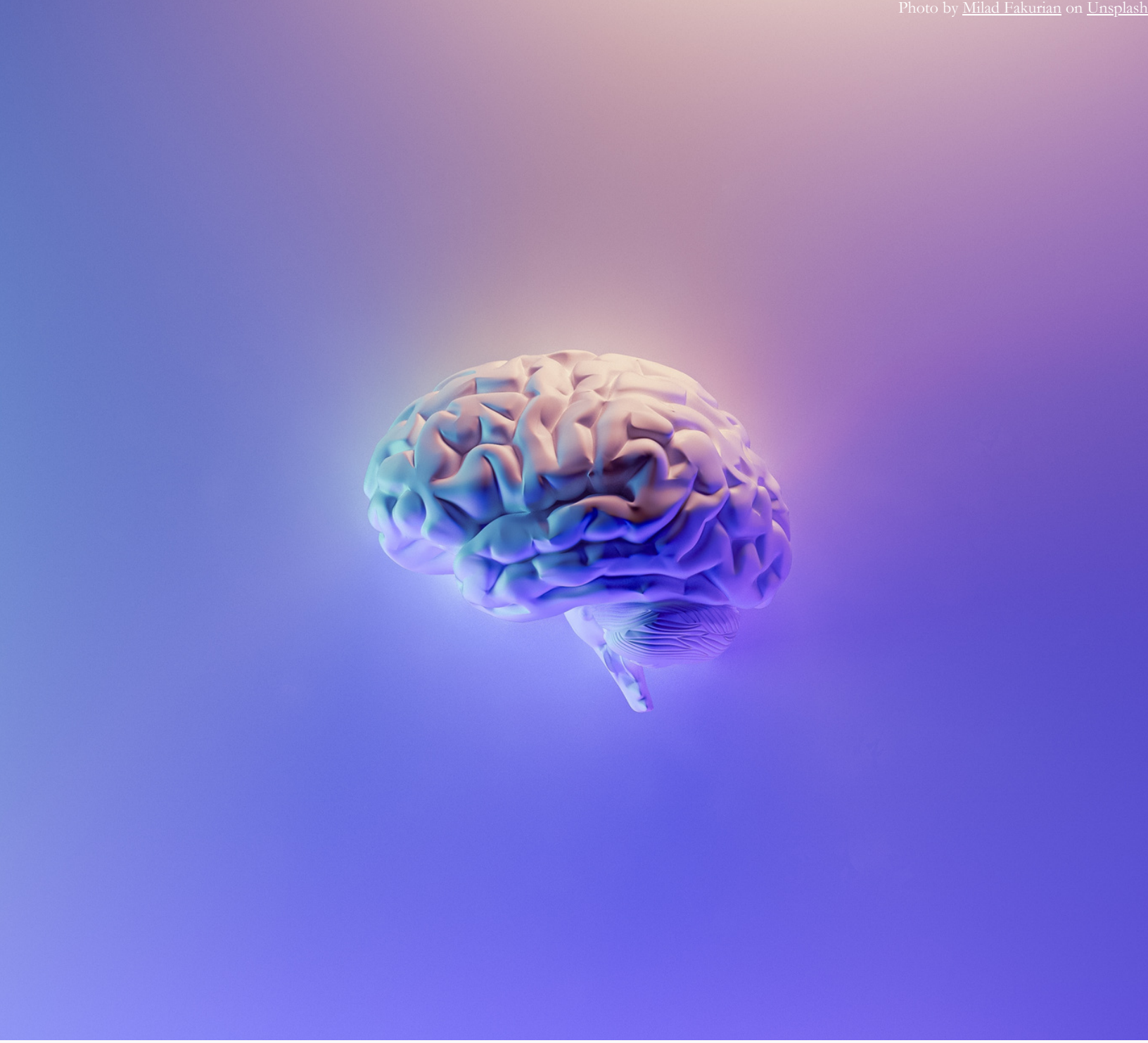


ARTIFICIAL INTELLIGENCE

An A-Z Guide for Lawyers

Walchet, September 2023

WELCOME

In an era defined by rapid technological innovation, Artificial Intelligence (AI) has emerged as a transformative force. AI has profoundly impacted the legal industry, bringing efficiency, accuracy and innovation to the practice and business of law. Even so, complex AI concepts often seem like a puzzle, with missing pieces to those unfamiliar with the technology and its jargon.

To help you navigate this paradigm shift, we have put together a glossary of AI words so that you can understand what they mean and how they could apply to legal teams. This eBook aims to bridge the gap between the legal field and the world of AI, ensuring that legal professionals and enthusiasts alike can grasp its concepts without confusion. With your feedback and input, we hope to keep adding words to the glossary in subsequent editions, to eventually create a more comprehensive guide.

Please write to us at priyadarshini@walchet.com for feedback.
Happy reading!

WALCHET
LAW + TECH + YOU

Legal Tech + Innovation Content Initiative

ABOUT WALCHET

Walchet is a legal technology and innovation content initiative.

Launched in 2022 by Priyadarshini Shetty, a legal tech and innovation expert and former corporate lawyer, Walchet's mission is to educate legal teams about the potential to grow and improve their practice through the adoption of innovative techniques and legal tech solutions.

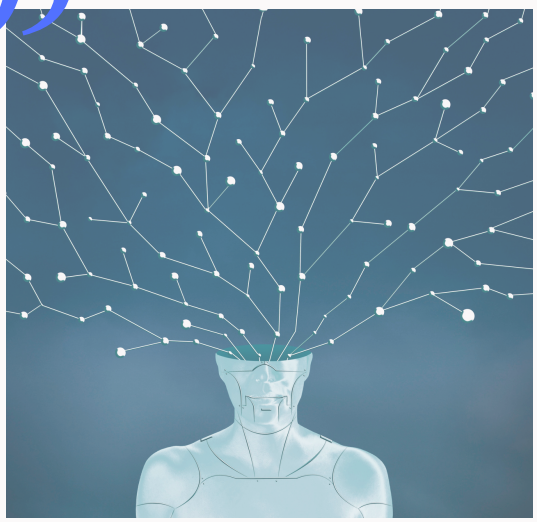
Walchet hosts a podcast on legal tech and innovation called Walchet Talks where we chat with other legal technologists and innovation experts. In each episode, we do a deep dive into one concept in the legal innovation space and discuss its application to legal teams. The podcast has covered a broad range of topics including document management systems, contract lifecycle management systems, document automation, design thinking, e-discovery and lots more!

Walchet also hosts a YouTube channel, Walchet Views, which posts a variety of content, including helpful tips on innovative legal practices, reviews of legal tech products, and Walchet Talks.

If there is a topic that you would like us to cover, please write in at priyadarshini@walchet.com.

CONTENTS

05



Deep Learning

11

Hallucinations

CONTENTS

18

Machine Learning

25

Training

CONTENTS

29

Weak AI

32

You Only Look Once

JARGON

Abbreviations

AI - Artificial Intelligence
LLM - Large Language Model
ML - Machine Learning
NLP - Natural Language Processing

Algorithm

A set of instructions or rules that enable machines to learn and analyse data and make decisions based on that knowledge [1]. It is essentially the programming that tells the machine how to learn to operate on its own [2].

Data

Information that has been converted into a form that is more efficient for processing and transfer. Data may be structured or unstructured, and is often collected to be measured, reported, visualized, and analyzed. [3]

Model

The output of an algorithm run on data. It represents what has been learned from the algorithm on the data. [4]

Chatbot

A computer program designed to simulate conversation with human users.

Systems

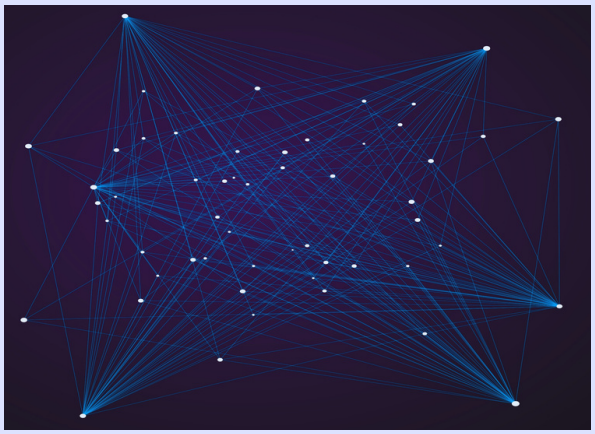
Systems can either be standalone computer programs or they can process data and produce behaviour in a robot body. [5]

AI Tools

A software application that uses artificial intelligence algorithms to perform specific tasks and solve problems. [6]

[1] Fred Tabsharani, "Types of AI Algorithms and How They Work" [2023] TechTarget <https://www.techtarget.com/searchenterpriseai/tip/Types-of-AI-algorithms-and-how-they-work>;
[2] "Artificial Intelligence Algorithm: Everything You Need To Know About It" [2023] Rock Content <https://rockcontent.com/blog/artificial-intelligence-algorithm/>
[3] "Definition of Data Applied to Artificial Intelligence" <https://www.meetbunch.com/terms/data>;
[4] Yahya Abi Haidar, "Difference between Algorithm and Model in Machine Learning" LinkedIn <https://www.linkedin.com/pulse/difference-between-algorithm-model-machine-learning-yahya-abi-haidar/>
[5] D Gamez and O Holland, "Artificial Intelligence and Consciousness☆," Elsevier eBooks (2017) <https://doi.org/10.1016/b978-0-12-809324-5.05918-6>
[6] "What Is an AI Tool?" Synesthesia <https://www.synthesia.io/glossary/ai-tool>.

ARTIFICIAL INTELLIGENCE



Artificial Intelligence (AI) is the simulation of human intelligence processes by machines [7]. AI research has been highly successful in developing effective techniques for solving a wide range of problems, from game playing to medical diagnoses.

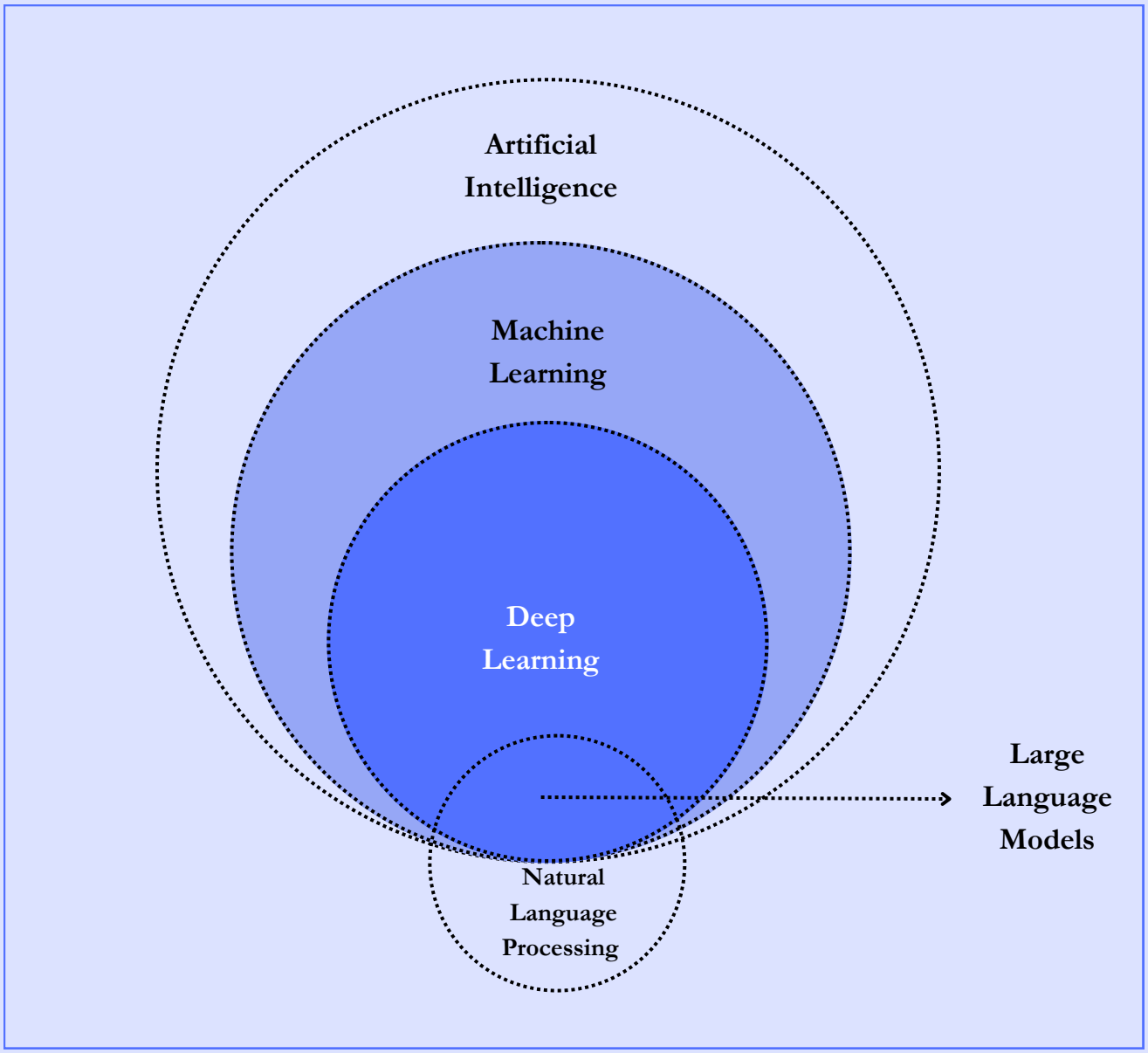
How is AI relevant to law?

AI is poised to fundamentally reshape the practice of law. AI can assist lawyers when reviewing, summarizing and drafting legal documents. AI can also be used to extract specific data points in documents, or to predict certain outcomes that can help lawyers assess the likelihood of success in a case. Ultimately, AI helps lawyers improve their legal service delivery.

Some Use Cases of AI for Legal:

1. **Drafting:** You can use Large Language Models (LLMs) to draft clauses or agreements once it has been exposed to legal data.
2. **Review**: You can tell an AI tool what your standard clauses are, so that it can highlight clauses that deviate from your standards.
3. **Search & Extraction**: Searching through volumes of data to identify and extract specific data points for further review.
4. **Summarisation**: Ask the AI to summarise caselaw, memos, opinions, agreements and other legal documents.
5. **Prediction**: If you fine-tune the AI on relevant jurisprudence, it can provide a prediction of what the outcome of a particular case may be.

AI is an umbrella term that comprises several concepts such as Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP) and Large Language Models (LLM). The diagram below provides an overview of some relevant AI concepts this eBook will address.



An Overview of AI

[7] Ed Burns, Nicole Laskowski and Linda Tucci, "Artificial Intelligence (AI)" [2023] TechTarget <https://www.techtarget.com/searchenterpriseai/definition/AI-Artificial-Intelligence/>.

BIAS IN AI

Bias in AI occurs when ML algorithms produce results that are systemically prejudiced. For instance, with Generative AI, there is a lack of ability to distinguish between right and wrong as it is a probabilistic engine that merely stitches words together, and will produce biased outcomes.

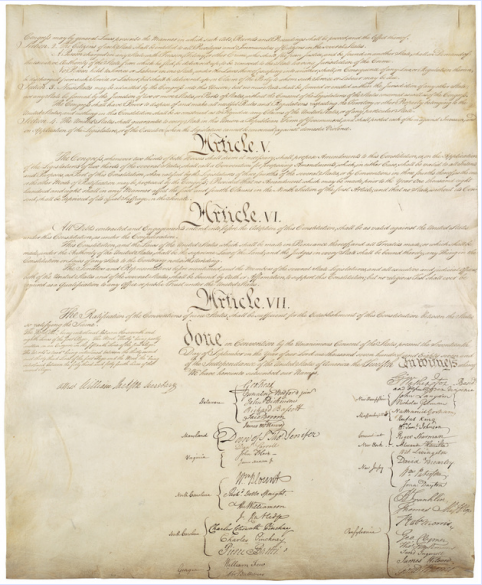
Photo by Jr Korpa on Unsplash

Why are outcomes biased?

- **The data used to train the algorithm is biased.** The quality of the output is determined by the quality of the input. For example, if an algorithm is trained on a dataset of only technology-related contracts, it will struggle to accurately interpret and assess contracts of other sectors like healthcare and manufacturing.
- **The algorithm itself is biased.** The algorithm may be designed in a way that favors one group of people over another. For example, an algorithm that is designed to predict whether a loan applicant is likely to default may be biased against minorities that are traditionally denied loans.
- **The way the algorithm is used is biased.** For example, an algorithmic credit scoring system that is solely reliant on traditional credit data (such as credit card history and loan repayment records) may exclude individuals that are financially responsible but lack conventional creditworthiness.

Biased AI produces inaccurate, misleading and sometimes harmful outcomes, leading to a loss of trust. To prevent AI bias, it is important to use diverse and representative data to design and train algorithms, test the algorithms in a real-life setting before deploying them, consider human feedback on outcomes, and educate users on AI bias and fair use of algorithms.

CONSTITUTION -AL AI


Article V.
Article VI.
Article VII.
Done
Attest William Jackson Secretary

Constitutional AI is an approach to AI safety that involves designing AI systems with built-in principles and constraints that align the systems' goals and behaviors with human values.

Anthropic, an AI safety and research company, uses Constitutional AI for its LLM Claude.

Anthropic's Constitution is based on a variety of sources, including the UN Declaration of Human Rights, Apple's Terms of Service, the Sparrow Principles from DeepMind, and principles encouraging consideration of non-Western perspectives.

Constitutional AI seeks to encode these principles directly into the AI's core architecture, so that it becomes a fundamental part of how the AI operates. This contrasts with more ad-hoc or external alignment approaches, where the principles are applied to the AI after it has been trained.

For instance, LLMs that are based on Constitutional AI would not answer a question asking how to bypass security systems and shoplift a store, since the accurate response would be harmful and deemed illegal per the constitutional principles. Accordingly, the LLM alters the response to state "shoplifting is illegal and can get you arrested".

Constitutional AI can help address concerns about AI safety, make AI systems more transparent, and possibly more robust than external observation.

DEEP LEARNING

Deep learning is a type of ML that uses artificial neural networks to learn from data. [8] Neural networks are inspired by the human brain, and they can be used to perform a wide variety of tasks, including image recognition, NLP and speech recognition.

The term "deep" refers to the multiple layers within the neural networks used for processing information. These layers are organized in a hierarchical fashion, with each layer learning to extract higher-level features from the data. This allows the deep learning algorithm to learn increasingly complex patterns as it progresses through the layers.

Deep learning algorithms can be trained on a wide variety of data, including images, text, and audio. For instance, when a photo is uploaded to Facebook, deep learning models are used to identify people and objects in the photo and suggest tags. In language translation, for example, deep learning models aren't just able to identify the raw meaning of the words, but the relationship between those words and how they interact to form sentences, and appropriate grammar rules governing the sentence. If the model has been instructed to identify the indemnity clause in a contract, the model will know to also look for the term "claim" (even if it has not been instructed to do so), since it knows that "claim" and "indemnity" are related terms.

You can also read about the related concepts of ML, Supervised Learning, Unsupervised Learning, and NLP.

[8] "What Is Deep Learning? Applications & Examples | Google Cloud" (Google Cloud) <https://cloud.google.com/discover/what-is-deep-learning>.

ETHICS IN AI

With the increased use of AI, there is a need to carefully consider the ethical implications of using AI, particularly in the legal field.

Key ethical considerations of AI systems:

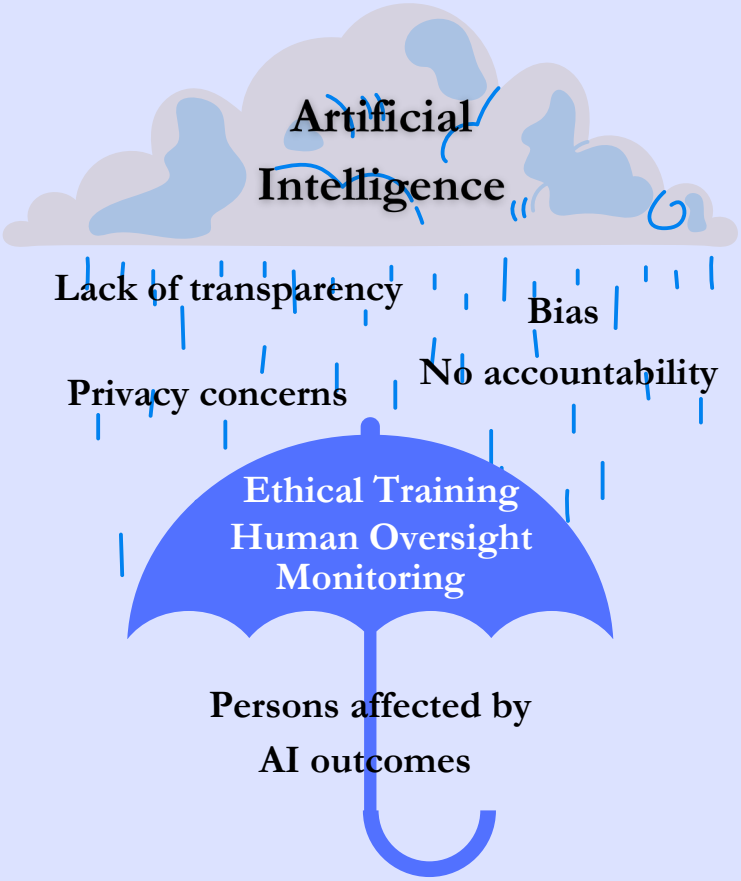
- **Transparency**: AI systems must be transparent. Both users and those affected by AI decisions should be able to understand how and why a particular result was reached. Ensuring transparency helps build trust, and helps legal professionals verify responses.
- **Accountability**: It is essential to establish human accountability for AI outcomes. When legal professionals use AI, they maintain ultimate responsibility for all responses, whether correct or erroneous. Verification of responses helps identify potential mistakes that legal professionals need to address, in order to maintain accountability while using AI. For instance, a New York attorney used OpenAI's ChatGPT for legal research, submitting a brief citing non-existent court cases, after failing to verify the LLM's responses. Ideally we would want AI systems to have built-in accountability so legal professionals can learn to trust and rely on the outputs they generate.
- **Bias**: AI systems are only as biased as the data they are trained on. Bias in data can perpetuate deeply ingrained societal biases, leading to unfair outcomes. It is crucial to regularly evaluate and mitigate biases within legal AI systems to avoid perpetuating discrimination or injustices. Here you can read more about bias.
- **Data Privacy and Security**: Legal AI algorithms often require access to large amounts of data. However today, rigorous data protection standards are mandated to safeguard client confidentiality and ensure that personal information is not misused or compromised. Accordingly, it is important to ensure compliance with those standards when training models.

To ensure maintaining the highest ethical standards, there are some guiding principles that can help strike a balance between AI usage and ethics.

- **Human Oversight**: Human involvement should remain integral, and legal professionals should always have the final say in making important legal decisions.
- **Continuous Monitoring and Auditing**: Regular monitoring and auditing of legal AI will help improve accuracy and reduce biases
- **Ethical Training**: Investing in education and training programs will enable legal practitioners to navigate the complexities of AI and make informed decisions.

Striking a Balance

With numerous benefits of AI come numerous ethical implications. Ensuring ethical behaviour and upholding ethical standards is a professional obligation. In order to strike a balance, legal professionals must continuously engage in conversations about transparency, accountability, bias, and privacy, to make the best use of AI while upholding the principles that guide our legal system.



FEEDBACK

Feedback systems are essential for improving the performance and accuracy of AI algorithms, particularly LLMs. They act as a continuous learning mechanism that helps address inaccuracies or biases that arise in outcomes by allowing AI algorithms to learn from real-world outcomes and continuously improve their performance.

Human feedback is a vital component of the feedback loop. Human feedback is deployed in the Reinforced Learning and Human Feedback (RLHF) approach, which is used by many LLMs and chatbots to improve answers. Legal professionals and end-users can provide insights that help refine and validate AI-generated results.

It is important for legal professionals to clarify whether legal tech tools provide an internal feedback mechanism (aside from the larger RLHF system that may be present in the underlying LLM) where incorrect responses from the tool can be highlighted and corrected, so the tool can improve its outputs. It is equally important to understand to what end the feedback provided is used - whether it is applied to all users of the legal tech tool, or just to the legal professional or just for the individual organisation.

Effective feedback systems in legal AI tools encourage collaboration among legal professionals, AI developers, and users of AI applications. By sharing experiences, insights, and best practices, the legal community can collectively improve the performance and reliability of AI systems. Collaboration can also foster a culture of continuous improvement while encouraging the sharing of knowledge which will benefit the entire legal industry.

GUARDRAILS

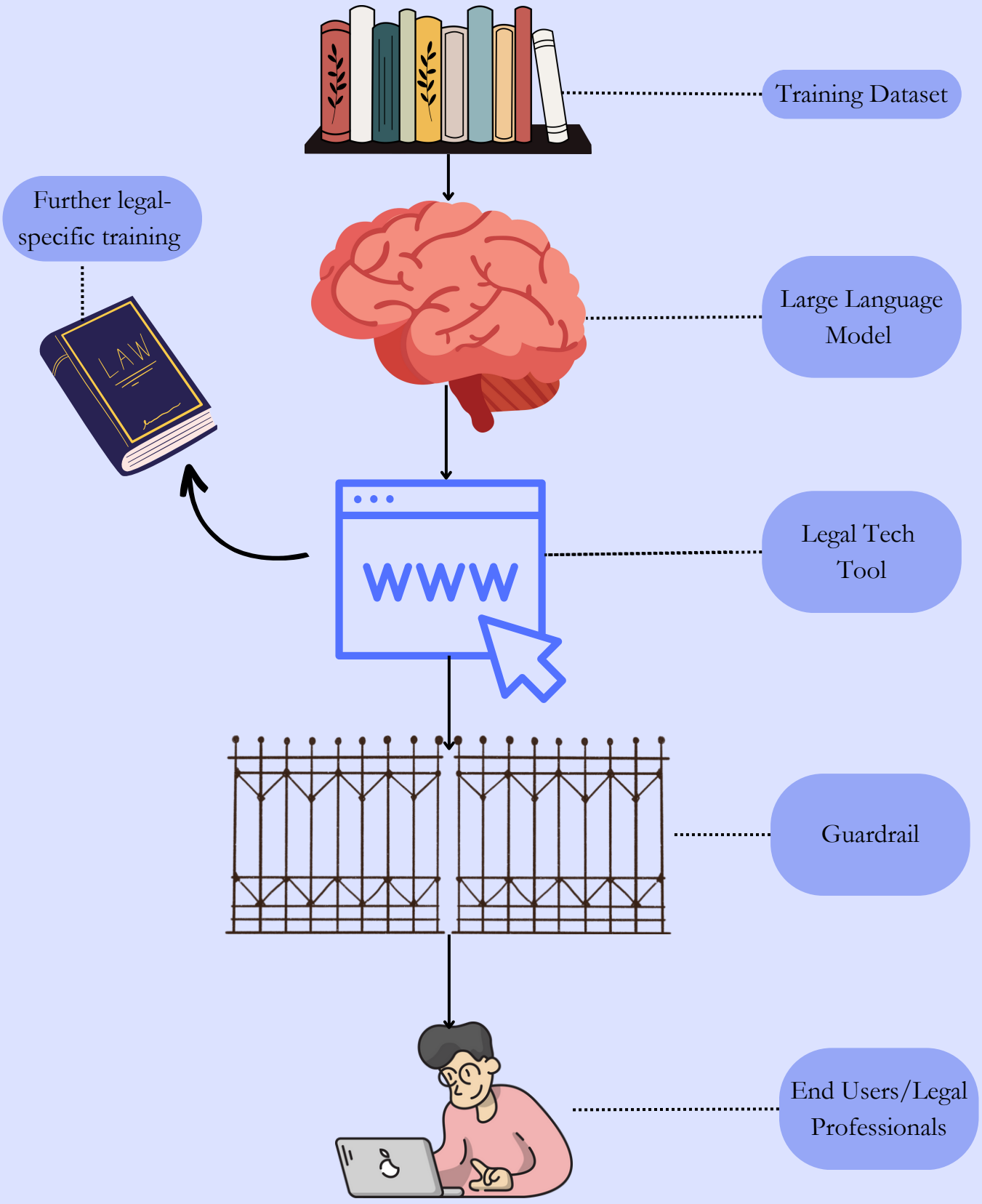
Guardrails in AI systems prevent LLMs from providing undesired responses. Undesired responses include unethical responses, hallucinations and providing a response by breaking data security parameters.

Asking an LLM how to tap into someone's phone could be considered a question that generates an unethical response. Hallucinations are when the LLM produces an incorrect or misleading response, like when GPT invented court cases. Accordingly, guardrails are safeguards placed within or over LLMs to ensure responses conform to the desired framework (like conforming to ethical standards, no hallucinations, and within security parameters).

Guardrails promote transparency, making sure that decisions made by AI systems can be understood and explained. This allows for accountability, ensuring that errors or misuse can be identified and rectified. In some cases, they encompass policies that mandate human oversight in critical decision-making processes. This is particularly important in high-stakes scenarios where AI mistakes could lead to significant harm, such as in autonomous vehicles.

Here are some specific examples of guardrails:

- Topical guardrails prevent AI systems from veering off into undesired areas. For instance, a legal chatbot might have a topical guardrail that prevents it from answering questions on how to execute criminal activities.
- Safety guardrails protect users from harm, like fact-checking responses to prevent hallucinations.



Where Guardrails Sit in the Legal AI Ecosystem

HALLUCINATIONS

AI hallucinations are when an LLM generates false or misleading information.

Hallucinations can be deviations from external facts, contextual logic, or both. [9] They often appear plausible because LLMs are designed to produce fluent, coherent text.

Hallucinations occur because LLMs are not conscious and have no unique perception of the world. Further, if the dataset it was trained on contains errors, or the algorithm itself is biased or inconsistent, or when the user prompt is unclear, the LLM is likely to hallucinate.

Hallucinations can take the form of sentence contradictions (generated sentences contradict each other), prompt contradictions (generated sentences contradict the prompt, inventions (when falsities are presented as fact) and random outputs (where generated sentences are not linked to the prompt).

The consequences of hallucinations are severe. User trust is betrayed and misinformation is more likely to be spread. In legal applications, hallucinations could result in incorrect information being generated and potentially recorded, which is disastrous to both the user and the persons who are affected by the information.

Guardrails serve to prevent hallucinations. Hallucinations are a result of black box AI, where it is unclear how and from where the LLM generated its response. Accordingly, guardrails help make the LLM accountable to its sources to help reduce or even prevent hallucinations.

[9] Ben Lutkevich, "AI Hallucination" [2023] TechTarget <https://www.techtarget.com/whatis/definition/AI-hallucination>.

INTEGRATING AI

AI has the potential to transform legal service delivery. Several legal tasks can be performed more efficiently by effectively incorporating AI in the process. Legal research, contract management, e-discovery are some examples where AI add lots of value.

That said, the deployment of any AI tool and successfully integrating it within existing legal service delivery processes, requires proper planning and execution. Here are a few pointers that can help.

1. **Identify the use case/ legal task**. Not all legal tasks lend themselves to AI application. So while AI cannot argue your case in court, AI can help identify which other cases are most relevant for your purposes. Take the time to spell out the use case clearly, and at which points in the process AI would be helpful.
2. **Scout out existing tools**. There are several AI empowered legal tech tools in the market, you need to figure out if any of them suit your requirements. Explain your use case to the vendor and request for a demo for that use case.
3. **Rolling out an AI tool**. There are several things to consider during the roll out.
    a. AI is only as good as the data it has access to, so if the tools require access to your proprietary data, spend the time cleaning your data to ensure that it is organised and machine readable.
    b. What other existing tools of your organisation do you need the AI tool to integrate with? Plan for that integration and any consequent data migration.
    c. Train your internal teams on the basics of AI software, its capabilities and limitations, in addition to the specifics of using the tool.
    d. Plan for change management internally. If the tool significantly impacts the existing way of working, you will need the buy-in of the users to consciously embrace the change.

Planning for the above helps in effectively integrating AI in your legal processes.

JUSTICE:ACCESS THROUGH AI

Access to justice is a fundamental principle underlying any democracy. Sadly today, in several countries, there is a huge gap between citizens that can access justice, because they have the means to do so, and those that cannot. Complex legal jargon, delayed justice delivery from courts, expensive legal advice are some contributing factors for this divide.

AI has the potential to improve access to justice in a number of ways and some of these are discussed below.

1. **Reducing the knowledge gap**: AI can help provide legal information and advice to people who cannot afford or access traditional legal services. AI-powered chatbots and legal websites can provide people with basic legal information and advice on a variety of topics, such as divorce, bankruptcy, and housing law. Further, AI-powered tools can help people understand complex legal procedures and complete legal forms and documents.
2. **Identification of criminals**: Image recognition AI can help with identification of criminal suspects through CCTV footage.
3. **Prediction of outcomes**: If the AI is trained on historical jurisprudential data, it will be equipped to provide a probabalistic outcome of a case at a specific court. The AI will be able to do this faster than a lawyer, as it can digest volumes of data quickly.
4. **Document review**: AI tools can review large datasets and flag transactions that appear fraudulent. Governments can use such tools to analyse financials statements of companies and individuals to detect potentially fraudulent transactions.

KNOWLEDGE-BASED SYSTEMS

Knowledge-based systems (KBSs) are a type of AI that use a large amount of data and rules to solve problems. In legal applications, KBSs can automate tasks such as legal research, document review, and case analysis.

KBSs typically have three components: a knowledge base, an inference engine, and a user interface. The knowledge base contains the data and rules that the KBS uses to solve problems. The inference engine uses the knowledge base to locate the data. This is similar to the way a search engine works. The user interface allows users to interact with the system.

KBSs can be used to automate a variety of tasks in legal technology, such as searching through large amounts of legal documents to find relevant information, and analysing legal documents to identify inconsistencies.

1 Knowledge Base

2 Inference Engine

3 User Interface

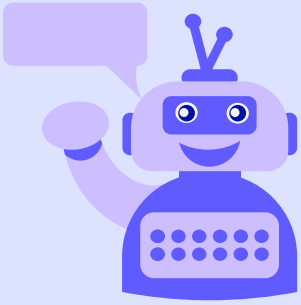
Components of a Knowledge-based System

Although KBSs are a type of AI, they differ from newer AI developments in their approach. Newer AI developments use top-down systems that detect patterns in data using statistical methods, such as big data, deep learning, and data mining. These systems also include neural network systems that use deep learning to solve pattern recognition problems, such as facial recognition. In contrast, KBSs handle large amounts of unstructured data and integrate knowledge based on that data on a large scale.

Feature	KBS	Newer AI developments
Approach	Bottom-up	Top-down
Data	Unstructured	Structured
Pattern recognition	No	Yes
Problem-solving	Rule-based	Statistical
Examples	Medical diagnoses tools	ChatGPT, self-driving cars

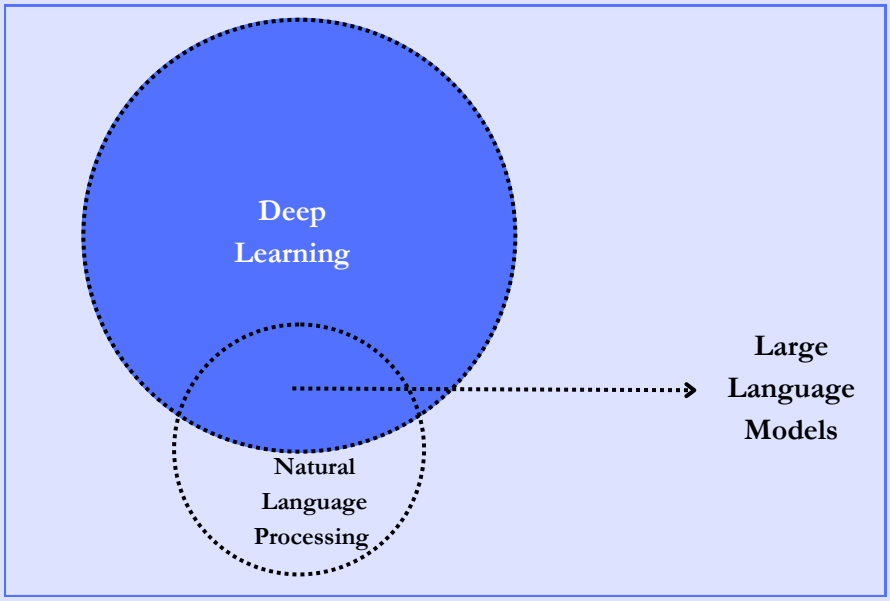
Differentiating Knowledge-based Systems with newer AI developments

LARGE LANGUAGE MODELS



A Large Language Model (LLM) is a trained deep-learning model that understands and generates text in a human-like fashion [10]. LLMs can recognize, summarize, translate, predict, and generate content using very large datasets.

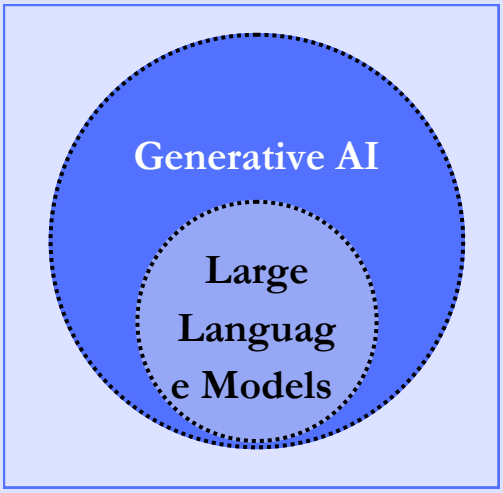
In the AI ecosystem, LLMs sit at the juncture of Deep Learning and NLP. It is a type of ML model that can perform NLP tasks like generating text, classifying texts and answering questions in a conversational manner. While ML is focused on learning from data, LLMs generate new data.



LLMs in the AI Ecosystem

[10] Adrian Tam, "What Are Large Language Models" [2023] MachineLearningMastery.com <https://machinelearningmastery.com/what-are-large-language-models/>.

Often, the terms Generative AI and LLMs are used interchangeably. While they are comparable concepts, there is a subtle difference. They are both types of AI, but they have different goals and applications. Simply put, Generative AI is the larger subset within which LLMs reside. Generative AI includes generating content that could be in the form of text, images, audio and video, while LLMs typically generate content in the form of text.



Generative AI and LLMs

Generative AI aims to mimic and enhance human creativity, while LLMs are specifically designed to comprehend and produce text-based content. LLMs are trained on massive amounts of text data, which allows them to perform tasks such as text summarization, question answering, and translation. Generative AI can also be used for more creative tasks such as image generation. Examples of models that generate text include OpenAI's GPT and Google's PaLM, whereas OpenAI's DALL-E and Midjourney can generate and create new images.

Chatbots are used to interact (provide instructions and receive responses) with an LLM. Sometimes, a chatbot is owned by the developer of the LLM, and sometimes a chatbot can be built by you on top of the LLM. Some examples of LLMs with corresponding chatbots are set out below.

Developer	LLM	Chatbot
OpenAI	GPT	ChatGPT
Google	PaLM	Bard
Anthropic	Claude	Claude.ai

LLMs and Corresponding Chatbots

MACHINE LEARNING

Machine Learning (ML) is a subset of AI that is centred around data. It uses algorithms and statistics to find patterns within huge datasets [11].

ML is widely deployed in real life settings. The YouTube recommendations page, for instance, is powered by ML to understand user preferences based on clicks and views, based on which it will generate recommendations that the user is most likely to watch. Virtual assistants like Amazon Alexa and Apple Siri use ML to interpret natural speech and supply context. If a person vocally instructed Alexa to play "Hey Jude", ML would help Alexa understand that the request is to play a song.

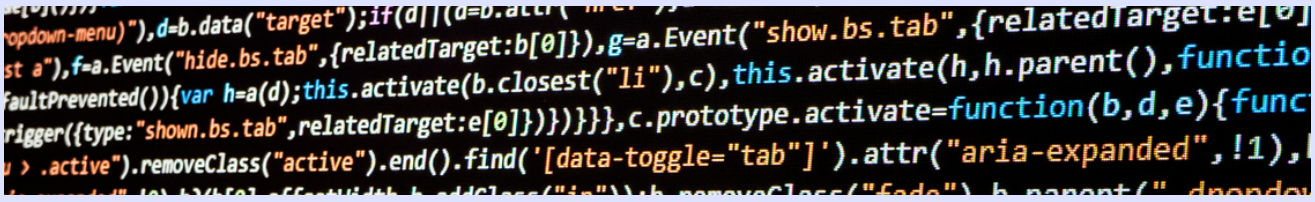
ML-based systems in the legal industry are fine-tuned on legal data including jurisdiction-specific statutes, case law, industry-specific clauses, policies, rules and procedures. An ML algorithm learns from this dataset and applies its knowledge to new data to discover patterns, and then automate tasks. This knowledge empowers it to identify anomalies. For example, in the context of contract review, the ML algorithm fine-tuned on commercial contracts understands that a termination clause typically exists. Accordingly, it is able to identify contracts where a termination clause is absent, as this is an anomaly from the identified patterns.

Outside of law, ML has successfully been applied to automate tasks that were once thought to require human intelligence, like language translation, fraud-detection, driving automobiles, facial recognition, and data-mining [12].

There are three primary types of ML – Supervised Learning, Unsupervised Learning and Reinforcement Learning.

[11] Exigent, "Why Machine Learning Is Key to Legal Transformation | Exigent" (Exigent, November 24, 2020) <https://www.exigent-group.com/blog/why-machine-learning-is-key-to-legal-transformation/>
[12] Harry Surden, "Machine Learning and Law" (Colorado Law Scholarly Commons) <https://scholar.law.colorado.edu/faculty-articles/81/>

NATURAL LANGUAGE PROCESSING



Natural language processing (NLP) is a branch of AI that gives computers the ability to understand and process human language.

LLMs are a type of machine-based NLP system that are trained on massive amounts of text data. This allows them to perform tasks such as text summarization, question answering, and language translation.

NLP has a number of legal use cases, including:

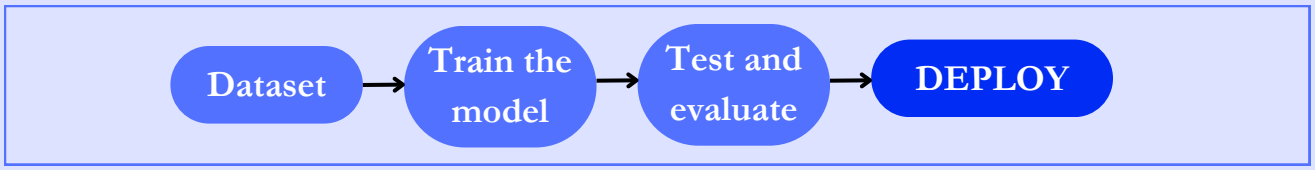
- **Legal research**: NLP can be used to find information relevant to a legal judgement by searching through large amounts of legal text.
- **eDiscovery**: NLP can be used to determine the relevance of documents to an information request by analyzing the text of the documents.
- **Contract review**: NLP can be used to check that a contract is complete and avoids risk by analyzing the text of the contract.
- **Document automation**: NLP can be used to generate routine legal documents by filling in templates with the relevant information.
- **Legal advice**: NLP can be used to provide tailored legal advice by answering questions from clients in a natural language format.

ON/OFFLINE LEARNING

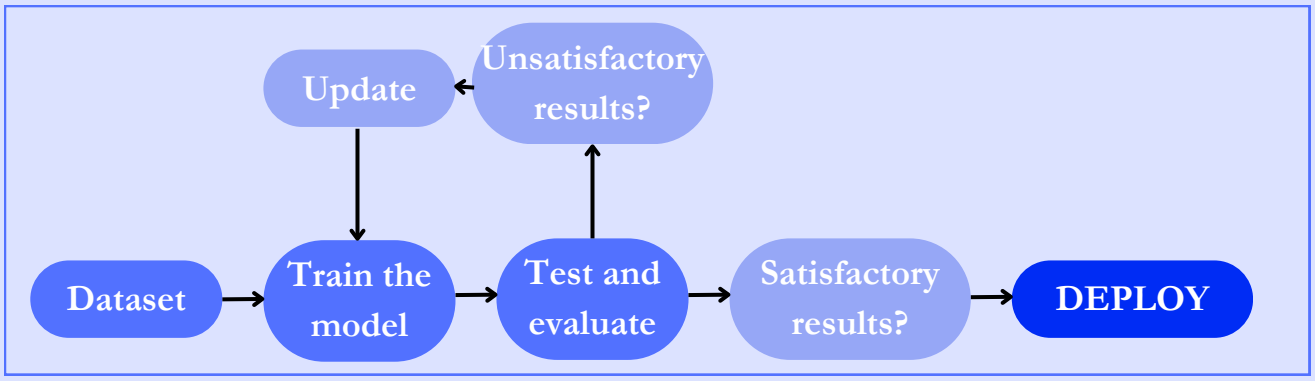
Online and offline learning are two different ways for reinforcement learning models to learn.

In **online learning**, the model learns from data as it is received. This means that the model is constantly updating its knowledge as new data comes in. Online learning is often used for systems that need to adapt to changing conditions, such as video streaming analytics and weather forecasting. Legal research tools typically use online learning techniques since there is always new jurisprudence and legislation being created.

In **offline learning**, the model learns from a dataset that is pre-collected. Once the model is trained, it is deployed to a live environment. Offline learning is often used for systems that do not need to adapt to changing conditions, such as image recognition and classification. A legal summarisation tool, for instance, uses offline learning techniques given that it does not need the most up-to-date legal knowledge to summarise text presented to it.

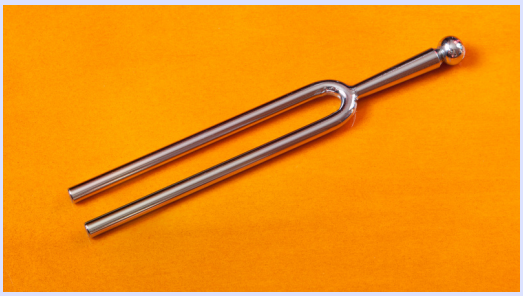


Online Learning Process Flow



Offline Learning Process Flow

PRE-TRAINING & FINE-TUNING



Pre-training and fine-tuning are two stages of transfer learning.

Transfer learning is an ML technique where knowledge of an already trained ML model is applied to a different but related problem. Transfer learning is able to reduce the need for large volumes of labelled data and speeds up the training time. For instance, an ML model that is pre-trained on vast amounts of data to summarise text can then be exposed to a smaller volume of legal data, which can be used to summarise legal text, like a court case or a statute.

Pre-training

Pre-training is when a model is trained on large datasets to accomplish a specific task. For instance, LLMs have been pre-trained on open source data. The term is often used interchangeably with training, which could be confusing. In fact, the term 'training' is often generically used to cover instances of pre-training and fine-tuning.

Fine-tuning

Fine-tuning is when the pre-trained model is further trained on an even more specific task typically using labelled data by making small adjustments to achieve desired outputs. It is used to speed up the training as it already contains vital information from a pre-existing deep learning algorithm.

For legal teams looking to harness the power of an LLM, know that you have to further train the base LLM model on legal data. You could do this by using LLMs from legal tech companies that have exposed the LLM to generic legal data (legislation, caselaw or contracts) or by exposing the LLM on your internal proprietary data, or both. Each stage of data exposure is 'training' the LLM.

QUERY EXPANSION

Query expansion is a technique in AI and information retrieval that broadens the range of relevant results. It operates by adding synonyms, related terms, or conceptually similar words to the user's original search query. This approach ensures that even if the exact words in the query don't match the terms in the documents, relevant information is still retrieved.

Legal documents often employ intricate language and specific terminology, which can lead to challenges in information retrieval. By expanding the initial query, these tools could uncover a broader array of legal documents that might contain relevant insights. For example, if a user inputs the phrase "copyright infringement" while using a legal research tool, query expansion determines that "fair use" is also a relevant search and includes searches for "fair use", despite the user not inputting the phrase.

It is worth noting that query expansion often includes search terms that may introduce false positives and outliers (for instance, including results for "plagiarism" when the initial query was for "piracy"), thereby making a lawyer's job more challenging as they now have more information to sift through. It is up to the user to determine the suitability of query expansion techniques for their specific use case.

REINFORCEMENT LEARNING

Reinforcement learning is an ML technique where the algorithm learns by trial and error to achieve a desired outcome.

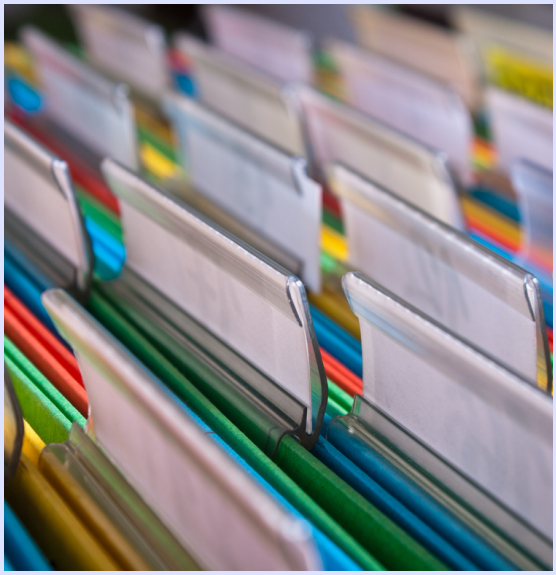
The algorithm is rewarded for taking actions that lead to the desired outcome, and penalized for taking actions that do not.

It is similar to how dogs are trained. Much like Pavlov's experiment where dogs were conditioned to salivate at the sound of a bell since they were presented with food at the same time, a dog is rewarded with treats for performing desired behaviours and penalised for performing undesired behaviours. Over time, the dog learns to associate certain behaviours with rewards and punishments. The reinforcement learning algorithm learns in a similar way, by associating certain actions with rewards, and finding solutions that allow it to obtain the most rewards.

Reinforcement Learning with Human Feedback (RLHF) combines reinforcement learning with human guidance. In RLHF, the algorithm is given human feedback on its actions, which helps it to learn more quickly and efficiently. RLHF is often used to train LLMs to generate text that is both engaging and accurate. Several LLMs like GPT and Claude have been trained using RLHF, among other techniques.

This is also why legal tech companies that use an underlying LLM in their tool, fine-tune the LLM on legal data, where their legal experts review and provide feedback on the responses provided by the LLM. In fact, some legal tech tools go a step further and enable their end users to provide human feedback to further improve the quality of the responses. The tool takes this human feedback and learns using reinforcement learning techniques.

SUPERVISED LEARNING



Supervised learning is a type of ML where the algorithm is trained on data that has been labeled with the desired output. This involves assigning labels and categories to the data to instruct the machine on identifying certain patterns. The model is trained to discern the inherent patterns and correlations between the input data and the output labels. This process equips the model to accurately assign labels when confronted with new data. In essence, the machine learns from labelled data which it uses to assign labels to new data.

For instance, supervised learning is the mechanism behind the Netflix recommendations algorithm. It is trained to offer show recommendations aligned with user preferences gathered from the thumbs up/ down feature. As users engage by indicating what shows they like and dislike, the predictive capabilities of the Netflix algorithm improve, leading to better forecasts of user preferences.

Supervised learning carries significant implications in the legal realm as well. In contract review, supervised learning helps users to locate pertinent information within contracts to address specific queries. For instance, when users ask the tool to identify limitation of liability clauses that extend beyond the termination of the agreement, the tool can flag this information for a large volume of contracts as it has been trained on existing limitation of liability of clauses. Additionally, supervised learning aids legal researchers in identifying pertinent laws, legal precedents, and other documents encompassing crucial legal principles articulated in judicial opinions.

TRAINING

Training an AI involves instructing an AI system to comprehend, interpret, and derive insights from data. After being trained, the AI will be able to draw inferences and conclusions on new data/ inputs based on the information it has acquired through the training.

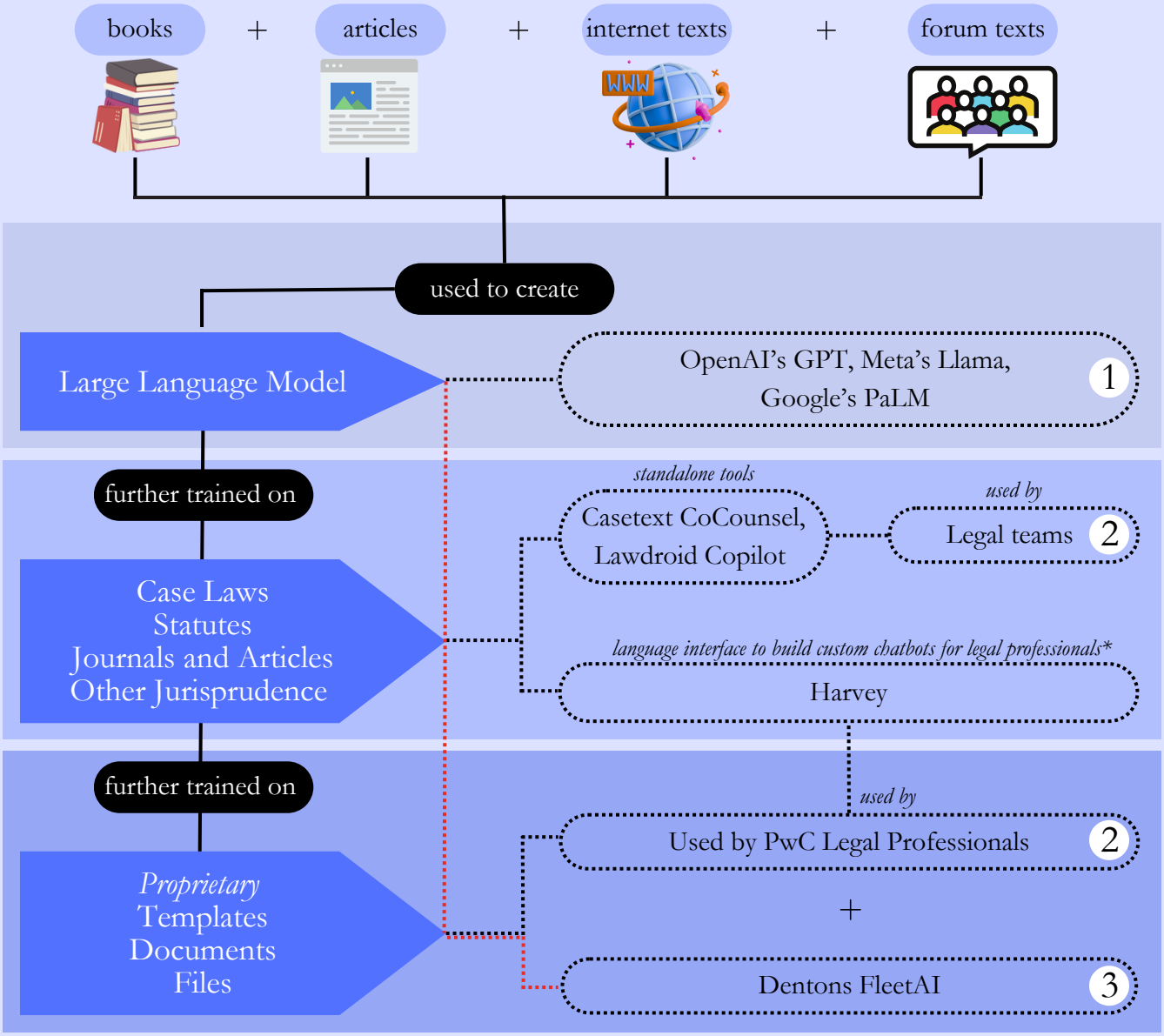
ML algorithms acquire knowledge from data, discern patterns, generate understanding, and make decisions in light of the training data provided to it. The quality of the training data profoundly influences the accuracy with which the model carries out its tasks. Higher volumes of data employed for ML training increases the probability of more variety in responses, thereby possibly reducing instances of bias.

For instance, most LLMs are trained on a mix of publicly available data, comprising various websites on the internet including public forums to understand human dialogs, books, articles, and non-public data sourced from third parties.

When legal tech tools deploy AI, specifically LLMs, to provide drafting, summarisation and research capabilities, these tools are typically further trained and fine-tuned on legal data. Depending on the tool and the jurisdiction, the data is typically a mix of case law, statutes, rules, regulations and legal documents such as contracts.

Legal teams can also choose to go a step further and deploy LLMs specifically for their team (whether in a firm or in a corporate legal department). Here, the training dataset is expanded to include the team's proprietary data including clauses, contracts and templates.

In the image below, there are multiple approaches to using an LLM for legal work.



Training LLMs

1. Use the LLM directly with a generic chatbot (like ChatGPT/ Claude.ai). Note that you would need to ensure security parameters are met.
2. Use an LLM that has been exposed to generic legal data (case law, legislation, contracts etc). There are some legal tech companies that have legal datasets for this purpose. You can directly use that on your proprietary data and build your own chatbot. Or some legal tech companies go a step further and provide a chatbot interface that can then be directly applied to your proprietary data.
3. Use an LLM directly on your proprietary data and build your own chatbot.

The point is that there is no single way of training an LLM for your purposes. Identifying the use case for your organisation will help in selecting what approach works best for that use case.

*(*based on publicly available information on LLMs and legaltech companies)*

UNSUPERVISED LEARNING

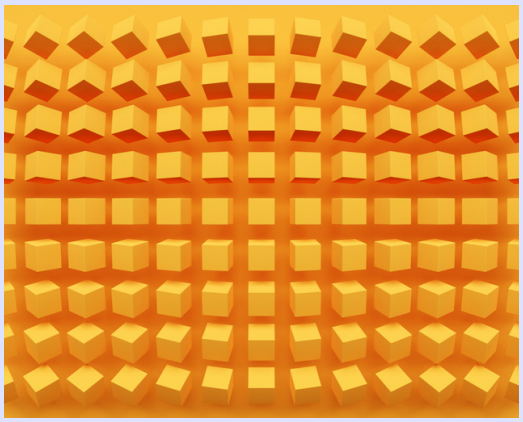
Unsupervised learning is the opposite of supervised learning. It is an ML technique that uses AI algorithms to identify patterns in data sets that are not classified or labelled. Subsequently, the machine looks for whatever patterns it can find and groups the data according to those patterns.

The lack of supervision makes it an ideal ML technique for discovering groupings and differences in unstructured data according to similarities and differences even though no categories are provided. It is well-suited for processes such as customer segmentation or image recognition.

For instance, if an algorithm is asked to differentiate between images of tigers and leopards, it will first classify the images into categories such as those with stripes and those with spots. This classification continues with increasingly specific sub-groups such as size of body, length of limbs, facial features etc. until the algorithm satisfactorily differentiates between the two animals.

For legal applications like document review, the unsupervised ML will cluster similar documents or clauses along with clear anomalies from those groups. This would reduce time spent on manual review, and attorneys could possibly only have to spend time reviewing clauses that are classified as outliers – those that are different to others.

VECTORS



Vectors are mathematical representations of attributes. They help in indexing data in granular detail, thus making it easier and quicker to search the data.

For instance, a movie streaming app like Netflix is likely to have a vector database of its content. This will include details like genre of movies, (horror, comedy, romance, etc.) PG rating (child or adult friendly), language and actors. These are all attributes of a movie and are stored in a vector format so that it is easier for a computer to process. Accordingly, it is easy for the algorithm to recommend similar movies that you may like based on movies you watch, as the algorithm identifies similarities of the vectors across the movie database.

The same example can be used in legal research, when searching for relevant case laws that are applicable to your present matter. Searching using a vector database will be far more efficient than the traditional keyword or Boolean searches as the algorithm will be able to more closely identify relevant case law based on vectors similar to the matter at hand.

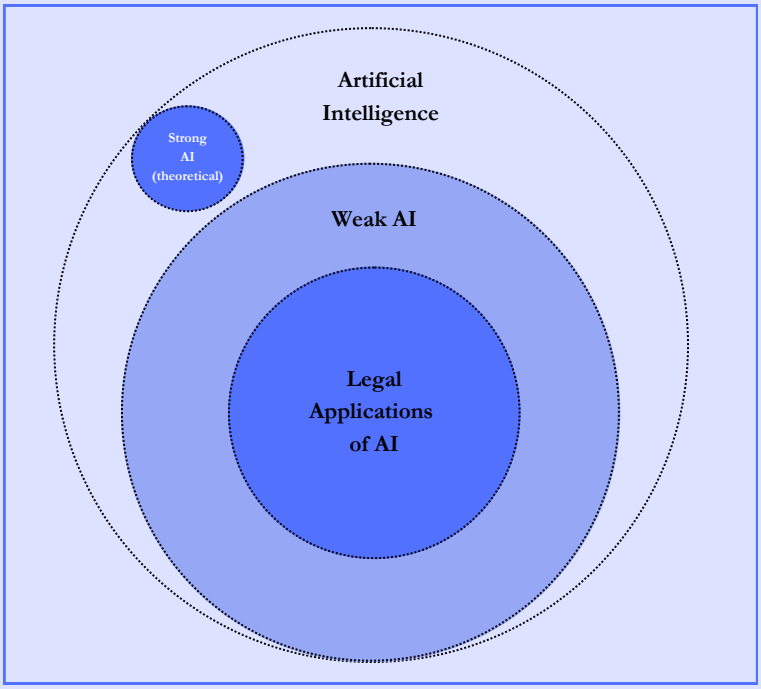
In short, using vectors on your dataset and then applying a search, will bring you relevant results much quicker than the traditional search methods. Vector databases allow models to more accurately understand data such as texts, images or videos by drawing comparisons and identifying relationships between the words. Storing data in vectors allows models to search the data based on how similar the data is, instead of exact matches, allowing for contextual understanding of the data.

WEAK & STRONG AI

AI has two categories based on what the AI system is capable of doing.

Strong AI (or General AI) is an AI system which has reached a level of human intelligence in every way - the system is able to learn, perceive and function completely like a human being. No such systems exist yet, and the concept is purely theoretical.

Weak AI is an AI system that can perform specific tasks, like extracting a legal concept from a contract. They are not meant to mimic human intelligence, but instead automate precise tasks. Despite its dismissive designation, Weak AI is anything but weak. All forms of modern AI are classified as Weak AI, including chatbots, image recognition systems, autonomous vehicles and recommendation engines.



Where Weak and Strong AI sit in the AI Ecosystem

XAI: EXPLAINABILITY

Explainable Artificial Intelligence or XAI is a set of processes and methods that allows human users to comprehend and trust the results and output created by ML algorithms. [13]

ML models are trained using vast amounts of data, and it may be difficult to verify the accuracy of every source. Further, the complexity of technology used in ML models make it difficult to interpret by humans. AI models are prone to various forms of bias, and performance can vary when specific fine-tuning data (such as legal specific data) differs from training data.

Explainable AI enables organisations deploying AI to gain access to the AI technology's underlying decision-making process, by which they are able to make adjustments. This helps build trust and confidence with users, and helps organisations adopt a responsible approach to AI usage with auditable AI models. It also mitigates compliance, legal, security and reputational risks of AI.

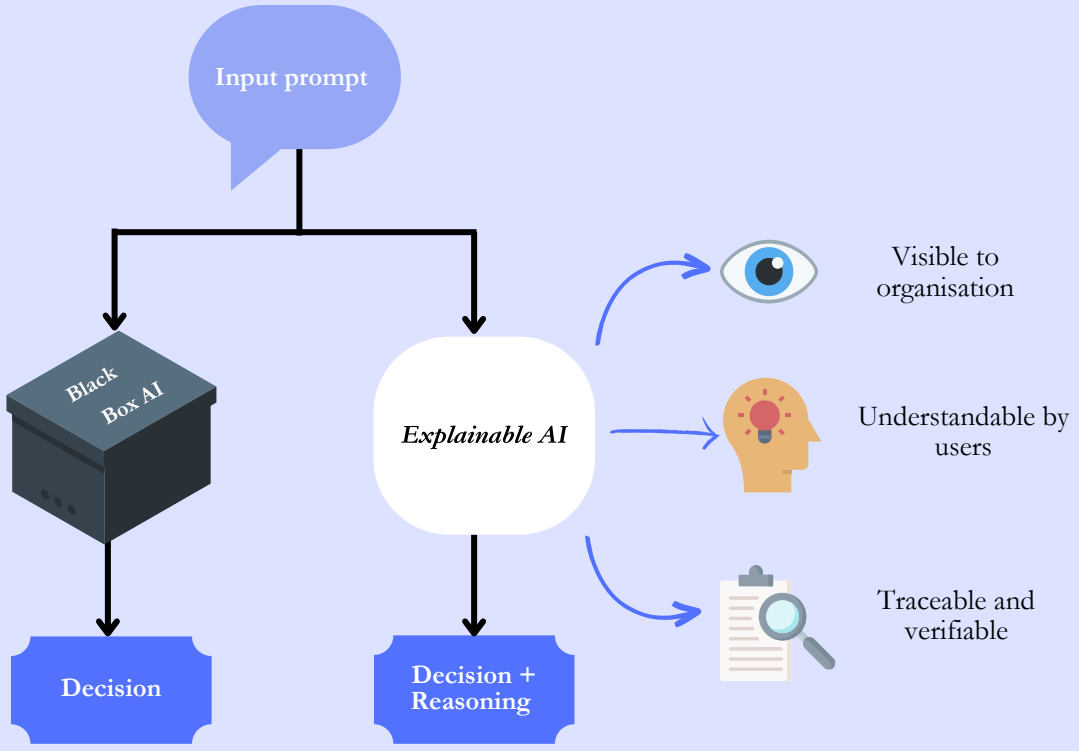
As referenced earlier, a New York attorney used OpenAI's ChatGPT, based on its GPT models, for legal research, submitting a brief citing non-existent court cases after failing to verify the LLM's responses. This is because OpenAI's GPT models are not directly explainable to end users - it operates as a black box, where it provides answers without justifying its sources. If the attorney used Explainable AI and asked the LLM to cite its sources, the answer would point to the specific court case, statute or other sources to back up its claims. This refers to traceability, which is a part of the explainable AI techniques. One specific example is using Perplexity AI, a chatbot interface that operates over ChatGPT and Claude, but provides the sources for its responses.

Another example is how insurance companies employ AI to calculate the amount

of premium payable on certain policies. Sometimes a formula function is provided on their website to calculate the amount of premium payable after keying in certain information. Nonetheless, it is not clear what was the basis for arriving at the final premium amount as the formula for that calculation is not disclosed, its a black box. However, if the insurance company were to use Explainable AI instead, they would disclose the basis of the formula. That could include statements like

- Tabular indication of premium payable age wise
- If you are a heart patient a weightage of x is applied to the regular premium payable.
- If you are diabetic, a weightage of y is applied to the regular premium payable
- If you are a smoker, your premium becomes 20% higher.
- If you have more than two of the above conditions, we will not be able to insure you.

The insurance regulator in fact, may require insurance companies to ensure that it provides cogent reasons when denying insurance to an individual as well as explaining why the premium appears discriminatory to different individuals of the same age. Using Explainable AI helps you comply with this requirement.



Explainable AI

[13] "What Is Explainable AI? | IBM" <https://www.ibm.com/topics/explainable-ai>.

YOLO:
YOU ONLY LOOK ONCE

YOLO (You Only Look Once) is a real-time object detection algorithm. Essentially, it helps you identify objects in an image very quickly because it uses neural networks to detect objects in a picture.

YOLO has a number of potential applications in law, particularly in law enforcement. YOLO can be used to detect license plate numbers of vehicles that commit traffic violations, and be able to identify motorcyclists riding without helmets.

Object detection with the YOLO algorithm is a two-step process. It uses a bounding box, i.e., an imaginary rectangle serving as a reference point, to locate the existence of objects in an image, and then assigns classes to the objects found. In the motorcycle example, riders with helmets would be assigned to one class and riders without helmets would be assigned to another class.

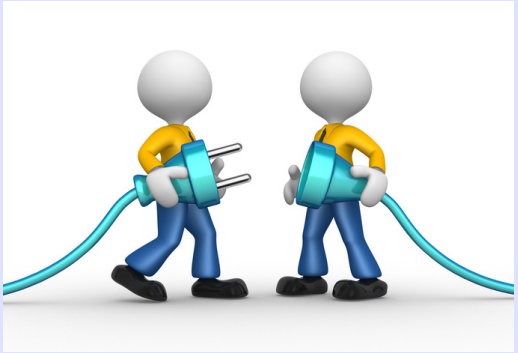
YOLO has two main advantages. Firstly, it is much faster than other object detection algorithms, making it suitable for real-time applications like video surveillance and self-driving cars. Secondly, YOLO uses global reasoning. This means that it makes predictions by considering the entire image at once, unlike other object detection algorithms that scan the image in smaller parts. This allows YOLO to learn contextual relationships between objects and their appearance, which improves its accuracy.

For instance, if YOLO is detecting a car in an image, it will take note of the surrounding environment like the road, other vehicles and pedestrians. This helps it better understand the context of the car like the car's location, size and distance in relation to other vehicles.

.

ZERO-CODE AI

Zero-code AI, or no-code AI, are tools that allow anyone to create AI applications without writing technical code. In lieu of coding, the core functionalities are accessible through visual interface, often alongside pre-built templates and integrations with other tools.



Traditionally, AI has the potential to transform businesses, but at a steep cost given that AI integration requires high investment in infrastructures. However, Zero-code AI allows small to mid-sized businesses to use AI to improve delivery of services. For instance, a law firm can provide guidance to companies on eligibility for government-related business loan schemes, by simply creating a checklist-style plug-and-play tool whereby clients have to answer relevant questions ("where are you located", "how many employees do you have" etc.) to narrow down criteria until the tool shows them their eligibility status for specific schemes.

Thus far, law could only be practiced by lawyers that possess the relevant licence, but with the rise of zero-code tools, regulators will have to evaluate the extent to which models and machines can provide legal services. The issuance of alternative business licenses in some countries like the UK and some states in the USA are indicative of regulatory reform that will help democratise access to legal advice.

There are several advantages to zero-code AI

- Accessibility - low investment to enable wider adoption among smaller organisations
- Usability - plug and play features allow AI implementation within a budget
- Speed - simple an intuitive tools allow for speedy implementation
- Scalability - solutions can easily be scaled for many users

FOLLOW WALCHET

YouTube | Spotify | Apple Podcasts | LinkedIn | Website | Priyadarshini Shetty

We would love to hear from you.
Write in to us at **priyadarshini@walchet.com**