstrated its ability to assist in tasks such as Legal Research, Drafting, Due Diligence, Contract Redlining and Visualisation based on analytics. Across each of these use cases, the underlying theme remains consistent: AI enhances efficiency by automating repetitive and mechanical aspects of legal work, while interpretation, strategy, and accountability continues to rest with humans.

Having examined how AI can assist lawyers at the level of individual legal tasks, it is now important to shift focus to a different but equally significant dimension of legal technology: Design-Based Augmentation.

"Design-based augmentation is a process that focuses on enhancing the overall efficiency of the legal process by thoughtfully designing the most effective workflows. Let us understand this concept using some practical use cases."

1. MS Word and Google Docs plugin

From a practitioner's perspective, this is where design-based augmentation starts to feel genuinely useful rather than

theoretical. Most legal work still begins and ends inside familiar tools like MS Word and Google Docs. Drafting contracts, reviewing agreements, preparing advisory, or redlining documents all happen here. Design-based augmentation recognises this reality and builds intelligence into these very environments instead of asking lawyers to shift to entirely new platforms.

When a contract is opened in MS Word or Google Docs with a legal AI plugin enabled, the experience remains familiar. The lawyer continues to read, edit, and revise the document in the usual manner. The difference is that intelligent assistance now operates quietly in the background. Within this familiar setting, **legal research** becomes far more fluid. While drafting or reviewing a document, relevant case laws, statutory provisions, or regulatory references can be accessed contextually, without leaving the document. The lawyer remains focused on the language being written or reviewed, while the supporting legal material becomes available exactly where it is needed, along with pertinent source citations.

Similarly, when a user **drafts** within MS word or Google docs, the plugin can provide real-time suggestions for augmenting the drafting process without requiring to switch to another interface. It may suggest alternative phrasing based on its library of pre-approved clauses, highlight key missing elements, or even offer ideas to strengthen an argument based on the context of what is being written. For more complex documents, the tool can generate a first-level draft or a skeletal structure to help the lawyer get started. Importantly, these suggestions remain optional and advisory in nature. The end user retains full control over the language, using the assistance only as a starting point or a source of inspiration, while the final expression remains a product of professional judgment.

Due diligence also benefits from this approach. References, defined terms, and factual statements mentioned in the document can be validated instantly, whether against internal repositories or reliable external sources. Any inconsistency, outdated reference, or missing linkage is flagged at the point where it appears, allowing the lawyer to address it immediately without breaking concentration.

Additionally, **redlining** can also occur within the same environment. Counterparty drafts can be assessed against pre-approved organisational positions, and deviations are highlighted directly within the document. Suggested changes appear alongside the existing text, enabling the user to review, accept, modify, or reject them using professional judgment.

By allowing research, drafting, diligence, and contract review to take place entirely within MS Word and Google Docs, one single plugin can remove all unnecessary frictions. The technology adapts to how lawyers already work, ensuring that efficiency is improved without altering familiar tools or established working habits.

2. Intelligent Storage

When legal work begins to scale across multiple matters, clients, and timelines, the way documents are stored becomes as critical as how they are drafted or reviewed. Intelligent storage is a natural application of design-based augmentation, where documents are not merely saved but are **organised with intent**, in a manner that reflects real legal practice. Before diving deep into this important concept, one should know that this application is only aimed at one thing in particular, that is, enhancing efficiency of legal work so that the predominant time spent by a practitioner is utilised in strategic legal work. I will take

the liberty of extrapolating this and terming the entire process of 'design based augmentation' as 'Justice by design'.

" **Justice by Design** is the conscious effort to structure legal systems, processes, and workflows in a manner that promotes efficiency without compromising fairness or accountability. It draws from the long-standing jurisprudential insight, articulated by thinkers such as Roscoe Pound, that efficiency is not antithetical to justice but is, in fact, one of its core pillars. A justice delivery mechanism that is slow or procedurally burdensome ultimately fails the very purpose it seeks to serve."

At its core, intelligent storage moves beyond static folders and file names. Instead, it automatically groups related documents based on their context and content. For instance, a contract, its subsequent amendments, related correspondence, and supporting regulatory filings are automatically linked together as part of a single document family. This contextual grouping allows a lawyer to view all legally connected documents together, without manually creating or maintaining complex folder hierarchies.

A practical illustration of intelligent storage can be seen in the context of a Master Services Agreement (MSA), which is rarely a standalone document in real-world practice. An MSA is often accompanied by multiple Statements of Work (SOWs), amendments, change orders, and exhibits that evolve over the course of a long-term commercial relationship. In a traditional storage setup, these documents are frequently scattered across folders, renamed inconsistently, or saved by different teams, making it difficult to reconstruct the contractual position at any given point in time.

An intelligent storage system addresses this problem by automatically organising all related documents into a **single,**

coherent document family. The original MSA acts as the *anchor document*, while subsequent SOWs, amendments, and exhibits are contextually linked to it. For instance, if an SOW is amended multiple times, each amendment is automatically associated with the specific SOW it modifies, rather than being stored as an isolated file. Similarly, exhibits that apply across multiple SOWs are grouped accordingly, allowing a lawyer to instantly understand how they fit within the larger contractual framework.

From the end user's perspective, this means that with a single click on the MSA, they can view the entire contractual ecosystem. They can see the base agreement, identify which SOWs are currently active, track how individual clauses have evolved through amendments, and access all supporting exhibits without having to manually search across folders. At every step, the underlying documents remain fully visible and accessible, ensuring that the system's organisation can be independently verified.

This form of intelligent organisation not only saves time but also reduces the risk of oversight, particularly in scenarios involving renewals, audits, or disputes. By presenting complex contractual relationships in a structured and intuitive manner, intelligent storage allows legal *professionals to focus on interpretation and strategy, rather than on the mechanics of organising documents.*

Now considering the application of this in courtrooms, one can easily imagine a digitally enabled judicial environment where judges, while dealing with a case, have seamless access to a well-organised repository containing all relevant case files along with their supporting documents. They no longer need to navigate through voluminous paper files or fragmented digital folders. Every pleading, annexure, order, and connected document is logically grouped and readily accessible within a clearly structured

digital workspace. The judge can search within documents, move across related filings, and verify references with just a few clicks.

The time savings in such a system are substantial. Tasks that would otherwise require repeated assistance from court staffs or lead to avoidable adjournments can be handled independently and efficiently by the judge. Importantly, this does not dilute existing safeguards. Even though the documents would already have been reviewed and verified by the registry staff during uploads, when a Judge is able to control the digital interface then it adds an additional layer of accountability and confidence. The judge remains in complete control of the material before the court, while *technology quietly removes procedural friction*. In this manner, intelligent design strengthens the justice delivery process by enabling faster, more reliable, and well-informed judicial decision making without interfering with judicial discretion.

Once documents are intelligently stored, organised, and made easily searchable, the next logical step is *Case Management*. An intelligent repository lays the foundation, but case management is where this organised information is put to active use within live legal matters.

3. Case Management using Legal AI

Case management builds directly upon an intelligent document repository by bringing together all case-related materials into a single, unified interface. Instead of merely storing documents, the system contextualises them within a specific matter. Pleadings, affidavits, annexures, correspondence, court orders, research notes, and internal drafts are automatically associated with the relevant case, ensuring that every document is available at the right place and at the right time.

Legal AI-enabled case management systems further enhance efficiency by tracking procedural timelines and case milestones.

Hearing dates, filing deadlines, compliance timelines, and pending actions are identified directly from case documents and court orders, allowing legal teams to stay organised without relying on manual reminders or fragmented tracking methods.

Advanced search capabilities allow practitioners to retrieve information based on parties, issues, statutory provisions, or specific legal questions, rather than depending on file names or folder structures. This is particularly valuable in complex litigation where matters span multiple years and involve voluminous records.

Crucially, while the system assists in organising and presenting information, all strategic and legal decisions continue to be taken by lawyers and judges. The technology merely ensures that the relevant information is structured, accessible, and reliable. In this sense, case management represents the practical culmination of intelligent storage, transforming a well-organised repository into a living, functional system that actively supports the efficient administration of justice.

Example:

Consider a sensitive matrimonial dispute pending before a Family Court, where the case file contains not only pleadings but also highly sensitive personal data within the exhibits. These documents may include medical records, counselling reports, financial disclosures, bank statements, and details relating to *minor children*. In such cases, merely linking all documents to a case number is insufficient. While procedural efficiency demands that every relevant document be available within the case file, *principles of privacy and data protection* require that access to certain documents be strictly limited.

An intelligent Legal AI enabled case management system addresses this challenge by embedding **access controls** directly

into the design of the repository. Documents containing critical personal data are automatically classified as sensitive and are made visible *only to authorised decision making authorities* such as judges, court appointed officers, or specifically permitted legal representatives. Other users who may have access to the case file for procedural purposes can view only those documents that they are legally entitled to see. In this way, efficiency is achieved without compromising confidentiality.

At a technical level, this is supported through encryption and decryption mechanisms. **Encryption** refers to the process of converting readable information into a coded form that cannot be understood without proper authorisation. **Decryption** is the reverse process, where the coded information is converted back into its original, readable form, but only by users who possess the appropriate permissions or keys. This ensures that even if data is stored or transmitted digitally, it remains protected from unauthorised access.

By combining intelligent document linkage, role-based access control, and strong encryption standards, modern case management systems ensure that privacy is not treated as an afterthought. Instead, it becomes an integral part of the justice delivery framework. Such design led safeguards allow courts and legal professionals to work faster and more efficiently, while continuing to uphold the fundamental duty to protect personal data and maintain public trust in the legal system.

Having understood the application of case management in litigation, the same mechanism can be used throughout various regulatory bodies and quasi-judicial setups for enhancing the overall process of justice and delivery of services to the public.

Building upon the principles of intelligent case management, the same design philosophy naturally extends into the commercial

sphere in the form of **Contract Lifecycle Management (CLM)**. If case management focuses on organising and safeguarding information, CLM applies similar intelligence to the entire life of a contract, from its inception to its termination. In many ways, contracts are the commercial lifelines of organisations, and managing them efficiently is as critical as managing cases in a courtroom.

The journey of a contract within an organisation is not confined to drafting and signing alone. A well-designed CLM framework views a contract as a living instrument, one that moves through clearly defined stages, each requiring legal oversight, operational clarity, and accountability. A contract, much like a case file, passes through a series of interconnected stages. When viewed through a *design based augmentation* lens, these stages can be streamlined into a coherent and manageable lifecycle, mentioned below:

A. Contract Initiation and Draft Assembly

Once a request is raised, the system automatically prepares a first draft of the contract using the relevant templates and clauses. Inputs such as Jurisdiction, nature of services, pricing structure, and counterparty details guide the system in generating a first draft by pulling relevant templates and clauses. At this stage, CLM systems often integrate with procurement, CRM, or finance tools, allowing commercial inputs to flow directly into the legal document.

B. Negotiation and Redlining

Draft contracts then move into negotiation. AI-enabled redlining tools analyse counterparty changes, flag deviations from approved positions, and suggest alternative language aligned with organisational standards. This helps legal teams negotiate

efficiently while maintaining control over risk and consistency across contracts.

C. Legal Review and Approvals

Negotiated drafts undergo legal and commercial review, where key terms such as pricing, liability, indemnities, and termination rights are extracted and assessed. Approval workflows are triggered based on predefined thresholds, ensuring that the right stakeholders sign off at the right stage. Every approval is logged, creating a clear audit trail.

D. Execution and Formalisation

Once approvals are secured, the contract is executed. CLM systems integrate with electronic signature platforms, enabling seamless signing, tracking, and storage of the final agreement within a single system of record.

E. Contract Enablement and Performance Monitoring

Post-execution, the contract shifts from a static document to an operational instrument. Key obligations, milestones, and responsibilities are made visible to relevant internal teams. AI-driven monitoring helps track compliance, deadlines, and performance metrics, reducing the risk of oversight or breach.

F. Renewal, Termination, and Closure

As contracts approach critical milestones, the system generates timely alerts and insights. Based on performance history and business impact, organisations can make informed decisions to renew, renegotiate, or terminate the agreement. This ensures that contracts conclude in a structured and deliberate manner.

This **above model** keeps the lifecycle intuitive, avoids unnecessary fragmentation, and aligns closely with how legal and business teams actually operate, while still reflecting the full breadth of Contract Lifecycle Management.

E-Signatures

Before moving to the next chapter, let us quickly understand the legal grounds of Electronic Signature or E-Sign. Electronic signatures have now acquired clear legal recognition across major jurisdictions, making them a reliable and enforceable method of contract execution. Our focus will be to understand their legal basis across the U.S, Europe and Indian jurisdiction.

In the **United States**, electronic signatures are recognised under the *Electronic Signatures in Global and National Commerce Act* **(E-SIGN Act)** and the *Uniform Electronic Transactions Act* **(UETA)**. Together, these laws establish that an electronic signature cannot be denied legal effect solely because it is in electronic form, provided that the parties have consented to transact electronically and the signature can be attributed to the signatory.

In **Europe**, electronic signatures derive their legal validity from the **eIDAS Regulation**. The regulation categorises electronic signatures into different levels, namely simple, advanced, and qualified electronic signatures. Qualified electronic signatures, which meet the highest standards of identity verification and security, carry the same legal effect as handwritten signatures across all EU member states. This framework has brought uniformity and cross-border recognition to electronic contracting within Europe.

In **India**, electronic signatures are legally recognised under the **Information Technology Act, 2000**. The Act provides validity to electronic records and electronic signatures, including digital signatures based on cryptographic authentication. Courts in India

have consistently upheld electronically executed contracts, provided the requirements of authenticity, integrity, and intention to sign are satisfied. With the increasing use of Aadhaar-based authentication and certified digital signature providers, electronic execution has become both practical and legally robust.

Taken together, these frameworks make it clear that electronic signatures are no longer a technological convenience but a legally accepted mode of execution. When integrated into CLM systems, electronic signatures not only accelerate contract execution but also ensure enforceability across jurisdictions, provided statutory requirements are complied with and human oversight is maintained.

Chapter 4

What Legal AI cannot do

Having understood the capabilities of Legal AI tools, it is equally important to understand their structural limitations. I would go to the extent of calling this the second most important chapter of the book (first being chapter 1: human in the loop). The motivation behind this chapter is simple yet crucial: to clear the air surrounding the integration of Artificial Intelligence with law. In recent times, there has been no shortage of exaggerated claims and fear-driven narratives suggesting that AI will soon replace judges, lawyers, or the justice delivery system itself. Most of these claims are not supported by sound technical understanding or credible legal literature.

From a policy making perspective as well, clarity on what AI can and cannot do is indispensable. Regulatory decisions taken without a clear appreciation of technological limits risk either stifling innovation or enabling irresponsible use. It is therefore essential to draw firm conceptual boundaries. In this chapter, I attempt to identify four such glass ceilings in the domain of Legal AI, limitations that cannot be crossed through logic.

1. Predictive Analysis

As we speak there are proprietary softwares out there in the market claiming to have the capabilities of predicting the outcomes of a case or an ongoing negotiation, which I think is

nothing short of a farce. The reason I am using such strong words against predictive analysis is because law is not an isolated subject and every single stage where law is practically applied takes into consideration various socio-political factors, larger good of the society and prevailing public policy, among other nuances . The world already has too many astrologers who claim to predict the future, let us keep law out of these prediction games. The role of Legal AI is only to assist, inform, and augment human decision-making, not to indulge in predictions that undermine the seriousness of legal reasoning. Any technology that claims otherwise not only misunderstands the nature of law but also misleads its users into placing unwarranted faith in algorithmic outcomes.

That being said, supporters of predictive analysis often cite '*civil law based jurisdictions*' as a supposed exception. Since judges in these systems operate within tightly written statutes and have almost negligible scope to interpret the law, decisions tend to follow fixed rules and are therefore assumed to be easier to predict. On the surface, this line of reasoning sounds convincing, and it is easy to see why many find it appealing. But this view rests on a very convenient oversimplification which is certainly a construct of the *Sales teams* of various legal AI softwares to sell their respective products. Let me dig in a little deeper and bust this myth. Even in civil law systems, judicial decision-making cannot be reduced to a purely mechanical exercise of applying statutory provisions. While judges may operate within narrower interpretative boundaries than their common law counterparts, those boundaries still leave room for application of mind. Judges must assess facts, weigh evidence, consider proportionality, and apply statutory rules in a manner that aligns with their constitutional values and prevailing public policy. Each of these steps requires human reasoning, and with human reasoning comes discretion, whether acknowledged or not. The presence of

codified statutes may reduce variance at the margins, but it does not eliminate the deeper contextual forces that shape legal decision making. Legal outcomes are still contingent on human judgment, institutional priorities, and societal values, all of which require broad interpretation.

At best, predictive analysis can help us understand what has happened in the past by *identifying patterns in historical data.* The moment it starts claiming that it can tell us what will happen in a future legal dispute, it stops playing the role of an assistant and turns into speculation, regardless of whether the system is based on common law or civil law. Technology can assist by organising information, identifying patterns, and reducing mechanical effort. What it cannot and should not replace is the discretionary judgment that lies at the heart of justice. The value of law lies not in its ability to predict outcomes, but in its capacity to respond thoughtfully to the complexities of human life.

2. Handling Edge cases

Legal AI systems, by their very design, are trained on historical data. They learn from past judgments, statutes, prior contracts, and documented legal outcomes, among other legal data forms. Their strength lies in recognising patterns within what has already happened. However, this very strength becomes a limitation when the system is confronted with an unprecedented scenario. When there is no comparable precedent, no settled interpretation, and no historical analogue to rely upon, the system has little meaningful guidance on how to respond.

This limitation becomes particularly evident when the law encounters what are often described as "Black Swan" events. A Black Swan event is a situation the law did not anticipate, where established rules, precedents, or contractual outlines suddenly prove inadequate, forcing decision-makers to rely heavily on

judicial discretion, interpreting public policy and constitutional principles rather than settled doctrine alone. In the legal world, such events are not anomalies confined to theory. They emerge during pandemics, financial crises, sudden regulatory overhauls, technological disruptions, or periods of intense social transformation. These moments reveal the outer limits of the legal framework, where existing rules were not contemplated to apply and offer definitive guidance.

Legal AI tools fundamentally operate on a data-based approach, not a first-principles approach. When confronted with a situation that has not previously appeared in their data, the system lacks the capacity to truly understand the problem. At best, it attempts to approximate an answer by drawing analogies from loosely related historical material, leading to hallucinations. This can create a dangerous illusion of certainty, where confident-sounding outputs conceal the absence of genuine contextual, moral, or constitutional reasoning.

"**First principle approach in law**: It is a method of reasoning that breaks a problem down to its most fundamental truths and builds conclusions from those core principles, rather than relying on precedent, convention, or assumptions. In essence, instead of looking backwards and asking how a similar issue was dealt with in the past, the court looks inward to the core values the law is meant to safeguard and reasons forward to reach a just outcome."

Edge cases in law often demand more than pattern recognition. They require judges, lawyers, and regulators to interpret legislative intent, balance competing constitutional values, and respond to real-world consequences that extend far beyond the text of a precedent or statute. These decisions are shaped by social realities, public policy considerations, and moral judgment factors that cannot be reliably encoded into training data.

The COVID-19 pandemic offers a clear illustration. Courts were suddenly confronted with questions around lockdowns, digital access to justice, and the tension between public health and individual freedoms. Much of the legal reasoning during this period was developed in real time, informed by evolving facts and societal priorities. No historical dataset could have fully anticipated these challenges, nor could any Legal AI system have resolved them autonomously (at best they could have given some suggestions).

This is why handling edge cases remains a firm boundary for Legal AI. The technology excels when the path is already mapped. When the law is forced to navigate into uncharted territories, it is human judgment, discretion, and responsibility that must lead. Expecting Legal AI tools to perform otherwise is not an advancement of justice, but a misunderstanding of both law and technology.

3. Interpretation

I remember when I was sitting in my 'Interpretation of statutes' class, a revered professor wrote in bold letters on the blackboard: "*Law is a game of Interpretation*". Everyone in class was startled and started looking at each other's blank faces. We could all sense that the professor was enjoying this phase of the class and it sure looked like a rehearsed act of teaching. After a couple of minutes, the professor further added to the buildup and told us that the class was over and asked us all to ponder over this. Strangely enough, it turned out to be one of the most memorable and impactful classes of my entire law school journey. With time, experience, and practice, the meaning of that sentence became increasingly clear. The professor was absolutely right. Law, at its core, is indeed a game of interpretation.

Interpreting law is one of the most human-centric functions within the legal ecosystem, and it represents another clear boundary that Legal AI tools cannot cross. Law is not merely a collection of words arranged in statutes or judgments. It is a living system of meaning, shaped by context, intent, history, and societal values. Interpretation is the process through which static legal text is translated into practical outcomes, and this process is deeply rooted in human reasoning.

In legal terms, **interpretation** refers to the process by which courts and legal authorities ascertain the meaning, scope, and intent of statutory provisions, constitutional texts, or legal instruments in order to apply them to specific facts and circumstances. It involves examining the language of the law, its context, purpose, and underlying principles so that the rule can be given effect in a manner consistent with legislative intent and the broader objectives of justice. Prima facie, it may appear that interpretation is a minor or mechanical exercise, something that merely fills gaps in the written law. In reality, the opposite is true. Interpretation lies at the very heart of legal decision-making. It has the power to reshape viewpoints, alter judicial outcomes, and even pave the way for legal and social reform. The same statutory provision, when interpreted differently, can lead to entirely different consequences. This is precisely why interpretation is not a peripheral skill, but one of the most consequential functions performed by courts.

Let us understand this through a practical example of **granting bail**.

At a surface level, bail provisions appear straightforward. Statutes often prescribe conditions, exceptions, and limitations under which bail may or may not be granted. However, the real decision does not flow merely from reading the section. Courts must interpret phrases such as "reasonable grounds," "likelihood of

absconding," "tampering with evidence," or "interest of justice." These are not mathematical expressions. They are open-textured concepts that demand judicial discretion.

In one case, a strict interpretation may lead a court to deny bail, emphasising the gravity of the offence or the statutory embargo placed on release. In another case, the same provision may be interpreted liberally, taking into account factors such as prolonged incarceration without trial, the health and age of the accused, or the constitutional mandate of personal liberty under Article 21. The statutory text remains unchanged, yet the outcome differs significantly because interpretation bridges the gap between law on paper and justice in practice.

This exercise cannot be reduced to pattern matching or statistical inference. It requires a judge to balance competing interests, weigh human consequences, and apply constitutional values to the facts at hand. Bail jurisprudence, therefore, serves as a clear illustration of why interpretation is inherently human. It is not about what the law says in isolation, but about what the law means in a given context. This is precisely the space where Legal AI can assist by providing information, but can never replace the interpretative judgment that lies at the core of judicial reasoning.

At best, a Legal AI tool can assist by presenting multiple possible interpretations that have emerged in past cases or academic discourse. It may highlight how similar provisions have been read in different contexts or point out divergent judicial approaches over time. However, the ultimate choice of interpretation, or even the decision to depart from all suggested interpretations and craft a new one, remains entirely within the domain of the Judge. That decision is shaped by judicial experience, constitutional conscience, and an understanding of societal realities. In this sense, Legal AI can serve as a reference point in the interpretative

exercise, but it can never assume the role of the human decision-maker who ultimately breathes life into law.

4. Strategic thinking

The word *strategy* traces its roots in the Greek term '*strategos*', which is referred to as the art of military leadership and generalship. At its core, strategy was never about rigid rules or fixed plans; it was about making informed choices under conditions of uncertainty, often with limited resources and evolving circumstances.

Another defining feature of strategy, particularly in the legal context, is that it is a living and evolving tool. Strategy is never static. It continuously responds to external variables such as shifting facts, intentional or unintentional errors, changes in public policy, regulatory interventions, opposing party conduct, and other unforeseeable factors. A legal strategy that appears sound at the time of filing may require recalibration midway through the proceedings, and in many cases, the most effective strategic decisions are taken in response to developments that could not have been anticipated at the inception. This adaptive quality is what makes strategy a *living tool* rather than a static blueprint.

This fluid nature of strategy stands in sharp contrast to how Legal AI tools operate. Legal AI systems function on pre-existing information. They analyse historical data, previously decided cases, established procedural patterns, and known legal positions. While this enables them to assist with research, drafting, and pattern recognition, it also confines them to what is already known and documented. They do not possess the ability to sense shifting dynamics in a courtroom, interpret subtle changes in judicial temperament, or reassess priorities based on emerging realities.

In practical terms, a human lawyer or judge can easily manoeuvre when an unexpected situation unfolds, when a witness testimony alters the narrative, or when external circumstances demand a different legal posture. Legal AI, on the other hand, can only update its outputs when new data is formally introduced into its system. *It reacts after the fact, whereas human strategy anticipates, adapts, and sometimes even pre-empts change.*

This distinction is crucial. Strategy thrives on uncertainty and adapts to it. Legal AI, by its very design, seeks certainty by relying on past information. Expecting a data-driven tool to replace a living, adaptive, and context-sensitive process is to misunderstand both the nature of strategy and the role technology is meant to play in the legal ecosystem. AI can support strategic thinking, but the act of strategy itself must remain firmly human.

This is precisely where the limitation of Legal AI becomes evident. Legal AI tools operate on pre-existing information and historical data. They can assist by highlighting past patterns or suggesting conventional approaches, but they cannot genuinely adapt strategy in real time as human actors do. Strategy, in its truest sense, requires continuous assessment, intuition, and judgment in the face of uncertainty which are qualities that remain firmly within the human domain.

Chapter 5

Ethical AI in law

Ethics is the foundational space that allows any new technology to grow in a responsible and sustainable manner. Whenever a new technology enters society, it naturally brings with itself a set of apprehensions and debates surrounding its possible misuse. These concerns are not misplaced. In fact, they serve an important purpose. Without well thought out guardrails, even the most powerful innovations can lead to unintended consequences and systemic harm.

Many governments across the globe have recently raised serious concerns over certain controversial features made available to users of a renowned Chatbot. One such feature, which drew widespread criticism, appeared to operate in complete disregard to individual privacy. It reportedly allowed users to generate obscene and, at times, lascivious images of other individuals without consent (as usual the target of this mischief was mostly females). This incident serves as a stark reminder of how powerful AI systems, when released without adequate ethical safeguards, can quickly cross the line from innovation into misuse, exposing deep vulnerabilities in privacy, dignity, and personal autonomy. The organisation soon realised its mistake and made a public apology. In technological terms this can be described as a case where **output moderation** was either inadequate or entirely absent. **Output moderation** is the process

of filtering and controlling AI-generated responses so that the output before being made publicly available goes through a series of virtual filters, which ensures that the final output as visible to the end user is safe, sound and in compliance with the prevailing public policy.

"Strong *output moderation* acts as the first line of defence against misuse, ensuring that technological power does not translate into social harm. The absence of such safeguards transforms AI from a tool of augmentation into a source of legal and moral liability."

In the context of Legal AI, ethics cannot remain an abstract ideal. It must translate into clearly identifiable principles that guide how these systems are designed, deployed, and used. Broadly speaking, ethical Legal AI rests on five core tenets, each of which addresses a distinct but interconnected concern:

- Confidentiality of data
- Bias
- Accountability
- Trustworthiness
- Transparency

i. Confidentiality: It sits at the very heart of the legal profession. Lawyers are entrusted with some of the most sensitive information imaginable, ranging from personal details to strategic legal positions. When AI tools are made aware of such information, the obligation to protect client data does not dilute but it multiplies. Ethical Legal AI systems must ensure robust data protection through encryption, access controls, and strict data usage policies. Client information should not be reused for training models without explicit consent, nor should it be exposed to unauthorised parties. Any breach of confidentiality

undermines not only individual cases, but also public trust in the legal system as a whole, as it constitutes a clear violation of attorney-client privilege and a serious breach of data privacy obligations.

In the context of Ethical Legal AI, the concepts of encryption and access control forms the practical backbone of confidentiality and data protection. Each addresses a different layer of risk within the legal technology ecosystem.

Encryption

Encryption refers to the process of converting data into a coded format that can only be read or accessed by authorised parties who possess the correct decryption key. In Legal AI systems, encryption ensures that sensitive legal data remains protected both when it is stored and when it is being transmitted. Even if data is somehow intercepted or accessed without permissions, encryption renders it unreadable and unusable. This is particularly critical in legal workflows, where documents often contain privileged communications, personal data, and strategic legal information.

For instance, imagine that in a set of confidential legal documents *all vowels are replaced with numbers*: *a* is replaced by *9*, *e* by *3*, *i* by *8*, *o* by *6*, and *u* by 5. Once this is done, the document no longer makes sense to a random reader. Even if someone somehow gains access to these files, they will not be able to understand what is written because they do not know the conversion rule. The conversion rule itself is the "encryption key".

In practical terms, this means that merely accessing a file is not enough. Without the key, the information remains protected. Modern Legal AI systems use far more advanced digital methods, but the idea is exactly the same-confidential legal data remains unreadable to anyone who is not supposed to see it.

Access Controls

Access controls are simply rules that decide *who can see what* inside an organisation. They recognise that not everyone needs access to every document, even within the same company or legal team. Take a typical organisational structure. The **C-suite executives** are usually privy to all documents, including the most sensitive strategic and legal materials. **Senior management** may have access to almost all data, but certain highly confidential documents such as sensitive negotiations, whistle-blower reports, or privileged legal opinions may still be restricted. At the **associate level**, access is even more limited: associates can only view documents that are directly relevant to the matters they are working on.

This layered access is not about distrust, it is about responsibility and risk management. Access controls ensure that confidential legal information is only available to those who are legally and professionally permitted to see it (permitted disclosure clause within the larger ambit of the Confidentiality clause). They prevent accidental disclosures, reduce the chances of internal misuse, and protect privileged material from being viewed by unauthorised persons. In the context of Legal AI tools, strong access controls are essential. Even if all data is stored securely, the system must still ensure that each user only sees what their role allows, nothing more, nothing less.

ii. Bias

At the core of every Legal AI system lies its training data. These systems learn by analysing large volumes of past judgments, statutes, contracts, and legal outcomes. They do not understand law in the human sense, they learn patterns from the inputs given to them. When this training data is biased, incomplete, or

historically skewed, the system absorbs those distortions as if they were neutral truths.

In simple terms, biased training data leads to biased outcomes. The principle is straightforward: **biased inputs produce biased outputs**. Legal AI tools cannot correct for unfairness embedded in the data unless humans consciously intervene. Without careful curation of training datasets and ongoing evaluation of system outputs, AI risks reinforcing existing inequalities rather than mitigating them.

Explicit Bias

Explicit bias refers to bias that is **conscious, intentional, and openly held**. It occurs when an individual or institution knowingly treats certain people or outcomes more favourably or more harshly based on identifiable characteristics such as race, gender, economic status, political belief, or social background. In law, explicit bias may surface through discriminatory policies, overtly prejudiced reasoning, or deliberate exclusion.

In the context of Legal AI, explicit bias can enter the system through **design choices**. For example, if a dataset is deliberately curated to exclude certain categories of cases, or if a tool is configured to prioritise one type of outcome over another to serve institutional convenience, the bias is no longer accidental, it is embedded by intent. Such bias is ethically indefensible and legally dangerous, as it undermines the rule of law and exposes institutions to serious accountability risks.

Implicit Bias

Implicit bias is far more subtle and, arguably, more dangerous. It refers to **unconscious attitudes or assumptions** that influence decisions without the decision-maker being aware of them. Judges, lawyers, policymakers, and data curators may genuinely

believe they are acting neutrally, while their choices are quietly shaped by social conditioning, professional culture, or past experiences.

Legal AI systems are particularly vulnerable to implicit bias because they are trained on historical data. If past judgments, policing patterns, sentencing trends, or contractual practices reflect unconscious preferences or systemic inequalities, the AI will learn those patterns as if they were neutral truths. The system does not understand *why* the data looks the way it does; it merely learns that "this is how things usually happen."

The danger here is that implicit bias, once encoded into an AI system, gains an appearance of objectivity. Outputs are presented as data-driven or statistically grounded, which can make biased results harder to question and easier to accept. This creates a false sense of neutrality, where human prejudice is repackaged as algorithmic logic.

Solution

One effective way to address bias in Legal AI systems is through the **RLHF (**often referred to as **Reinforcement Learning from Human Feedback)** approach. Under this method, humans remain actively involved in reviewing and refining the data before it is used as input for training the Legal AI tool. This allows potential explicit biases such as those based on gender, race, caste, creed, or other social markers to be identified and corrected at an early stage. Rather than allowing the system to blindly learn from historical material, human reviewers apply judgment and contextual understanding to ensure that problematic patterns are not absorbed as neutral truths. In the legal domain, this human oversight is crucial, as it introduces ethical awareness and professional responsibility into the training process. While RLHF cannot completely eliminate bias, it serves as an important

safeguard, helping ensure that Legal AI systems reflect deliberate human values rather than unintentionally amplifying historical inequities. It is also important to note that implicit biases are not covered in this process, since they are very subtle in their nature and can easily escape the RLHF process.

iii. **Accountability**: As discussed in Chapter 1 of this book, *"Human in the Loop"*, accountability is essentially about answering a very simple but critical question: *where does the buck stop?* In the legal ecosystem, accountability has never been abstract. Every opinion, filing, argument, or decision making ultimately traces its origin back to a human that can be questioned, challenged, and held responsible.

Legal AI tools, regardless of how advanced they may appear, cannot occupy this position. They do not possess the legal personality, moral agency, or the capacity to bear consequences. Their role is strictly that of an assistant, one that efficiently processes information, discovers patterns, and offers legal outputs as suggestions. They do not make decisions; they only support decision making. Consequently, assigning legal liability to an AI system for its outputs is both legally unsound and conceptually flawed.

Accountability, therefore, must always rest with the human at the other end of the spectrum, such as the lawyer, judge, or other legal professionals who choose to rely, modify, or reject the output generated by these tools. It is this human actor who applies their judgment, weighs context, and understands the real-world consequences of legal action. Legal AI can inform and augment human reasoning, but it cannot replace responsibility. Treating AI outputs as authoritative without human scrutiny risks diluting accountability, which is one of the foundational pillars of the legal profession.

Accountability also plays a crucial ethical role. Knowing that responsibility cannot be outsourced encourages caution, verification, and professional judgment. It prevents blind reliance on algorithmic outputs and reinforces the duty of care owed to clients, courts, and society at large. Without this clarity, there is a real risk that AI becomes a convenient scapegoat used to justify poor decisions under the guise of technological objectivity. Any attempt to treat AI as a decision-maker rather than a decision-support tool is not progress; it is an erosion of responsibility. The law does not and must not permit accountability to be automated.

iv. Trustworthiness: In Legal AI, trustworthiness is not a theoretical concept, it is the backbone on which the entire system rests. For lawyers and judges, the real question is whether an AI-generated output can be safely relied upon or not. It is not enough for an answer to sound convincing or well-written, it could even be hallucinated data (when an AI tool produces a confident-sounding answer that is factually incorrect). In law, an output is trustworthy only if it is accurate, verifiable, and capable of being independently examined. An untrustworthy output is not just unhelpful but it can actively mislead legal decision-making.

A key determinant of output trustworthiness is '**provenance**', which is the ability to clearly trace each part of the AI's response back to an authoritative legal source. When a Legal AI tool generates a conclusion, extracts a clause, or summarises a topic, the user must be able to see exactly which statute, judgment, paragraph, or clause supports that output. Trust arises not from the fluency of the response, but from its grounding in verified legal data.

Trustworthy outputs allow lawyers and judges to verify, cross-check, and independently interpret the source text. This ensures that the AI's response is not taken at face value but used only as

an assistive layer that remains open to scrutiny. Without this capability, outputs risk becoming black-box assertions that cannot be confidently relied upon in pleadings, opinions, or judicial reasoning.

Trustworthiness also demands restraint. A reliable Legal AI output should clearly reflect the limits of the underlying data and avoid overstating certainty where the law is unsettled or dependent on facts. Outputs that present speculative or probabilistic conclusions as settled law erode trust, even if they sound authoritative. In the legal domain, cautious accuracy is far more valuable than confident generalisation.

v. Transparency: Transparency in Legal AI goes beyond merely arriving at the "right" answer. It also requires that the *reasoning* behind the answer is also visible and understandable. In law, outcomes are rarely accepted at face value, what matters equally is *how* a particular conclusion was reached. A legal opinion, a judgment, or even an interim order derives its legitimacy from the rationale that supports it. The same expectation applies to Legal AI outputs.

Transparent Legal AI systems should therefore be capable of explaining how an answer was generated, what sources were relied upon, which assumptions were made, and how competing considerations were weighed. This is particularly important in complex legal questions where multiple interpretations are possible. If a Legal AI tool merely produces a final answer without revealing its reasoning, it functions as a black box, making it difficult for lawyers or judges to assess reliability, challenge errors, or exercise independent judgment.

This is where thinking models or explainable reasoning frameworks become relevant. Instead of delivering a single opaque output, these systems can present a step-by-step logic,

highlight relevant statutory provisions, citing relevant portions of judgments, and showing how these authorities were applied to the present facts. Much like a junior lawyer setting out their reasoning in a note, the AI's value lies not just in the conclusion but in the structured explanation that precedes it.

Transparency ultimately reinforces accountability and trust. When users can see *how* an answer was formed, they are better positioned to verify it, disagree with it, or refine it using their own expertise. In the legal domain, where reasoning is as important as results, transparency is not a luxury but a necess

Chapter 6

Technology behind Legal AI tools

Legal AI tools may appear seamless on the surface, but they rely on a combination of well-established technologies working together to convert raw legal material into usable outputs. This chapter demystifies some of the most important of these technologies in simple and practical terms. This chapter focuses on four core pillars that power most Legal AI systems today: OCR, NLP, RAG, and Vector Databases & Embeddings.

i. Optical Character Recognition (OCR)

The legal ecosystem heavily relies on scanned documents, legacy contracts, annexures, certified copies of judgments, and handwritten or typed records that have been digitised as images. Optical Character Recognition, commonly known as OCR, is the foundational technology that allows Legal AI systems to read documents that are not originally in digital text form. OCR converts the documents into machine readable text. In practical terms, it enables a Legal AI system to *read* a scanned document in the same way a human reads printed text.

Importance of Document Quality

OCR accuracy is highly dependent on the quality of the scanned document. One of the most important technical factors here is **DPI (Dots Per Inch)**. **300 DPI** is widely regarded as the minimum acceptable standard for reliable OCR in legal documents.

Documents scanned below this resolution often result in:

- - Misread characters (e.g., "l" mistaken for "1")
- - Missing punctuation
- - Distorted words or incomplete clauses

Two preprocessing steps play a crucial role here: binarization and noise removal. Binarization simplifies the document by converting it into a clean black and white format so that the text clearly stands apart from the background, where text denotes 1 and background denotes 0. Noise removal on the other hand eliminates irrelevant visual elements such as stains, ink smudges, scanner shadows, or stray marks that often appear in legal documents. While a human reader may effortlessly ignore these imperfections, a machine may misinterpret them as characters or punctuations, leading to unintended hallucinations.

This stage is often invisible to the end user, yet its impact is profound. Any error introduced during the OCR stage becomes part of the data that the Legal AI system later relies upon for analysis. A misread word, section number, or date can quietly distort downstream outputs, all while appearing confident and authoritative. In legal practice, where outcomes rely on precise language, OCR is not a peripheral technical step but is the foundation wheel upon which the credibility of every subsequent AI-assisted output rests.

Limitations

One of the most significant limitations of OCR, especially in the legal domain, is its difficulty in handling handwritten content. Handwriting varies widely from person to person, often lacking consistency, and includes abbreviations, symbols, or personal shorthand that do not follow any standard pattern. Court notes, marginal annotations on contracts, or handwritten statements can therefore be misread, ignored or partially captured by OCR systems. Even sophisticated, AI-assisted OCR struggles when handwriting is unclear, hurried, or stylistically unique. Unlike a human reader, the system cannot rely on context, experience, or intuition to "fill in the gaps," which makes handwritten documents a persistent blind spot which reinforces the need for human verification.

ii. Natural Language Programming (Prompt Engineering)

At its simplest, Natural Language Programming which is more commonly known as *prompt engineering* is the practice of giving clear, structured instructions to an AI system (such as LLM's) in plain language so that it produces useful and relevant outputs. Instead of writing complex computer codes, the user directs the system by carefully choosing words, context, and constraints. In the legal domain, the quality of the output often depends less on the sophistication of the AI and more on how well the question or instruction is framed. There are several recognised ways of structuring prompts, each suited to different tasks. Some of the most commonly used techniques are as follows:

- **Zero-shot prompting** is the most basic form of prompting that is widely used. The system is given a task without any examples, such as asking it to summarise a judgment or identify issues in a contract. This works well

for straightforward tasks but may lack precision in case of complex legal analysis.

- **One-shot prompting** involves providing a single example along with the instruction for extracting outputs . For instance, showing the AI tool one sample clause analysis before asking it to analyse another clause. This helps the system to better understand the expected format and depth of response. This is beneficial when we wish to get our answers in a specific format.
- **Few-shot prompting** extends this idea by providing multiple examples. In legal drafting or classification tasks, a few well-chosen examples can significantly improve consistency and accuracy of the output by guiding the system toward the desired reasoning pattern.
- **Many-shot prompting** uses a larger number of examples, often in specialised or repetitive tasks such as contract categorisation or document tagging. While effective, it requires more effort and longer inputs and is typically used in controlled or enterprise environments, where the margins for error are negligible.
- **Chain-of-thought prompting** is a technique where the user not only instructs the system on *what* answers are required, but also guides *how* the reasoning should proceed, by indicating the logical steps or thought processes to be followed for generating the output
- **Negative prompting** works by clearly stating what the system should *not* do. For example, instructing the AI not to speculate beyond the given facts, not to cite non-existent case law, or not to quote statutes beyond a certain jurisdiction. This is especially important in law, where overconfidence can lead to misleading outputs.

Each of these techniques reinforces a central truth about Legal AI: it does not "think" like a lawyer, but it responds to how it is guided. Prompt engineering, therefore, is less about technical skill and more about clarity, discipline, and precision in asking the right questions.

iii. Retrieval Augmented Generation (RAG)

Retrieval Augmented Generation (RAG) is a technique in which an AI system retrieves relevant information from defined sources before generating its response, instead of solely relying on what was learned during the training phase. This retrieval can happen in two ways, locally and externally. **Local RAG** operates on documents that are uploaded directly into the tool, such as case files, contracts, pleadings, or internal research notes. **External RAG**, on the other hand, retrieves information from outside sources, such as legal databases, online repositories of judgments, or statutory portals. In both cases, the system first identifies the most relevant material and then uses that content as the basis for its answer, ensuring that responses are grounded in actual legal text rather than abstract predictions.

The relevance of RAG becomes especially pronounced in **legal research**, where accuracy, source integrity, and completeness are non-negotiable. Legal research is not about producing the most eloquent answer; it is about identifying the correct statute, the applicable precedent, and the precise reasoning adopted by courts in a certain scenario. A RAG-enabled Legal AI tool approaches research much like a diligent junior associate: it first scans through a corpus of judgments, statutes, regulations, commentaries, or internal research memos, identifies the most relevant portions, and only then synthesises an answer based on those materials.

In practical terms, when a lawyer searches for the judicial position on a specific legal issue, a RAG system retrieves relevant cases, extracts pertinent paragraphs, and uses them to frame a coherent response. This ensures that the output is anchored in existing law rather than abstract reasoning, heavily countering the scope for hallucinated responses. More importantly, it allows the researcher to trace the answer back to its source, enabling verification, citation, and deeper analysis. This source-based grounding significantly reduces the risk of hallucinated or fabricated legal propositions, which can otherwise be misleading in a research context.

RAG also addresses one of the core challenges of legal research: **the evolving nature of law**. Statutes are amended, precedents are overruled, and interpretations shift over time. Because RAG works on a live or regularly updated document repository, whether local or external, it can reflect the most recent legal position provided the underlying database is regularly updated. This is a substantial improvement over traditional language models that remains static after training.

Another key advantage of RAG in legal research is the **controlled scope**. Law firms, courts, and legal teams can limit the retrieval process to jurisdiction-specific databases or firm-approved materials. This prevents contamination from irrelevant foreign laws or unreliable sources which further ensures that the research remains contextually accurate. For example, a RAG system used for Indian constitutional law research can be restricted to Supreme Court and High Court judgments, constitutional provisions, and authoritative commentaries on constitution.

Despite its strengths, RAG does not replace the researcher's role. It organises, retrieves, and presents legal material efficiently, but it does not evaluate the strength of a precedent, reconcile

conflicting judgments, or assess strategic relevance. These remain inherently human functions. RAG simply ensures that legal research starts from the right materials and is grounded in verifiable law.

In essence, within the domain of legal research, RAG transforms Legal AI from a speculative answer generator into a **source-driven research assistant**, enhancing speed and accuracy while preserving the human's central role in analysis and decision making.

iv. Vectors and Embeddings

Vector Databases and Embeddings are the core background technologies that allow Legal AI systems to work meaningfully. To understand them properly, it is important to familiarize ourselves with a key term that is often used in this context, that is, **semantic**.

In simple terms, **semantic** refers to *meaning*. When we say "semantic similarity," we are not asking whether two sentences use the same words, but whether they *mean the same thing*. For example, the phrases *"breach of contractual obligation"* and *"failure to perform contractual duties"* are semantically similar even though they share very few words. Humans recognise this instantly but Legal AI systems attempt to replicate this ability using the concept of embeddings.

An **embedding** is a way of converting texts, such as a judgment, statute, clause, notice, or pleading into numbers that represent its meaning. Think of it as translating legal language into a mathematical form that a computer can work with. Each document or paragraph is converted into a numerical key that captures what the text is about, not just the words it contains. Texts that deal with similar legal ideas will have embeddings that

are closer to each other, even if the language used is very different.

Once these embeddings are created, they are stored in what is known as a **vector database**. A vector database is not like a traditional database that searches for exact words or phrases. Instead, it searches for *similarity of meaning*. When a user asks a question or uploads a document, that input is also converted into an embedding. The system then compares it against the embeddings already stored in the database and retrieves the content that is closest in meaning.

The use of vector databases and embeddings is not limited to legal research alone. They are equally relevant in **contract analysis**, **due diligence**, **litigation support**, **compliance review**, and **document classification**. For example, when reviewing hundreds of contracts, a Legal AI tool can quickly identify clauses that are semantically similar to a 'dispute resolution' clause, even if the wording of the clause varies from contract to contract. In litigation, it can group pleadings or evidence based on underlying legal issues rather than file names or keywords. In compliance work, it can flag documents that are conceptually aligned with regulatory obligations, even when the regulation is paraphrased.

This meaning-based approach is especially useful in law because legal language is rarely uniform. The same principle may be expressed differently across jurisdictions, courts, or time periods. A keyword-based system may miss relevant material simply because the phrasing has changed. *Vector databases avoid this problem by focusing on substance rather than form, meaning rather than words.*

That being said, these systems still have their limitations. They do not understand the law in a human sense, nor do they assess

correctness, fairness, or applicability. They only determine *which documents are meaningfully similar*. The responsibility of interpreting, weighing, and applying those documents remains firmly with the human lawyer or judge.

In essence, **embeddings teach the system what legal text "means," and vector databases organise those meanings so they can be retrieved efficiently**. Together, they allow Legal AI tools to move beyond simple word-matching and operate at a conceptual level, while still remaining dependent on human judgment for legal reasoning and decision-making.

v. Large Language Models (LLMs)

Large Language Models, commonly referred to as **LLMs**, form the core "thinking engine" behind most modern Legal AI tools. It is important to understand that the majority of legal AI products available in the market today are **not standalone intelligence systems**. They are mostly **wrapper tools** built around one or more existing LLMs, which operate in the background as the primary reasoning and language-generation mechanism. The legal interface, document uploads, citations, or workflows may differ, but the underlying engine that processes language and generates responses is typically the same class of LLM technology.

At a fundamental level, LLMs do not understand law, language, or meaning in the human sense. They operate on **probability**. When an LLM produces an answer, it is not applying legal logic or interpretation. Instead, it is statistically predicting what word is most likely to come next, based on patterns it has learned from vast quantities of text such as statutes, judgments, academic writing, and general language data, among others.

This process happens **one word at a time**. The model does not pre-plan an entire response. Each new word is generated by

asking a single question repeatedly: *given everything written so far, what is the most probable next word?* The model then selects that word and moves forward. The entire output is therefore a sequence of probabilistic guesses, each building on the last.

This probabilistic design explains why LLM outputs often appear coherent, structured, and legally styled. The model has seen enough legal writing to reproduce its tone and form. However, it also explains why these systems can confidently generate incorrect or misleading content. When sufficient grounding is absent, such as verified source documents or human supervision, these models may still produce a fluent answer even when the underlying information is flawed. Confidence, in this context, is not a marker of correctness; it is merely a reflection of statistical likelihood.

In the legal ecosystem, this distinction is critical. LLMs do not verify facts, assess evidence, balance rights, or apply discretion. They predict language. This is why Legal AI tools must be supported by mechanisms such as Retrieval Augmented Generation, source traceability, output moderation, and, most importantly, human oversight. Without these safeguards, probabilistic outputs risk being mistaken for legal conclusions.

In essence, LLMs are powerful linguistic prediction systems, not legal decision-makers. Legal AI tools that rely on them must be understood for what they are: **assistive technologies built on probabilistic models**, designed to support legal professionals, not replace legal reasoning or responsibility.

Chapter 7

Legal Frameworks supporting AI development

Law changes the society or the society brings about changes in law? This is a widely discussed topic in the legal cohorts and the answer lies somewhere in the middle. Sometimes law brings about changes in the society and sometimes society brings about changes in law. When we compare the two dimensions, clearly society brings more changes in law vis-a-vis law bringing changes in society. Similar is the story with the technological advancement that has taken up in the field of Artificial Intelligence. The rate at which these advancements have been made, have compelled the regulatory frameworks to catch up. Some jurisdictions anticipated these changes and accordingly have evolved their legal systems. They have introduced well researched legislations which are non-reactive and some legal systems are yet to devise a framework to meaningfully address these disruptive technologies. In this chapter we are going to discuss some of the most well thought out legislations and judgements which are essential in order to properly understand the applications and limitations of AI based technologies (which includes legal AI).

1. EU-AI Act: The European Union's Artificial Intelligence Act marks a decisive shift in how technology, particularly Artificial Intelligence, is regulated at a systemic level. It is the first serious

attempt by any jurisdiction to move away from reactive regulation and instead introduce a forward-looking, structured framework that governs AI across its entire lifecycle. The Act applies not only to entities based within the European Union, but also to any AI system whose outputs are used within the EU. In other words, geographical boundaries do not act as a shield if the effects of an AI system are felt inside the Union.

At its core, the EU AI Act is built on a *risk-based regulatory mechanism*. Rather than treating all AI systems alike, it differentiates based on the potential impact of the system on individuals, society and rights. This approach recognises a ground reality that not all AI systems pose the same level of risk, and regulation must reflect that reality.

One of the most important contributions of the Act is that it clearly identifies *who is responsible* at different stages of the AI value chain. The law introduces distinct roles: **providers, deployers, and importers**, each carrying its own set of responsibilities.

A **provider** is an entity that develops an AI system, or arranges for its development, and then places it on the marketplace or puts it into use under its own name. This includes developers of foundational models as well as companies that repackage or substantially modify existing AI systems and present them as their own.

A **deployer** is the person or organisation that actually uses the AI system in practice. This could be a law firm using an AI-based research tool, a corporation deploying an automated compliance system, or a public authority relying on AI-assisted decision-making. The key idea is control and have an authority over how the system is used.

An **importer** is an entity that brings an AI system developed outside the EU into the European market. Even if the AI was

lawfully built elsewhere, the importer becomes responsible for ensuring that it complies with EU standards before it is made available or used within the Union.

An important nuance is that responsibility can shift. If a deployer or importer significantly alters an AI system, changes its intended purpose, or rebrands it, the law may treat that party as a provider. This prevents responsibility from being diluted through contractual arrangements.

Exceptions and Carve-Outs: The EU AI Act is not blind to practical realities. It consciously avoids over-regulation by carving out certain limited exceptions. AI systems used purely for *personal and non-professional purposes* fall outside the ambit of the Act. An individual experimenting with AI at home, without any commercial or professional intent, is not the target audience of this legislation. Similarly, *scientific research and experimental development* receive a degree of regulatory breathing room. The intention here is to ensure that academic inquiry and controlled innovation are not stifled, provided that such research does not translate directly into market deployment or real world decision making affecting individuals. The core concept behind the statute is to avoid *'over regulation'* and *'jeopardising technological innovation'*.

Risk-Based Framework

The EU AI Act adopts a **risk-based regulatory approach**. Instead of treating all AI systems as same, it categorises them according to the level of risk they pose to individuals and society:

- **Unacceptable Risk AI (Prohibited):** Under Article 5 of the EU AI Act, certain AI applications are considered to pose an inherently unacceptable risk and are therefore outrightly prohibited. These are systems whose very nature or intended use is incompatible with fundamental rights and democratic values. The law targets AI systems

that exploit human vulnerabilities or use manipulative or deceptive techniques to influence behaviour in a way that causes serious harm. This category includes systems that assign social scores to individuals based on their behaviour, personal traits, or socio-economic status, as well as tools that attempt to predict an individual's likelihood of committing a crime solely on the basis of profiling or personal data. The Act also prohibits practices such as the large-scale scraping of facial images from the internet or CCTV footage to build biometric databases, the use of AI to infer emotions in workplaces or educational institutions, and the categorisation of individuals based on biometric characteristics. Limited exceptions are carved out for law enforcement purposes, such as locating a missing person or preventing serious threats like terrorism. The underlying principle is clear: certain uses of AI are considered so intrusive or dangerous that no level of technical accuracy can justify their deployment. Similarly, any attempt to push Legal AI into behavioural prediction or moral assessment risks crossing into prohibited territory. This is precisely the reason why predictive analysis was clubbed in the category of 'What legal AI cannot do' under chapter 4.

- **High-Risk AI:** Under Article 6 of the EU AI Act (classification rules for high-risk AI systems), AI systems are classified as *high risk* when their intended use creates a significant possibility of harm to fundamental rights, safety, or legally protected interests. The classification does not depend on how advanced or complex the technology is, but on the function the system performs and the context in which it is deployed. This marks a deliberate shift away from technology-centric regulation

toward impact-based regulation. High-risk AI systems generally fall into two broad categories.

i. First, AI systems that function as safety components of regulated products or are themselves regulated under existing EU product safety laws.

ii. Second, AI systems deployed in specific sensitive domains identified in the Act, where errors or bias can have serious legal or social consequences. These domains are detailed in Annex III of the EU AI Act, which serves as the practical guide for identifying high-risk use cases.

Once an AI system is classified as high risk, it becomes subject to a strict set of obligations (Article 8: Compliance with the requirements). Providers must implement continuous risk management measures, ensure that training data is appropriate and well-governed, and maintain detailed technical documentation explaining how the system operates, among others. The system must be designed to allow meaningful human oversight, ensuring that humans can intervene, override, or correct outputs where necessary. Accuracy, robustness, and resistance to misuse are not treated as optional features but as legal requirements. Deployers, in turn, are required to use such systems strictly within their intended purpose and to monitor their performance in real-world conditions.

A particularly important clarification emerges from **Annex III**, which emphasises that risk classification is **use-based and contextual**. The same AI system may be high risk in one setting and low risk in another. Certain narrowly defined, low-impact uses listed in Annex III may be exempted from high-risk obligations, especially where the system performs purely procedural or administrative tasks and does not involve profiling or

evaluative judgment. This ensures that the regulatory framework remains proportionate and does not stifle benign or routine applications.

Legal AI tools come into focus most clearly under **Annex III, point number 8**, which includes AI systems used in judicial and democratic decision-making. This covers systems that assist with legal interpretation, assess evidence, or influence outcomes affecting rights and obligations. While basic Legal AI tools used for drafting, research, or document management may remain outside the high-risk category, systems that shape or influence legal reasoning risk being classified as high risk under Article 6(2) of the Act. The decisive factor is not technological sophistication, but whether the system plays a substantive role in processes where legal rights are at stake. This reinforces a foundational principle of the EU AI Act: AI may support the administration of justice, but it must never silently steer it.

- **Limited and Minimal Risk AI:** It must be clearly stated that the EU AI Act does not create a separate or formally labelled category called "limited" or "minimal" risk AI. However, Recital 53 of the Act provides interpretive guidance by explaining that certain AI systems may, in practice, pose a reduced level of risk where they do not materially influence decision-making or affect the substance or outcome of legally relevant processes. Recital 53 clarifies that risk assessment should focus not on the domain in which an AI system is deployed, but on the degree of influence the system exerts. Where an AI system merely supports a human process—without shaping, steering, or determining outcomes—the legal and rights-based risk associated with its use is substantially lower. This reasoning is

particularly relevant for AI systems performing narrow, procedural, or preparatory functions. Examples include tools that structure unorganised information, classify or index documents, detect duplicates, assist with file management, or translate documents at an initial stage.

Similarly, AI systems that improve language quality, refine drafting, or flag deviations after a human assessment has already been completed operate only as an additional layer of support rather than as decision-making engines. In such cases, the AI does not replace human judgment, nor does it meaningfully alter legal consequences. Recital 53 therefore embeds a principle of proportionality into the regulatory architecture of AI-based applications. It acknowledges that not every use of AI in sensitive contexts warrants the same regulatory burden. At a policy level, this approach is crucial: it avoids over-regulation of assistive technologies while preserving strict safeguards for AI systems that genuinely shape or affect rights and obligations.

2. OECD Principles on Artificial Intelligence

In contrast to the EU AI Act, which is a binding legislative instrument with enforceable obligations, penalties, and a detailed risk-classification framework, the OECD Principles on Artificial Intelligence introduced in 2019, adopts a **non-binding, values-based approach**. They do not impose legal duties or compliance mechanisms. Instead, they function as soft law or normative guidance intended to shape policy making, corporate governance, and international consensus around the responsible development and use of AI.

Despite their non-binding nature, the OECD Principles are historically significant. Adopted in 2019, they represent the **first**

intergovernmental framework on Responsible AI, endorsed by OECD member states and several non-member countries. Their influence extends far beyond the OECD itself: they have informed national AI strategies, inspired regional frameworks, and directly shaped later binding instruments, including the EU AI Act. Where the EU AI Act operationalises risk and responsibility through legal rules, the OECD Principles articulate the ethical and societal foundations upon which such regulation is built.

The OECD framework is structured around five core principles, each addressing a critical dimension of trustworthy AI.

a. Inclusive Growth, Sustainable Development and Well-Being (Principle 1.1)

This principle recognises AI not merely as a technological innovation, but as a societal force with the potential to advance inclusive economic growth, environmental sustainability, and human well-being. Trustworthy AI should contribute positively to individuals, communities, and the planet, rather than concentrating the benefits in narrow segments of society.

The emphasis here is on **shared prosperity**. AI systems should support global development objectives, reduce inequality, and enhance quality of life. This principle implicitly warns against AI deployments that optimise efficiency or profit at the expense of social cohesion, labour dignity, or environmental sustainability. Unlike the EU AI Act, which focuses on risk mitigation, this principle foregrounds AI's affirmative social purpose. Principle 1.1 positions trustworthy AI as an enabling force for broad-based societal progress rather than a purely efficiency-driven technology. It calls upon all relevant stakeholders to take an active and responsible role in the stewardship of AI systems, with the objective of generating outcomes that are beneficial not only

for markets and institutions, but for individuals, communities, and the environment at large.

By integrating social equity, economic development, and environmental responsibility into a single framework, the OECD principles articulate a vision of AI governance that supports inclusive growth, collective well-being, and sustainable development as interconnected objectives, rather than competing priorities.

b. Respect for the Rule of Law, Human Rights and Democratic Values, Including Fairness and Privacy (Principle 1.2)

Principle 1.2 anchors the development and deployment of AI systems firmly within the framework of the rule of law and universally recognised human rights. It requires AI actors to ensure that democratic, human-centred values are respected at every stage of the AI lifecycle, from design and training to deployment and post-deployment monitoring. This includes a clear commitment to principles such as equality and non-discrimination, individual freedom and dignity, personal autonomy, privacy and data protection, as well as respect for diversity, fairness, social justice, and internationally accepted labour standards.

The principle also recognises the evolving risks associated with AI-driven information ecosystems. In particular, it calls for attention to the ways in which AI systems may contribute to the creation or amplification of misinformation and disinformation, while simultaneously safeguarding freedom of expression and other fundamental rights protected under democratic setups. The emphasis is not on blanket restrictions, but on responsible design choices that minimise systemic harm without undermining democratic discourse.

To operationalise these values, Principle 1.2 encourages the adoption of appropriate safeguards and governance mechanisms. These include preserving meaningful human agency and oversight over AI systems, especially in contexts where misuse, whether deliberate or accidental, or deployment beyond the system's intended purpose could lead to rights based harms. Such safeguards must be proportionate, context-specific, and aligned with the industry best practices, reinforcing the idea that respect for rights is not an abstract aspiration but a practical and continuous obligation in responsible AI governance.

While the EU AI Act translates these concerns into concrete prohibitions and compliance duties (for example, in relation to biometric categorisation or social scoring), the OECD principle operates at a normative level, setting ethical boundaries rather than legal thresholds.

c. Transparency and Explainability (Principle 1.3)

Principle 1.3 places transparency and responsible disclosure at the core of a trustworthy AI. It requires AI actors who design, develop, deploy, or operate AI systems to proactively communicate how AI systems function and how they affect individuals and organisations. The emphasis is not on technical disclosure for its own sake, but on providing *meaningful information* that is appropriate to the context and consistent with the prevailing state of the art.

At a foundational level, this principle seeks to foster a general understanding of AI systems, including their capabilities as well as their limitations. Users and affected stakeholders should not be left with an inflated or misleading perception of what an AI system can or cannot do. Closely linked to this is the obligation to ensure that individuals are made aware when they are interacting with an AI system, including within workplace

environments, where power asymmetries and dependency on automated tools may be particularly pronounced.

Where feasible and useful, transparency must extend beyond surface-level disclosures. AI actors should provide plain, clear, and intelligible information regarding the sources of data or inputs used by the system, as well as the key factors, processes, or logic that contribute to a recommendation, content generation, or decision. The objective is not to expose proprietary algorithms, but to enable those affected to understand how an outcome was reached in a manner that is comprehensible and practically relevant.

Finally, transparency under Principle 1.3 is inseparable from procedural fairness. Information must be provided in a way that enables individuals who are adversely affected by an AI system to question, contest, or challenge its outputs. Transparency, therefore, is not treated as an abstract ethical value, but as a functional safeguard—one that underpins accountability, trust, and the ability to seek redress in AI-mediated decision-making.

d. Robustness, Security and Safety (Principle 1.4)

This principle requires AI actors to embed resilience into system design and deployment. Appropriate safeguards should exist to detect failures, unexpected behaviour, or emerging risks at an early stage. Where an AI system begins to behave in a manner that could cause harm, there must be clear technical and organisational pathways to intervene whether by overriding system outputs, correcting faults, limiting functionality, or, where necessary, withdrawing the system from use altogether. The ability to safely repair or decommission AI systems is therefore a core component of responsible AI governance.

In addition, Principle 1.4 acknowledges the broader societal implications of AI-generated and AI-amplified information.

Where technically possible, measures should be adopted to support the integrity and reliability of information ecosystems, including reducing the spread of harmful or misleading content. At the same time, such measures must be carefully calibrated so that they do not erode freedom of expression or other fundamental rights. The principle thus reinforces a balanced approach, one that treats robustness, security, and safety not as purely technical goals, but as safeguards essential to maintaining public trust in AI systems.

This lifecycle-based thinking strongly influenced later regulatory models, including the EU AI Act's requirements on accuracy, robustness, and post-market monitoring. The difference lies in enforcement: under the OECD Principles, robustness is a **best practice expectation**, not a statutory mandate.

e. Accountability (Principle 1.5)

Accountability is the connective tissue of responsible AI. This principle states that organisations and individuals who develop, deploy, or operate AI systems should be accountable for their proper functioning. Principle 1.5 establishes accountability as the foundation of responsible AI use. It makes clear that AI systems do not operate in a vacuum and human actors remain responsible for their proper functioning.

A key aspect of accountability is traceability. AI actors are expected to maintain the ability to trace how an AI system was developed, trained, and operated, including the datasets used, key processes followed, and decisions taken throughout the system's lifecycle. This traceability allows outputs to be examined, questioned, and explained when concerns arise, and ensures that AI-driven outcomes can be meaningfully reviewed in response to complaints, audits, or legal inquiries.

Principle 1.5 also emphasises continuous risk management. AI actors should adopt a systematic approach to identifying and addressing risks at every stage of the AI lifecycle, rather than treating risk assessment as a one-time exercise. This includes risks relating to bias, human rights, safety, security, privacy, labour protections, and intellectual property, among others.

Unlike the EU AI Act, which carefully allocates responsibility across providers, deployers, and importers, the OECD principle does not define legal roles or liabilities. Instead, it establishes a **moral and organisational expectation**: AI should not become a tool to evade responsibility. Human actors must remain answerable for outcomes, even when decisions are supported or automated by AI.

Summary of OECD principles:

The OECD Principles and the EU AI Act serve complementary functions. The OECD framework provides the **ethical and policy blueprint** for responsible AI, while the EU AI Act translates similar concerns into **binding legal architecture**. Together, they illustrate the principle as to how society's values first shape norms, which later crystallise into law. It reinforces the broader idea that technology usually advances first, prompting regulation to catch up, and that once introduced, regulation then influences how technology is developed and used.

3. U.S Regulatory Approach

The United States has adopted a pro innovation approach for the governance of Artificial Intelligence, one that stands in deliberate contrast to the comprehensive, binding, and ex ante regulatory architecture exemplified by the European Union's AI Act. Rather than enacting a single horizontal statute that classifies AI systems and imposes uniform obligations across sectors, the U.S.

framework has evolved through a combination of executive action, sector-specific regulation, voluntary standards, and the application of existing legal doctrines. This approach reflects a long-standing regulatory philosophy in U.S. technology governance: Innovation should be enabled by default, and regulation steps in only when real harm is clearly visible, not based on hypothetical risks.

At the centre of the U.S. model is the belief that premature or overly prescriptive regulation risks constraining technological leadership, economic competitiveness, and national security interests. The U.S Administration has repeatedly emphasised that the United States' strategic advantage in AI depends on its ability to foster research, entrepreneurship, and rapid deployment at scale, which can only happen when the barriers to innovation, that is, over-regulation is countered.

Centralised AI policy: The federal government has expressed increasing unease with the emergence of a patchwork of state level AI related laws. From a U.S. policy perspective, such fragmentation risks discouraging investment, slowing deployment, and jeopardizing domestic firms in global competition. As a result, federal executive action has increasingly emphasised harmonisation, coordination, and the primacy of national AI policy over disparate local interventions.

This preference for federal uniformity is closely tied to the U.S. reliance on existing legal frameworks rather than AI-specific statutes. Consumer protection agencies, such as the Federal Trade Commission, have asserted jurisdiction over deceptive or unfair AI practices under established consumer law. Civil rights agencies have addressed algorithmic discrimination through existing anti-discrimination statutes. Competition authorities have scrutinised AI-driven market concentration under antitrust law. In this model, AI is not regulated because it is AI, but

because it produces effects that fall within already regulated legal domains. This approach allows regulators to intervene where harm materialises, without imposing ex ante design or deployment constraints on all AI systems irrespective of risk.

A defining feature of the U.S. approach is its reliance on soft law, guidance, and voluntary standards rather than mandatory compliance obligations. The NIST AI Risk Management Framework is emblematic of this philosophy. Rather than prescribing legal duties, it provides a structured methodology for identifying, assessing, and managing AI risks across the system lifecycle. Adoption is voluntary, but the framework has gained normative traction across industry, government, and international partners. It reflects the U.S. preference for multi-stakeholder governance, where industry, academia, civil society, and government collectively shape responsible AI practices. This stands in contrast to the EU's top-down legislative model, which imposes legally enforceable obligations based on predefined risk categories.

The U.S. approach also reflects a pragmatic understanding of the pace of technological change. Policymakers have acknowledged that AI capabilities are evolving too rapidly for static legal classifications to remain accurate over time.This is particularly evident in the federal government's own use of AI, where agencies are encouraged to experiment, pilot, and deploy AI tools while maintaining internal governance controls, human oversight, and accountability mechanisms. The emphasis is on learning through use rather than restricting use through precautionary bans.

This innovation-first orientation inevitably entails trade-offs. By avoiding comprehensive ex ante regulation, the U.S. model accepts a higher degree of uncertainty and potential downstream harm. Risks related to bias, transparency, explainability, and

accountability may not be fully addressed until the systems are deployed and impacts become visible. Critics argue that this reactive posture places disproportionate burdens on affected individuals and communities, particularly where AI systems influence access to employment, credit, healthcare, or justice. Supporters counter that rigid upfront controls would freeze innovation and entrench incumbent technologies, ultimately harming both economic dynamism and societal progress.

The U.S. model illustrates a regulatory philosophy that prioritises innovation over precaution. Where the EU seeks to shape AI development through predefined legal boundaries, the U.S. seeks to innovate at God speed and hit the brakes only when faced with an obstacle . Neither approach is inherently superior; each reflects different risk tolerances, institutional traditions, and strategic priorities. What is clear, however, is that the U.S. model has played a significant role in accelerating global AI development, even as it continues to grapple with questions of safety, accountability, and public trust.

This model accepts imperfection as the price of flexibility, betting that the capacity to innovate quickly will, over time, generate both economic value and the institutional learning necessary to govern AI responsibly.

4. Indian framework (balancing innovation and regulation)

The Indian framework on AI was recently shared at the **India AI Impact Summit 2026**, which has been widely recognised as the first major global AI summit led from the Global South. India used this forum to articulate a governance philosophy that consciously blends civilisational values with regulatory pragmatism. Unlike the European story where a standalone statute on Artificial Intelligence was adopted, India has followed

a framework based (normative) approach using existing laws to regulate this sector . India has decided to 'prioritise innovation over restraint'. For fostering Innovation, India has decided not to introduce a standalone legislation to regulate this sector. The existing statutes such as the DPDP Act and the IT Act, among others will have special carve outs to regulate this sector. This approach allows AI governance to evolve incrementally through standards, executive guidance, and sectoral regulation, rather than through rigid statutory mandates.

Even though India is prioritising 'innovation over restraint', India has also clarified its stance on building a culture of 'Responsible use of AI', with regulatory interventions 'as and when required'. This was made conspicuous by the establishment of bodies such as the 'India AI Safety Institute (AISI)', 'Technology and Policy Expert Committee (TPEC)' and 'AI Governance Group (AIGG)' to provide regulatory insights and administer AI governance. We can safely say that the Indian AI governance approach is a balancing act of Innovation and Regulation, slightly leaning towards Innovation.

The Three Sutras

The Indian AI Regulation framework is anchored in the **three Sutras**, which operates as guiding philosophical principles rather than enforceable legal mandates. These Sutras provide the baseline ethical guidance to steer India's AI governance discourse. These are:

i. **People** sutra foregrounds a human centric and inclusive approach, requiring AI systems to enhance accessibility (democratisation of AI), and fairness, particularly within large, diverse, and unequal societies to counter bias at all possible levels. The underlying theme of this sutra is Inclusivity.

ii. **Planet** sutra emphasises on sustainability, energy efficiency, and environmentally responsible AI development, recognising the long term ecological implications of compute intensive technologies. The underlying theme of this sutra is Sustainable development.

iii. **Progress** sutra focuses on innovation, collective economic growth and capacity building, with particular attention to the needs of developing economies and emerging markets. The underlying theme of this sutra is Conscious Innovation.

Taken together, these Sutras reflect India's attempt to frame her AI governance around social welfare, sustainability, and developmental objectives, rather than approaching regulation solely through the lenses of risk, liability, or control.

To translate these normative commitments into operational priorities, the summit articulated the **seven Chakras**, which convert abstract principles into actionable outcomes. The Chakras identify the key areas for institutional focus and international cooperation, including themes such as human capital development, social inclusion, transparency, safe and trusted AI, scientific advancement, resilience, robustness, democratisation of AI resources, and AI for economic development. Collectively, these 7 Chakras signal that India conceives AI governance as a multi sectoral and flexible exercise, rather than driven by a rigid regulatory regime.

The Seven Chakras

i. **Human Capital**: Advancing equitable skilling and inclusive workforce transitions for an AI-enabled future of work.

ii. **Inclusion for social empowerment**: Advancing AI systems that are inclusive by design, empowering diverse communities and ensuring equitable representation.

iii. **Safe and trusted AI**: Building globally trusted AI systems anchored in transparency, accountability, and shared safeguards for innovation.

iv. **Science**: Harnessing AI to accelerate frontier science, foster scientific collaboration, and translate breakthroughs into shared global progress.

v. **Resilience, Innovation, and Efficiency**: Driving sustainable, resource efficient AI systems that strengthen climate resilience and sustainability.

vi. **Democratizing AI Resources**: Promoting equitable access to foundational AI resources for inclusive innovation and sustainable development worldwide.

vii. **AI for Economic Development & Social Good**: Leveraging AI to enhance productivity, innovation, and inclusive development across economies and societies.

A defining feature of the Indian framework is its explicit commitment to a **minimal intervention or middle path approach** to AI regulation. India has clearly indicated that it does not intend to enact a single, comprehensive AI statute in line with the **EU AI Act**. Instead, the Indian approach relies on the strategic use of existing legislative frameworks, supplemented by targeted carve outs, sector specific rules, and adaptive regulatory guidance. This position is informed by the view that premature or overly rigid regulation may inhibit innovation in a rapidly evolving technological ecosystem, particularly one driven by startups and experimentation.

In practical terms, AI related risks and obligations are expected to be addressed through laws that already regulate data, digital platforms, and online conduct. Major legal instruments include:

- The **Digital Personal Data Protection Act**, which governs data processing, consent, data retention, and accountability, and is expected to apply to AI systems that are reliant on personal data, subject to calibrated exemptions and sector specific rules.
- The **Information Technology Act**, which continues to provide the foundational framework for India's digital governance.

Overall, the Indian AI framework combines ethical vision with regulatory pragmatism. The three Sutras provide philosophical direction, the seven Chakras offer concrete areas for policy action, and the minimal intervention model ensures that innovation is not unduly constrained by over-precautionary regulations. Through this approach, India seeks to balance innovation and regulation while advancing a governance model that resonates with the realities and developmental aspirations of the Global South.

Chapter 8

Future trends in Legal AI

Having examined the present regulatory, technological, and institutional landscape of Legal AI, it is equally important to look ahead at the trends that are most likely to shape the future trajectory. The developments discussed in this chapter are forward looking indicators, drawn from practical exposure, industry discussions, and engagement with leading experts in this niche domain. They should be understood as **signals rather than settled facts**, particularly given the pace at which AI capabilities and business models continue to evolve. Traditionally, technological growth in computing was understood through **Moore's Law**, which observed that the number of transistors on a microchip would roughly double every two years, leading to predictable and incremental increases in computational capacity. However, recent developments in large language models and generative AI suggests that AI capability growth is no longer strictly anchored in this settled progression. For Legal AI, this means that future transformations may occur more rapidly and unpredictably than past technological cycles would have suggested.

Some of the key trends are mentioned below:

1. **Multi- Agents**: One of the most consequential developments is the rise of **multi-agent AI systems**. Unlike single model

applications that perform singular tasks, multi-agent systems involve multiple AI agents collaborating, delegating, and validating outputs among themselves. In the legal domain, this enables task decomposition across research, drafting, review, compliance checking, and risk assessment. For Legal AI, this could mean autonomous systems capable of handling end to end workflows such as, research followed by drafting followed by compliance of the generated draft, all done as part of a single instruction with human in the loop only coming at the very end. Importantly, it allows human oversight to be strategically positioned at critical checkpoints or at the final output stage, rather than requiring constant supervision at every step of the process.

2. **SaaSapocalypse hype:** Another emerging development often described as the "SaaSapocalypse" refers to the potential disruption of traditional Software as a Service (SaaS) models. This trend has been highlighted due to the emergence of light-weight locally available plugins. This development has profound implications for Legal AI vendors. If a relatively inexpensive plugin can deliver high quality drafting assistance, research acumen, or document review , the traditional subscription based SaaS model faces substantial pressure. For law firms and in house teams, this reduces dependency on external servers, lowers recurring costs, and enhances data control and confidentiality. For Legal AI vendors, this could fundamentally alter pricing, deployment, and product design. Having said that, it is also important to understand that the plugins do not come with 'service providers'. In order to use any software successfully there is a dedicated team of professionals that are readily available to guide their clients in case they are having trouble in resolving a complex task, fixing a bug, taking accountability and ensuring resolution of client's grievances. All these elements require human touch and a dedicated team of trained professionals. This

development is truly noteworthy but it would be a misnomer to equate it with an apocalypse of SaaS models.

3. **Mobile first approach**: A further trend shaping the future of Legal AI is the move towards **mobile based applications**. As legal services expand beyond traditional law firms into in-house teams, startups, small practices, and even individual professionals, there is growing demand for Legal AI tools that function seamlessly on mobile devices. Mobile accessibility enables real time contract review, compliance checks, legal research, and document analysis outside conventional office environments. This trend aligns with broader digital behaviour patterns and may accelerate the adoption of Legal AI among practitioners who were previously excluded due to cost, infrastructure, or usability constraints.

4. **Decline of billable hours:** Perhaps the most structurally disruptive trend is the gradual erosion of the billable hour model. As Legal AI systems increasingly automate research, drafting, due diligence, and review tasks, the correlation between time spent and value delivered becomes harder to justify. Clients are already questioning why efficiency gains enabled by AI should continue to be billed on a time basis. This is accelerating a shift towards fixed fees, outcome based pricing, subscription legal services, or hybrid models where AI enabled productivity is factored into pricing structures. While this does not signal the immediate end of billable hours, it does suggest a long term reconfiguration of how legal value is measured, priced, and delivered.

In this moment of transition, the writings of **Rabindranath Tagore** provide a useful lens to reflect on the evolving role of Legal AI. Tagore observed that *"the highest education is that which does not merely give us information but makes our life in harmony with all existence."* Read in this context, Legal AI should not be assessed solely by metrics of speed, efficiency, or cost optimisation.

Rather, its true value lies in its capacity to align technological innovation with the broader objective of inclusive growth and equitable access to justice. If guided with care, Legal AI can move beyond automation for its own sake and contribute to a legal ecosystem where innovation coexists with justice, efficiency is balanced with accountability, and technological capability remains firmly anchored in human values.