

Machine Learning & The Future of Law

WHAT IS MACHINE LEARNING (ML)?

“Machine learning is a subfield of artificial intelligence, which is broadly defined as the capability of a machine to imitate intelligent human behavior.”¹

Machine Learning (ML) primarily serves three crucial functions that leverage data to extract meaningful insights and drive informed decisions. Firstly, the descriptive function utilizes data to provide explanations of past events, answering the question of "what happened?" This analysis is essential for understanding historical behaviors and trends which can provide a foundation for further examination. Secondly, ML extends into the realm of prediction, where data is employed to attempt to forecast future occurrences, addressing the query of "what might happen?" This predictive function is instrumental in anticipating potential scenarios and preparing for them accordingly. Lastly, the prescriptive function of ML takes the analysis a step further by using data to suggest viable courses of action, tackling the question of "what to do?" This function is pivotal for decision-making, as it provides recommendations based on analyzed data, aiding in navigating complex situations and optimizing outcomes.

REAL WORLD SCENARIO: ML CONTRACT ANALYSIS

In a real-world scenario involving Machine Learning (ML) for contract analysis, a law firm is enlisted to represent and carry out due diligence for a company engaged in executing a high volume of data processing agreements with various vendors. Customarily, a standard service agreement is utilized for these transactions. However, at times, vendors propose custom agreements, presenting a need for a detailed analysis to ensure compliance and safeguard the interests of the client.

Addressing this need, the law firm partners with a third-party consulting firm specializing in ML. This consulting firm crafts a tailored ML model capable of meticulously scrutinizing these custom agreements. The model is designed to identify clauses that could potentially shift liability onto the client in the event of third-party data breaches. It also flags the absence of a standard arbitration clause, an essential element for resolving disputes. Moreover, the model detects clauses falling short of the client's data protection standards and those that automatically renew the agreement without explicit consent, which could potentially lead to undesired long-term commitments.

With the implementation of this ML model, the law firm significantly enhances its contract review process. It can now promptly identify contracts that necessitate further scrutiny, enabling a more efficient allocation of resources. The model also aids in listing vendors whose relationships

¹ Sourced and adapted from <https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained>

warrant a reassessment to avert unexpected complications. Consequently, the law firm can deliver its services faster and more efficiently, which is beneficial not only for the current client but sets a precedent for handling future clients with similar needs. Furthermore, the law firm sees an opportunity for continual improvement. The insights gained from utilizing the ML model pave the way for refining the model further, reducing error rates, and expanding its capability to serve a broader client base with varying contract analysis needs. Through this collaborative and tech-driven approach, the law firm significantly elevates its contract analysis and due diligence processes, demonstrating the potent capabilities of Machine Learning in legal practice.

What are some challenges arising from developing and using such model?

The development and use of a ML model for contract analysis in a legal framework comes with its own set of challenges, primarily revolving around privacy, language ambiguity, and computational as well as financial resources.

Firstly, contracts are laden with personally identifiable information (PII) and sometimes sensitive personally identifiable information (SPII), alongside crucial business data. This necessitates the application of anonymization techniques during data preparation to comply with privacy regulations and mitigate associated risks. Ensuring privacy compliance while retaining the data's utility for analysis is a complex endeavor that requires a meticulous approach and may entail additional processing time and resources.

Secondly, contracts are notorious for their ambiguous language, which can be subject to varying interpretations. This ambiguity poses a significant challenge when labeling data, as inconsistencies and discrepancies can easily arise. A robust ML model for contract analysis would need to navigate the nuanced language of legal documents accurately. The challenge of ambiguous language not only impacts the data labeling phase but also the model's ability to provide precise and reliable insights, which is critical in a legal setting where stakes are high.

Lastly, the computational and financial needs for developing, deploying, and maintaining an ML model in production are substantial. The endeavor demands a considerable investment of effort, expertise, and financial resources. Particularly, finding ML talent with a solid understanding of the legal industry can be a rare commodity, thus making it a costly affair. The blend of legal and technical expertise is crucial for developing a model that accurately interprets legal contracts and provides valuable insights. Moreover, the ongoing costs of maintaining, updating, and improving the ML model to adapt to evolving legal standards and contract practices add to the financial burden.

ML APPLICATIONS FOR LITIGATION LAWYERS: A REAL WORLD-SCENARIO

A law firm specializing in litigation often grapples with the enormous task of sifting through vast amounts of discovery documents to build or defend cases for their clients. The documents can vary widely, from emails and contracts to financial records and other formal documents. The traditional manual review process is time-consuming and prone to human error.

To address this challenge, the law firm engages a third-party consulting firm proficient in ML. The consulting firm develops an ML model tailored to automate the document review process. The model is crafted to:

- Identify and flag potentially relevant documents based on predefined criteria, such as keyword occurrences, thematic consistency, and contextual relevance.
- Recognize and highlight discrepancies or inconsistencies in the document sets that may signify erroneous or fraudulent information.
- Automate the categorization of documents into predefined buckets based on their content and relevance to different aspects of the case.
- Extract critical pieces of information such as dates, names, and financial figures that are pertinent to the case.

With the deployment of this ML model, the law firm experiences a paradigm shift in how it handles document review for litigation. It can now:

- Rapidly pinpoint relevant documents amidst a sea of data, drastically reducing the time spent on the review process.
- Identify potential issues or inconsistencies more accurately, thereby strengthening their case strategy.
- Allocate their human resources more effectively, freeing up their lawyers to focus on more strategic aspects of litigation.
- Continually refine the model based on feedback and new case experiences, progressively reducing error rates and improving efficiency.

INSIGHTS: TECHNOLOGY & THE PRACTICE OF LAW

Excerpt from: [AI is Reinventing the Legal Industry](#), Morgan Beller & James Currier, August 2023.

Why is legal leading the rollout of AI?

There are some industries that appear to be natural fits for AI disruption. It's a matter of two variables: the percentage of a task that can be accomplished by AI, and the value of the task it can accomplish.



Now that we have LLMs that are exceptional at reading, writing, and analyzing, the business areas where they apply most are fields like law, customer service, software development, marketing, and sales.

The legal profession is built upon language data. The only product of legal is words. LLMs crack open language-based data the way that other forms of AI have been predicting patterns from numerical data. LLMs identify patterns in that language data and can offer up nearly complete products to lawyers and their clients.

That's the same thing that great lawyers do. Language inference and generation based on years of study.

You can ask a great lawyer if a certain privacy policy is compliant, and the best can give you an answer instantly. At this point, so can an intelligent AI trained on a corpus of high quality legal data. You can ask a great lawyer to draft a contract, and they can do so in a matter of hours thanks to their analytical skills and intuition informed by past experiences. So can an LLM trained on a dataset of similar contracts.

But it goes beyond just that. AI doesn't just replicate the skills of great lawyers. It improves them.

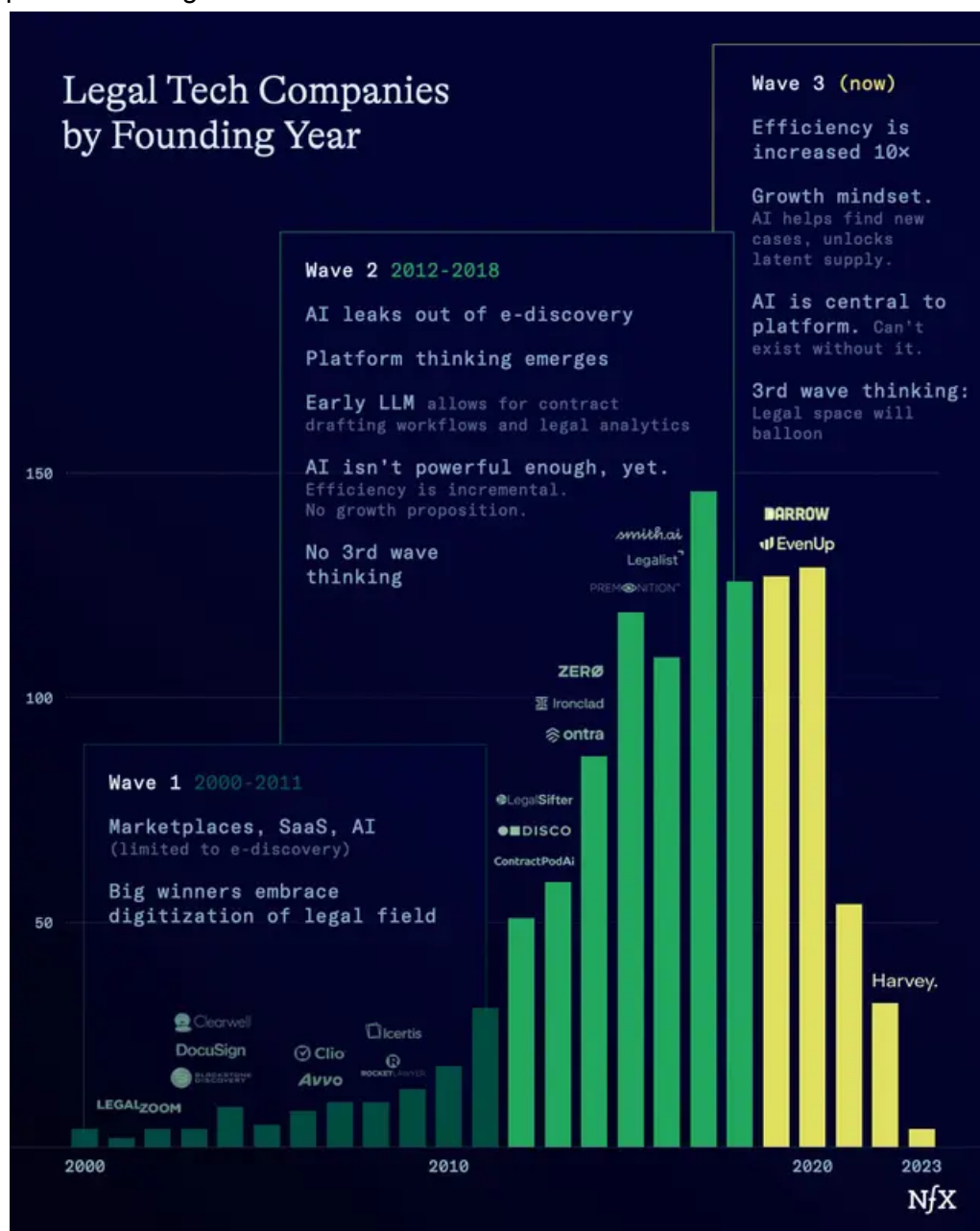
Evyatar Ben Artzi, the CEO of NFX portfolio company Darrow, points out that most litigators have a wealth of experience and finely trained instincts that allow them to excel. AI augments that natural talent.

One promise of AI in the legal field is the ability to incorporate a wider corpus of data, see patterns that human cognitive limitations and biases will overlook, and create more efficient legal processes.

AI is about to finally deliver on a promise that the legal field has been waiting for for years.

This is Actually the Third Wave of Legal Tech

There have been three waves of Legal Tech since 2000, with many of the companies trying to implement AI as part of the value proposition. The AI thinking was right, but the AI was never powerful enough. Now it is.



First Wave Examples: 2000-2011

- LegalZoom

- Docusign
- Blackstone Discovery
- Clearwell
- Avvo
- Clio
- Rocket Lawyer
- Icertis

The most successful of the first wave Legal Tech companies were traditional SaaS workflows like Docusign and Clio, or marketplaces like Avvo and LegalZoom. These companies were not particularly AI-centric.

However, a few companies in this first wave were already beginning to use AI to tackle legal problems, particularly with the e-discovery companies in the mid 2000s.

E-discovery is a process where electronic data is searched for use in legal cases. AI is used as a tool to search, rank, and categorize this data. An early example of this was Clearwell, which was used to search for concepts in documents rather than just keywords. The idea was to reduce the manpower needed to wade through troves of legal documents and save time.

A later example was Icertis, founded in 2009, to use early versions of artificial intelligence to read and analyze contracts. (Today Icertis is leaning into generative AI).

There were not many of these companies – between 2000 and 2010 there were fewer than 10 Legal Tech companies founded each year. The biggest winners like Docusign and LegalZoom weren't AI companies, but were transformative for the profession, and democratized access to legal expertise, information, and documentation. The AI companies were more incremental. AI was seen as a better way to search, or a way to cut costs by 1.5x. It was a tool, not a way to unlock latent supply, or scale a business 10x.

Second Wave Examples: 2012-2020

- ContractPod AI
- DISCO
- Legal Sifter
- Ontra
- IronClad
- Legalist
- Premonition
- Leading Edge: ZEROSystems/Hercules
- Leading Edge: Smith.ai

In the second wave, AI began to spread out of e-discovery and into a variety of other workflows. They started thinking less like tools, and more like platforms.

Disco, for example, combined the tools of e-discovery with case management software.

Other companies leaned more heavily into the contract process. They focused not just on contract drafting, but contract storage and analytics.

For example, Ontra is an AI-powered legal operating system for private markets. The company automates contract intake, negotiation and execution. It also can transform physical contracts into structured data, and compare and analyze contracts across an entire organization.

This wave was also powered by the emergence of early natural language processing which helped reduce friction within existing processes.

As The New York Times [put it](#) in 2017, LLMs “proved useful in scanning and predicting what documents will be relative to a case...yet other lawyer’s tasks like advising clients, writing legal briefs, negotiating and appearing in court seem beyond the reach of computerization, for a while.”

Many didn’t see what LLMs would eventually become – generative partners across a variety of specialized writing tasks. But the few companies that did have powered a transition toward the third wave (which we’re in now).

[Smith.ai](#), for example, saw the potential for Natural Language Processing to improve lead qualification, intake, and outreach for law firms (among other high-value service businesses). The customer engagement platform known best for its virtual receptionists has built AI tools that improve both the experience and outcomes of conversations with clients. By equipping human agents with the right information at the right time based on the nature of the conversation, the AI elements ensure that prospects and clients receive a prompt and accurate response, and that data from those interactions is passed to key business programs, all while the attorneys complete billable work uninterrupted.

Agents, supported by AI, can navigate all of the nuances of potential client inquiries, regardless of the type of law or client. The idea of LLM-powered firms has arrived in wave 3 (today), but Smith.ai was one of the earliest companies to realize that AI can improve the efficiency and quality of work and provide a superior customer experience.

NFX Portfolio company ZERO Systems/Hercules is one of them. They began by creating secure email and mobile email products for law firms. Then they expanded. Now, they have AI applications across an enterprise, not just the legal departments: policy enforcement, document extraction, time capture, filing compliance, offboarding, etc., and have transitioned into wave 3: the maturation of industry-specific LLMs.

The Third Legal Tech Wave is LLM powered by AI, and this is the big one

- EvenUp
- Darrow
- Harvey
- ZERO Systems/Hercules
- Smith.ai
- ...

The internet allowed wave 1’s big Legal Tech winners to see a world where the law was executed online. The winners in this third wave will see the future paved by LLMs and emerging generative AI technologies.

Full Commentary: [How AI will revolutionize the practice of law](#), John Villasenor (March 2023)

Artificial intelligence (AI) is poised to fundamentally reshape the practice of law. While there is a long history of technology-driven changes in how attorneys work, the recent introduction of large

language model-based systems such as GPT-3 and [GPT-4](#) marks the first time that widely available technology can perform sophisticated writing and research tasks with a proficiency that previously required highly trained people.

Law firms that effectively leverage emerging AI technologies will be able to offer services at lower cost, higher efficiency, and with higher odds of favorable outcomes in litigation. Law firms that fail to capitalize on the power of AI will be unable to remain cost-competitive, losing clients and undermining their ability to attract and retain talent.

Efficiency Improvements

Consider one of the most time-consuming tasks in litigation: extracting structure, meaning, and salient information from an enormous set of documents produced during [discovery](#). AI will vastly accelerate this process, doing work in seconds that without AI might take weeks. Or consider the drafting of motions to file with a court. AI can be used to very quickly produce initial drafts, citing the relevant case law, advancing arguments, and rebutting (as well as anticipating) arguments advanced by opposing counsel. Human input will still be needed to produce the final draft, but the process will be much faster with AI.

More broadly, AI will make it much more efficient for attorneys to draft documents requiring a high degree of customization—a process that traditionally has consumed a significant amount of attorney time. Examples include contracts, the many different types of documents that get filed with a court in litigation, responses to interrogatories, summaries for clients of recent developments in an ongoing legal matter, visual aids for use in trial, and pitches aimed at landing new clients. AI could also be used during a trial to analyze a trial transcript in real time and provide input to attorneys that can help them choose which questions to ask witnesses.

The Legal Tech Startup Ecosystem

These opportunities will spur the creation of new legal tech companies. One example is Casetext, which was featured in an early March [broadcast](#) of MSNBC's Morning Joe, and recently [announced](#) an AI legal assistant called CoCounsel. CoCounsel, which is powered by technology from OpenAI, the company that created ChatGPT, allows an attorney to ask the same sort of questions that he or she might ask of a junior associate, such as, "Can you research what courts in this jurisdiction have done in cases presenting similar fact patterns to the case we are working on?" Casetext is part of what is sure to become a rapidly growing ecosystem of legal tech companies offering AI products based on large language models.

There are also opportunities to use AI for more fully automated provision of legal services. Legal and policy frameworks will need to be updated to facilitate innovation in this space, while also identifying and protecting against the associated risks.

New Skills Required

For attorneys, getting the most out of AI tools will involve far more than just pushing a button. AI is most effective when it is used to complement human skills, and the people who learn how to leverage this collaboration well will get the most mileage out of AI tools.

This will require developing new skills, including knowing how to choose the right AI tool for a particular task, knowing how to construct the right queries and evaluate the relevance, quality, and accuracy of the responses (and then update the queries as needed), and being able to synthesize the overall results into a cohesive, actionable picture. Attorneys will also need to be

attentive to ensuring that any use of AI tools is done with appropriate attention to protecting confidentiality.

Law firms will need to institute new training so that practicing attorneys can adapt to this new environment. Law schools should update their curricula to ensure that they provide law students with instruction in how to use AI writing and research tools, as these skills will be in high demand among recruiters.

Broadening Access to Legal Services

AI also has the potential to dramatically broaden access to legal services, which are prohibitively expensive for many individuals and small businesses. As the Center for American Progress has [written](#), “[p]romoting equal, meaningful access to legal representation in the U.S. justice system is critical to ending poverty, combating discrimination, and creating opportunity.”

AI will make it much less costly to initiate and pursue litigation. For instance, it is now possible with one click to automatically generate a [1000-word lawsuit](#) against robocallers. More generally, drafting a well-written complaint will require more than a single click, but in some scenarios, not much more. These changes will make it much easier for law firms to expand services to lower-income clients.

Of course, there is a downside to lowered litigation costs as well. There will be people who exploit the nearly frictionless ability to automatically generate legal complaints to rapidly flood court systems in multiple jurisdictions with frivolous AI-written lawsuits. These plaintiffs will be expecting that in the resulting numbers game, a sufficient number of defendants will quickly settle, thereby making the enterprise profitable. It may be necessary to update laws (such as [this law](#) in California) regarding vexatious litigants to impede these sorts of abuses.

Human Intelligence Still Critical

Even with widespread adoption of AI, attorneys will still be vitally important. AI can’t make a convincing presentation to a jury. The technology can’t fully weigh the factors that go into the many strategic decisions, large and small, that get made over the course of any litigation matter. It can’t replace the human element of relationships with clients. And a computer can’t play a leadership role in motivating a team of attorneys to produce their best work.

In short, it would be a mistake to use the extraordinary advances of AI to minimize the importance of the human element in the practice of law. But it would be just as big a mistake to dismiss the role of AI, which will fundamentally reshape the landscape for both providers and users of legal services.