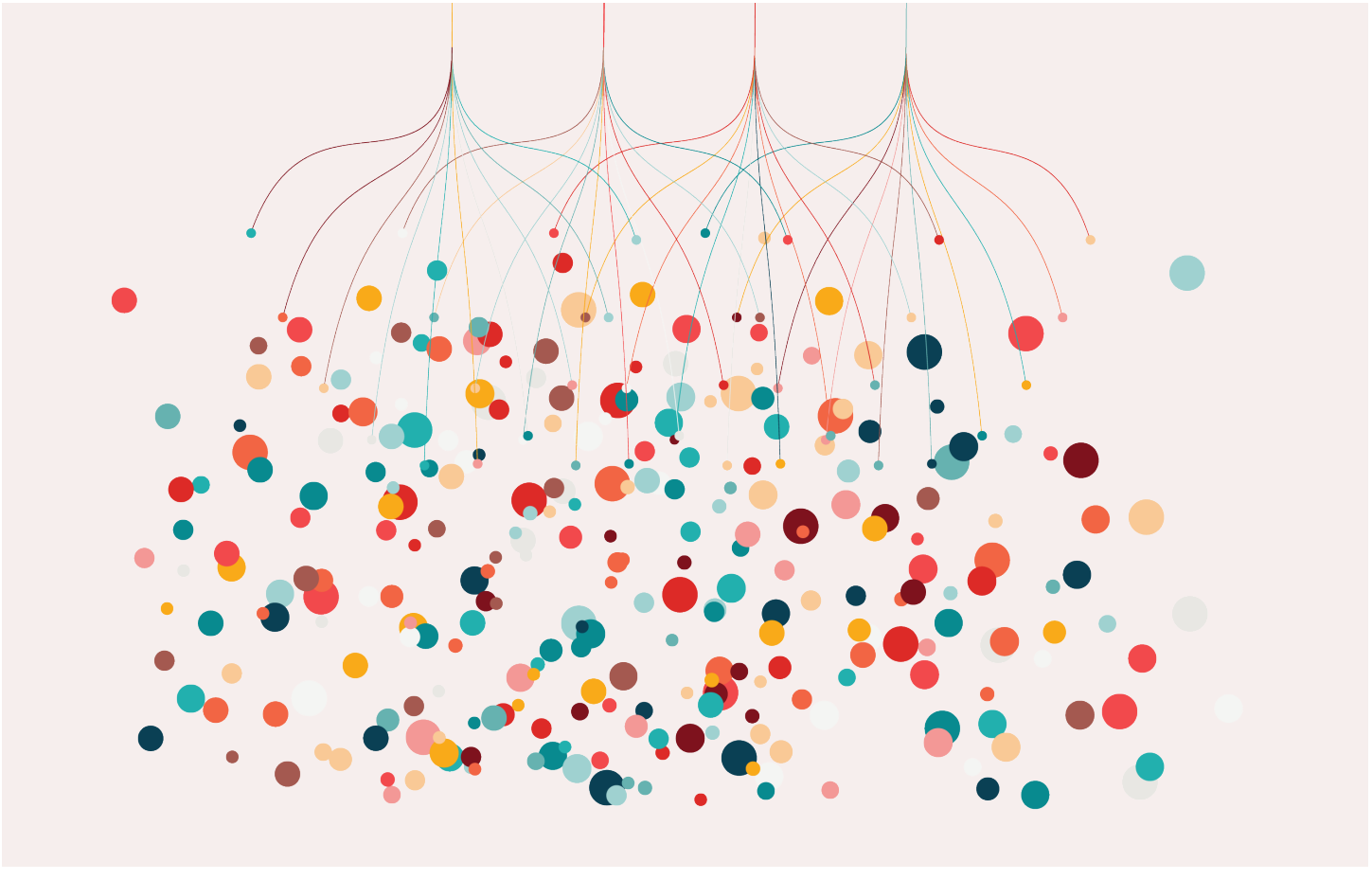
ISSUE #3 | VOL. 40

Peer to Peer

ILTA'S QUARTERLY MAGAZINE

Taxonomy, Data, Governance, and AI... Oh My!



FROM THE CEO

Joy Heath Rush
joy@iltanet.org

At the risk of being cliché, doesn't it seem as though we were just in December 2023? Calendar year 2024 whizzed by!

We wrap up 2024 and kick off 2025 with an issue that is top of mind for many legal tech leaders, whether in law departments, law firms, or tech providers. In this issue of Peer to Peer, we are focusing on Taxonomy, Data, Governance, and AI… Oh My!

Oh, My, indeed.

Let's take these words in pairs and highlight some of the key conversations taking place.

AI AND GOVERNANCE.

The increased use of AI-enabled tools, as well as the advent of - and hype around – Gen AI (happy second birthday, ChatGPT) has evolved, and in some cases transformed, application governance. Organizations without tech committees were seen to form specific committees of lawyers to discuss the ethical implications of Gen AI in the practice of law. In other instances, organizations with a tech committee added an AI committee. For nearly all organizations with whom I have spoken, the information governance function is now a well-established part of the application vetting and implementation planning teams for legal tech. We are charting this brave new world together.

DATA AND AI.

Garbage in, garbage out. Are you willing to perform a fearless inventory of your data before embarking on a major Gen AI project? This is the time.

GOVERNANCE AND DATA.

How many of your applications are associated with data? Where is that data? Is it discoverable? What about retention? Governance

and Data have to go hand-in-hand, and having data in the cloud does not absolve us of responsibility for governance of that data.

TAXONOMY AND AI.

The big question here seems to be whether AI will render taxonomies obsolete. Will well-engineered prompts replace well-architected taxonomies? I for one am in the camp of the continued value of taxonomies. In the past two years, we have learned about the value of curated content populating a large language model. If curated content is a key to success with Gen AI, then a taxonomy can be a key to success in delivering a well-curated content repository.

Enjoy the articles that follow! Hopefully they will provide you with food for thought on organizational roles and responsibilities, as well as processes, going into the new year!

Let me end by thanking you all for contributing to an amazing 2024 for ILTA!

We wish you a wonderful 2025, and we are looking forward to seeing you, speaking with you, and learning from you in the new year!

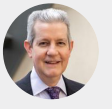
Hugs to all! **ILTA**

ILTA's Board of Directors

PRESIDENT
Tony McKenna
Howard Kennedy LLP
tony.mckenna@howardkennedy.com

EXECUTIVE VICE PRESIDENT
Julie Brown
Vorys, Sater, Seymour and Pease LLP
jkbrown@vorys.com

SECRETARY
Kara Portwood
Armstrong Teasdale LLP
kportwood@atllp.com



TREASURER
Paul Wittekind
Porzio, Bromberg & Newman, P.C.
pcwittekind@pbnlaw.com

DIRECTOR AT LARGE
Gary Berger
Ogletree, Deakins, Nash, Smoak & Stewart, P.C.
gary.berger@ogletreedeakins.com

DIRECTOR AT LARGE
Jason Dirkx
Paul Hastings LLP
jasondirkx@paulhastings.com

DIRECTOR AT LARGE
Catherine Monte
Fox Rothschild LLP
cmonte@foxrothschild.com

CEO
Joy Heath Rush
ILTA
joy@iltanet.org

A+I: YOUR INTELLIGENCE AMPLIFIED

The collective knowledge of your organization, enhanced with AI.

Connected knowledge:
Unified information powers smarter work

Practical AI:
Delivered at scale, securely, on the platform you rely on every day

Boost performance:
AI-assisted apps boost productivity and collaboration

Strategic advantage:
Turn collective knowledge into a competitive edge

Learn More

iManage
Making knowledge work

FROM THE EDITOR

Last year's Winter issue of *Peer to Peer* featured articles exploring facets of the paradigm-shifting legal tech advancements poised at the beginning of 2024 to reshape the legal sector. As expected, legal tech professionals devoted much of 2024 to exploring generative AI (Gen AI) and implementing governance, security, and operational protocol to responsibly manage the rapid explosion of data as we transition into humanity's Digital Age.

While this 2024 Winter *Peer to Peer* spotlights the ongoing importance of Gen AI advancements and the increasingly co-mingled (but equally critical) management of data within the legal sector, the information inside shows distinct progress in both areas over the past year. Standing on the cusp of 2025, it's clear that ILTAns have moved beyond broad brush-stroke generalizations about AI and Gen AI. Instead, community members focus on practical implementation and unique solutions to address specific issues.

Settle in for a cozy read as you review comprehensive guides for merging AI and data analytics to enhance efficiency and decision-making capabilities. Review case studies from law firms utilizing AI-driven automation and strategic partnerships to revolutionize their workflows and make remarkable data security improvements. From AI-driven automation to cloud-native solutions, "Taxonomy, Data, Governance, and AI… Oh My!" provides generous insight and practical advice for legal professionals modernizing their workflows and enhancing efficiency through AI and robust data governance policies in 2025.

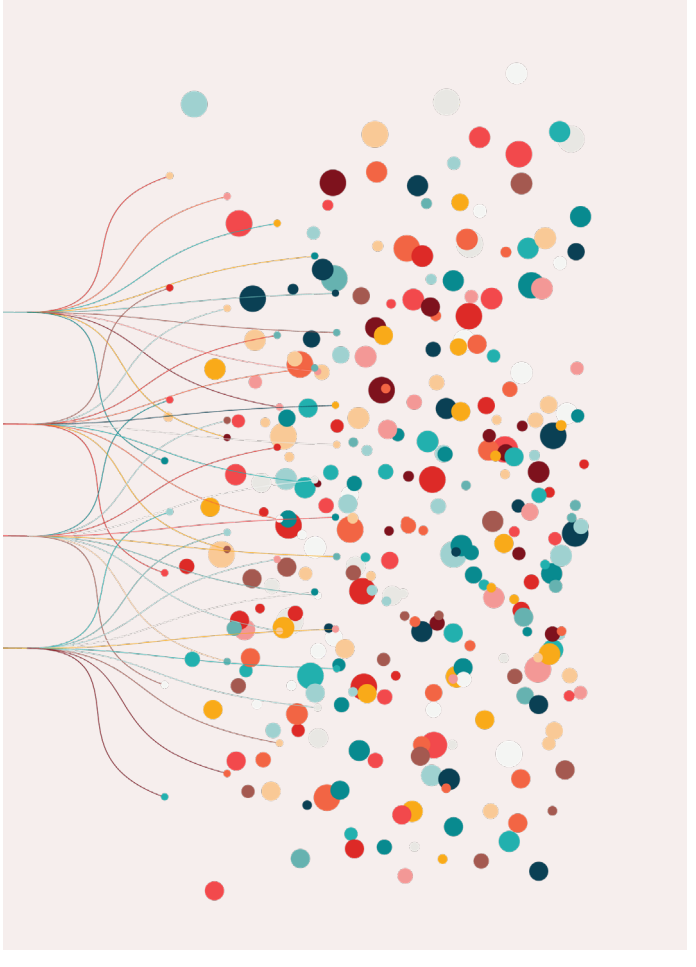
Happy Reading ILTANs! **ILTA**

Crystal Little

Editor of Content & Publications, ILTA

IN THIS ISSUE

Taxonomy, Data, Governance, and AI... Oh My!



UP FRONT

INSIDE ILTA

FEATURES

Peer to Peer
ILTA'S QUARTERLY MAGAZINE

WINTER 2024	ISSUE #3	VOL. 40

Editor
Crystal Little

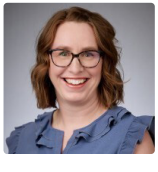
Creative Director
Kendall Lazorchak

Ad Sales
Tyler Howes

MYMENTOR

HAVE YOU CONSIDERED MENTORSHIP?

Are you starting to think about your 2025 goals or priorities? Have you recently thought about areas of professional development you want to tackle next year?



CHRISTINE BECKER
Membership Coordinator
ILTA

Most of us who have a yearly review process pause to reflect on questions like these at the end of each year. I want to suggest an opportunity. Have you ever considered partnering with a mentor but didn't know where to start? Do you want to explore being a mentor for someone? If you want to be a mentor or are looking for one, we would love for you to participate in ILTA's myMentor program in 2025!

WHAT'S IN IT FOR YOU?

Participants in myMentor expand their professional network by meeting new colleagues serving in similar roles. More importantly, there are opportunities to increase your self-confidence, improve your skills, strengthen your knowledge base, and work towards accomplishing your goals.

HERE ARE A FEW QUOTES ABOUT THE PROGRAM'S IMPACT ON SOME OF OUR 2024 MYMENTOR PARTICIPANTS*:

*Some quotes have been edited to protect anonymity.

I feel like I have a cheerleader and support system outside of my firm that I can go to for help or advice in the future.

I gained a connection with an experienced person in my field who helped to give me a new perspective. I also came up with a resource for myself, which I have shared with my team and use to help me run better meetings.

Partnerships and dedication are EVERYTHING.

I gained renewed confidence in my job and role at the firm. I have a playbook for tackling unknowns as they arise in my job, and I no longer feel like an outsider when it comes to navigating the legal or tech worlds.

I've gained valuable insights and fresh perspectives from participating in the mentor program. It's been a rewarding experience that has helped me refine my leadership skills, build stronger professional relationships, and enhance my overall effectiveness in my role.

Gained knowledge and perspectives on how to create value with everyday tasks and projects.

It was great to hear from someone who is in a similar role to people who report to me but without the baggage of actually reporting to me. I think I learned how to be a better manager by meeting with my mentee and hearing about her experience with her manager.

I can't even begin to say all the ways working with my mentor has helped me hone in on what is important to me as an employee and for my career moving forward. She was so open in sharing her own experience and answering all the questions I had.

Received the joy of enabling my mentee to shine, enhanced listening skills, and deeper self-awareness.

It has been a breath of fresh air and incredibly reassuring to engage openly with an experienced practitioner, to get practical, immediate advice when needed, and also to hear how similar, common challenges exist in all organizations, thus avoiding the risk of moving jobs as a result and jumping from the frying pan into the fire.

Also (and more importantly), I gained a new friend for life. Being a mentor also helped me to gain confidence and to acknowledge some accomplishments of my own.

If this sounds like something you want to participate in, we will open applications in January. Watch **this page** for all updates, and feel free to email **myMentor@iltanet.org** with any questions.

I truly hope you join us, and I can't wait to see you accomplish everything you set out to in 2025! **ILTA**

WWL INTERVIEWS

INFLUENTIAL WOMEN IN LEGAL TECH

Interviews by Cathrine Monte and Jima Kato

Each March, during International Women's History Month, ILTA publishes an Influential Women in Legal Tech list. Honorees are selected based on their history of mentorship and level of impact on Legal Technology.

ABOUT THE AWARDS

Since 2020, ILTA has honored outstanding women leaders in the global legal technology community through an annual application and nomination process. Honorees are selected based on their history of mentorship and level of impact on the Legal Technology industry.

ALMA ASAY
Chief Innovation and Value Officer
Crowell & Moring LLP

AALIA MANIE
Head of Webber Wentzel Fusion
Webber Wentzel

TERRI MOTTERSHEAD
Executive Director
Centre for Legal Innovation at the College of Law

GINEVRA SAYLOR
National Director, Innovation and Knowledge
Gowling WLG LLP

CARYN SANDLER
Partner + Chief Knowledge and Innovation Officer
Gilbert + Tobin

2025 APPLICATIONS NOW OPEN!

The Influential Women in Tech Awards are open to all women in the legal community across the globe. Nominees do not need to be an active ILTA member or business partner or live in the U.S.

Learn more about the 2024 Honorees

view online

For 2025, ILTA encourages our members, business partners, and those in the legal technology fields to **nominate** those who are making an active difference in the world! You may nominate more than one person or even apply yourself!

- Nominations close Friday, 24 January 2025 at 11:59 PM EST
- Applications close Friday, 31 January 2025 at 11:59 PM EST.

Winners will be announced in March 2025, highlighted at our Peer Award Ceremony on 29 April during EVOLVE 2025, and featured in our Annual Peer Award Publication. ILTA

ILTA'S DEIC

THE IMPORTANCE OF ACCURATE DATA AND ANALYTICS IN DEI INITIATIVES

BY MARI SUN & ANNA CORBETT

In recent years, corporate sectors around the globe began to shift towards creating cultures that acknowledge and address Diversity, Equity, and Inclusion (DEI) within the workplace and other professional groups.

Professional entities approach DEI-related efforts in various ways, from holding space for members of marginalized communities to share personal stories and experiences to profit-focused arguments showing the impact DEI improvements can have on a business's success. Others openly address the belief that DEI is divisive or a waste of community energy and resources. The formats are equally varied, from companies issuing statements regarding their commitment to DEI practices, panel discussions, or diversity training, to name a few.

Determining the best path forward for your organization to implement DEI strategies can be overwhelming, given the potential scope of approaches to DEI and the possible positive or negative outcomes of undertaking DEI initiatives. These efforts depend heavily on collecting and analyzing data to identify disparities, track progress, and inform decision-making. Leveraging artificial intelligence (AI) can enhance the depth of insights in the DEI context thanks to its capacity to process and analyze vast amounts of information. When discussing AI, we often focus on the ability of technology rather than the impact of data on specific processes. When it comes to DEI initiatives, having the right pool of data and statistics is just as important as the AI technology we use to analyze it.

Data provides the opportunity to progress and can help answer two of the most fundamental questions organizations should assess as a part of their DEI

endeavors: (1) does the organization need to address aspects of DEI, and (2) how do they want to approach advancing DEI improvements on an organizational level? Thoughtful and intentional data collection and analysis are critical for decision-making in every successful law firm or legal department. DEI efforts benefit from the same data-driven decision-making process and approach.

THE ROLE OF DATA IN DEI: A CLOSER LOOK

DEI initiatives aim to create more inclusive environments, promote fairness, and ensure that individuals from different backgrounds are represented and supported. However, the success of these initiatives hinges on an accurate understanding of the current landscape, assessed through data collection.

Those who believe DEI efforts in their organization are counterproductive or unnecessary may be correct. It is also possible that their DEI avoidance is rooted in a more profound psychological cause known as the "fear of finding out" (FOFO). This phrase, commonly used in the medical context, references the psychological barrier to learning information due to anxiety about what someone may learn. For example, during the pandemic, FOFO contributed to some people refusing to take COVID-19 tests, thus leading to the false conclusion that those individuals never had COVID. The fear of receiving a positive result was understandable, given possible consequences like isolation, quarantining, and missing significant personal, family, or professional events and celebrations. But there are also consequences to refusing diagnoses.

Undertaking the work of collecting and analyzing this data is essential for those with a genuine interest in improving outcomes. Missed data is a missed opportunity, and fear should not be the basis for drawing false conclusions. The insights revealed may be uncomfortable, but accurate data and careful analysis allow one to be accountable to facts rather than reactive to popular trends.

IMPROVE DATA QUALITY: ENCOURAGE PARTICIPATION AND MANAGE EXPECTATIONS

If the diagnostic data reveals opportunities to improve, it then becomes critical to understand what data you can collect and measure for any proposed DEI initiatives. When considering what data to use, organizations should be mindful of their ability to collect from an adequate sample size. Suppose a company decides to address gender disparities in leadership positions. However, the data they use to inform their decisions about the issue is collected from a small, unrepresentative sample of leadership staff. In that case, conclusions from the limited dataset will be skewed accordingly.

The quality of insights and conclusions we gather from analyzing any data set directly reflects the quality and comprehensiveness of the data collected. Like other tools and technologies, AI works with the data it ingests. The AI tool's work quality reflects the data it receives. For example, large language models (LLMs) are built on massive datasets and generate responses based on patterns and knowledge they have been trained on. However, if the data fed into these models is incomplete, biased, or inaccurate, the output will reflect those deficiencies. Ensuring data quality is a weakness of AI, however advanced it may be. Since understanding diverse populations' experiences, challenges, and needs is vital to DEI, having accurate and actionable data is essential.

DEI initiatives should strive to achieve end goals. However, moving the needle on DEI efforts looks like progress rather than victory. It is vital to celebrate incremental progress instead of focusing on completing a final outcome. AI that uses high-quality data can help the legal industry gauge progress toward DEI goals. Encouraging participation requires intentional outreach and communication to improve the data pool. To achieve this, organizations must clarify that every voice matters and ensure that insights gained from data collection efforts will drive meaningful change. Effectiveness in this type of data collection involves recruiting underrepresented groups as active participants and increasing the accessibility and inclusivity of the data

READ MORE ONLINE

find this recent DEI article from the Fall 2024 issue of Peer to Peer Magazine.

collection process. Organizations must also address the concern of privacy. Transparent communication regarding data usage and the necessary steps to protect individual privacy taken by the organization will establish trust and increase overall participation.

Whether through surveys, focus groups, or other data collection methods, the more diverse and comprehensive the data pool, the more accurate the analysis will be. Diversity representation within the data pool is essential in industries like legal technology, where diversity can vary significantly across regions, practice areas, and organizational types. AI may reveal more precise insights with bigger and better data sets, thus allowing teams to celebrate meaningful progress toward the more significant goals of belonging.

THE ILTA DEI CLIMATE SURVEY: A CASE STUDY IN DATA ACCURACY

To illustrate the importance of data quality in DEI, we can refer to the ILTA DEI Climate Survey as a case study to gauge the state of diversity, equity, and inclusion within the legal technology community. This survey is an essential resource for understanding the experiences of underrepresented groups and the effectiveness of current DEI initiatives to increase everyone's belonging within ILTA. However, one challenge that arises is the representativeness of the data collected.

Like most surveys, the ILTA DEI Climate Survey relies on voluntary participation. While the data collected offers valuable insights, it naturally represents a percentage of the legal and legal technology community. Accordingly, collecting data that captures a full-spectrum snapshot of DEI-related experiences within a broader community is often challenging. Diversity, Equity, and Inclusion

CLIMATE SURVEY

view the 2024 ILTA DEI Climate Survey Results.

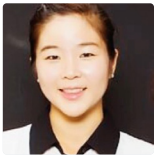
iltanet.org/dei

efforts strive to improve or increase a sense of belonging among traditionally marginalized individuals within a larger group, collective, or community. More importantly, we must remember that historically marginalized voices typically do not form the majority within a collective of survey participants. Underrepresentation of marginalized voices in survey results can produce skewed conclusions and can, ultimately, misinform DEI strategies.

Encouraging participation in DEI-related surveys and data collection efforts among community members with diverse backgrounds is crucial. By gathering a more comprehensive and diverse set of responses, we can improve the accuracy and relevance of the insights derived from the data. In the case of the ILTA DEI Climate Survey, increasing participation from a broader range of individuals within the legal technology community would allow us to better understand the current DEI landscape, identify improvement areas, and design targeted approaches for 2025.

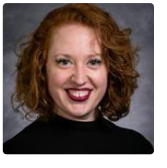
CONCLUSION

The effectiveness of any DEI effort is limited by the quality of the data driving its implementation. AI can play a decisive role in managing DEI initiatives but cannot deliver value without accurate and comprehensive data. This information is essential to understanding the experiences and needs of diverse populations, and DEI initiatives are vulnerable to failure if designed without meaningful data as the foundation. Even if AI is used, deficient data will lead to poor insights, and resulting DEI initiatives may reinforce existing disparities rather than help to address them. Furthermore, organizations miss opportunities to celebrate their progress without measuring the outcomes. Therefore, by gathering comprehensive, representative data, we can ensure that accurate insights inform our DEI efforts and that we are making positive and meaningful steps toward creating more inclusive and equitable environments. **ILTA**



MARI SUN

is a legal fellow at the World Intellectual Property Organization (WIPO). She is a member of the ILTA Diversity, Equity, and Inclusion Committee. Mari can be reached at mari.sun@wipo.int.



ANNA CORBETT

is a Practice Solutions Analyst at Akin Gump Strauss Hauer & Feld, LLP. She can be reached at acorbett@akingump.com

ADVANCING LITIGATION SUPPORT TO FUEL LAW FIRM GROWTH IN 2025

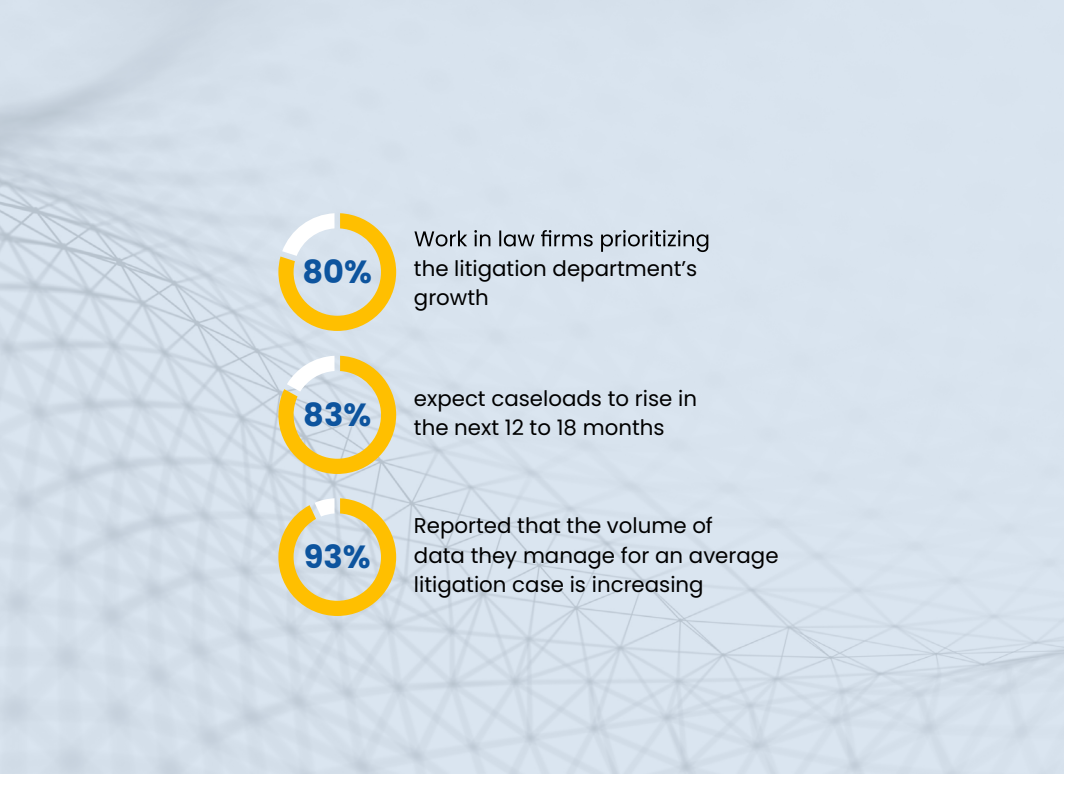
BY ARI KAPLAN

To better understand how law firms can earn a competitive advantage in a crowded market for legal services, I was honored to collaborate with **Opus 2** to identify challenges and opportunities in litigation support and case management. My conclusions and recommendations are based on interviews I was privileged to conduct with 30 litigation support directors between July and September 2024.

The research explores how litigation leaders manage increasingly complex dockets, navigate a rapidly evolving technology landscape, and set growth objectives for their litigation teams. It emphasizes the correlation between success in litigation support and case management in midsize and large law firms throughout North America (with a median number of lawyers of 488), the effect of artificial intelligence on how dynamic legal teams approach disputes, the influence of innovation and investment on these practice groups, and the expectations for change in the industry.

Download the full report here: https://www.opus2.com/litigation-support-industry-research/.


ARI KAPLAN ADVISORS
Advancing Litigation Support to Fuel Growth in a Changing Legal Market
Sponsored by
OPUS2



LAW FIRMS ARE BALANCING THEIR GROWTH OBJECTIVES WITH RISING DATA VOLUMES AND CASELOADS

The median number of cases the respondents manage in a typical year is 200, and 83% expect their caseloads to rise in the next 12 to 18 months. 93% reported that the volume of data they manage for an average litigation case is increasing, and 60% believe that the continued growth of case data volumes presents challenges for their team.

These increasing document quantity and complexity concerns are amplified because 80% of the law firms represented in this research also emphasize business growth within their litigation department. Given the profitability interests of their firms, litigation support teams must rely on technology to effectively adapt to and manage increasing data volumes. "Our technology can scale effectively, but the growth presents resource challenges," said one participant.

To address this convergence, law firm leaders need to monitor the talent and technology of the team supporting the firm's marquee matters to ensure that they have the necessary strength to compete. They should also adopt advanced solutions rather than simply adequate ones.

OUTDATED TECHNOLOGY IS A COMMON CONCERN

Beyond knowing how to use technology, savvy teams deploy it to drive litigation forward, especially since "We are in a transition phase where generative AI is being incorporated into case management and will enhance the experience and increase the potential engagement of users at a level to which they have not been accustomed," advised one participant, who echoed the perspectives of several peers.

Almost a third (30%) of the participants reported using a case management solution that is ten or more years old and described it as "inadequate." To minimize technology deficiencies and maximize user acumen, law firm leaders should align the talent of their litigators and litigation support leaders while also providing comprehensive and foundational training with emerging technology.

CASE MANAGEMENT MATTERS MORE THAN EVER

Over two-thirds (70%) use a fit-for-purpose case management system (https://www.opus2.com/solutions/cases/) designed specifically for litigation. Even so, two-thirds of those with a case management platform would like to upgrade it, as many see their competitive edge diminishing due to using suboptimal or outdated technology. This concern is growing, given the emergence of generative AI and the complexity of pre-trial obligations. "Since we collaborate so much, we need better and more modern tools," one participant explained. "If we were using a true case management tool, the efficiencies we would gain would give us a competitive advantage," offered another.

Consider deploying case management tools outside of litigation to maximize support, secure buy-in, and counter complacency and resistance. Harmonize the tools and processes to promote adoption over advancement and create a cohesive suite of tools legal teams can use according to trusted playbooks to advance dispute resolution.

"People are getting nervous and want to use it because they are concerned about missing out."

AI ADOPTION IS STILL A COMPETITIVE ADVANTAGE

Most participants noted that their firms already use generative AI to support their litigation processes through case management platforms with integrated AI or standalone AI solutions. Yet, they are in an early deployment stage. 90% are developing protocols and guardrails for generative AI use, 93% are preparing for changes driven by generative AI, and 87% indicated that AI-assisted case management software is a competitive advantage.

The most popular use cases include legal research, coding for ediscovery, interpreting filings, and drafting documents. Despite many admitting they only use the most basic functions of early-stage generative AI, "People are getting nervous and want to use it because they are concerned about missing out," acknowledged a participant.

To better understand the advantages of adopting AI, offer transparency and address concerns directly. Rather than dismissing fears, discuss

ARTIFICIAL INTELLIGENCE USE CASES:

- Legal research
- Coding for ediscovery
- Document summarization
- Drafting billing narratives

Given the current market conditions, it's unsurprising that many believe that AI will play an increasingly important role in meeting firm and client expectations.

them and include detractors in the deployment strategy to build their confidence. Recognizing that adequate case management and litigation support technology is becoming inadequate without Gen AI is also vital.

INNOVATION IN LITIGATION SUPPORT IS ESSENTIAL

Beyond Gen AI, litigation support leaders report innovation and investment in migrating to the cloud, automating reporting and data management, deploying continuous active learning in ediscovery, collaborating with an alternative legal services provider, streamlining transcriptions and translations, and creating chatbots. While most respondents reveal that their innovation investments are adequate or market-leading, some still face barriers to further advancement, including limited time, budget, and adoption. After all, sustained innovation requires leaders to empower change management. "The challenge is selling innovation to internal stakeholders and securing buy-in from them," offered a participant.

Dynamic teams view innovation through a lens of incremental improvement and add meaningful value iteratively. To that end, emphasize the rewards of innovation to promote progress and participation. Also, encourage new ideas and, when appropriate, invite clients into your innovation journey. Although clients may not be asking detailed questions about how you use legal technology to service their matters, it does not necessarily mean they are disinterested. Firms that engage their clients in this area may find new opportunities.

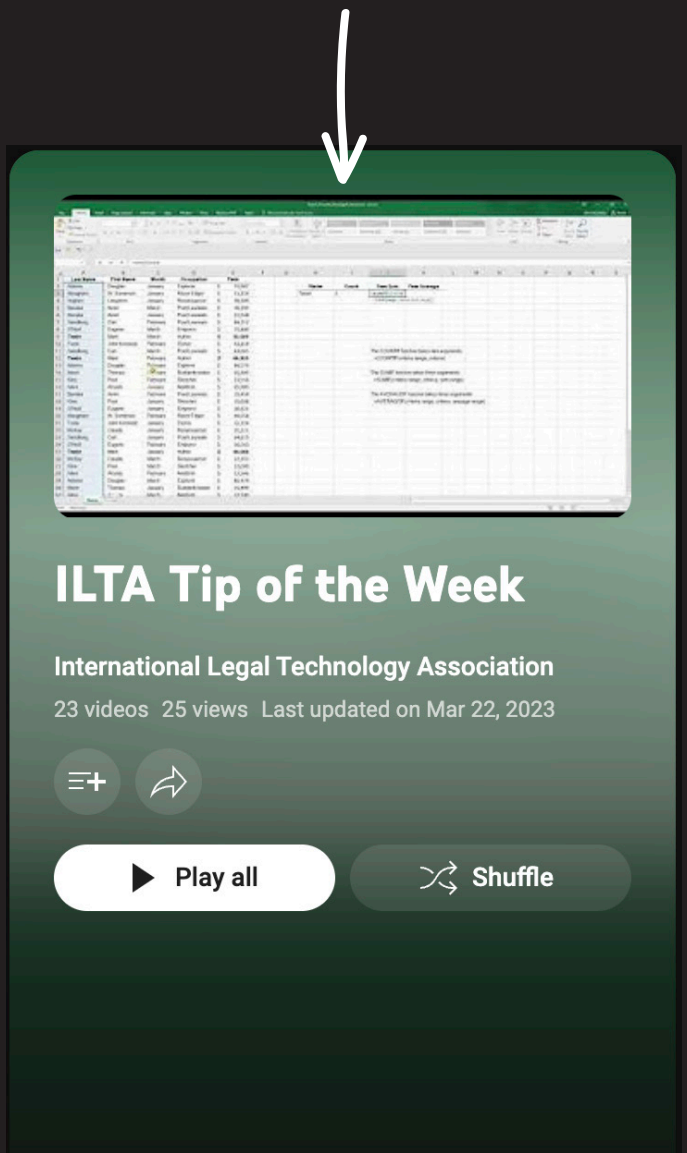
CONCLUSION

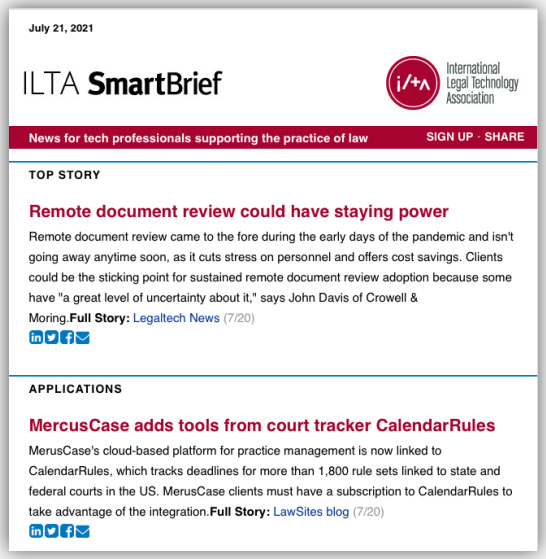
Ultimately, law firms face a convergence of complex litigation and unprecedented technological advancements driven by AI, forcing them to enhance their case management acumen, accelerate their skill-building, and artfully accommodate rapid change. Competitive pressures, growth objectives, and shifting risk factors reshape legal and make success unusually elusive. Teams that understand the critical trends in leadership, advocacy, and digital transformation will likely thrive. **ILTA**

ARI KAPLAN

an attorney (formerly with McDermott, Will & Emery in NYC) and a legal industry analyst, is an inaugural Fastcase 50 honoree, a College of Law Practice Management fellow, and a finalist for ILTA's Thought Leader of the Year award. He is the author of Reinventing Professional Services: Building Your Business in the Digital Marketplace and The Opportunity Maker: Strategies for Inspiring Your Legal Career Through Creative Networking and Business Development. Kaplan serves as the principal researcher for various widely distributed benchmarking reports, hosts the Virtual Lunch daily on Zoom, and is a keynote speaker for events worldwide. He is an avid swimmer, a self-taught, struggling guitarist, and a two-time Ironman triathlon finisher. Visit https://www.AriKaplanAdvisors.com to learn more.

Have you checked out our newest Tip of the Week?





July 21, 2021

ILTA **SmartBrief**

International Legal Technology Association

News for tech professionals supporting the practice of law SIGN UP · SHARE

TOP STORY

Remote document review could have staying power

Remote document review came to the fore during the early days of the pandemic and isn't going away anytime soon, as it cuts stress on personnel and offers cost savings. Clients could be the sticking point for sustained remote document review adoption because some have "a great level of uncertainty about it," says John Davis of Crowell & Moring.**Full Story:** Legaltech News (7/20)

APPLICATIONS

MercusCase adds tools from court tracker CalendarRules

MerusCase's cloud-based platform for practice management is now linked to CalendarRules, which tracks deadlines for more than 1,800 rule sets linked to state and federal courts in the US. MerusCase clients must have a subscription to CalendarRules to take advantage of the integration.**Full Story:** LawSites blog (7/20)

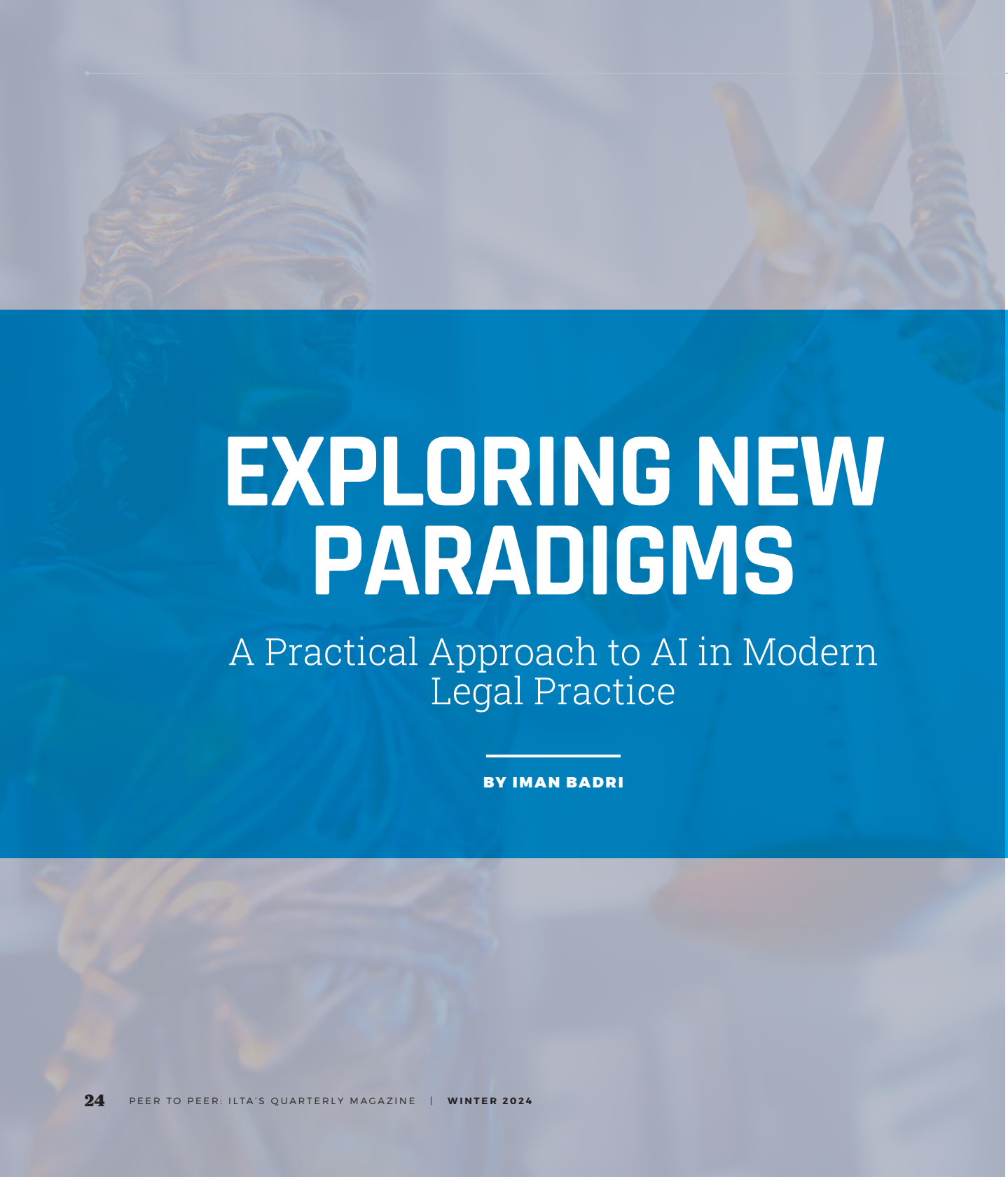
We know that the news cycle never stops and that's why we've designed ILTA SmartBrief to deliver you snapshots of the need-to-know news affecting tech professionals supporting the practice of law.

Here's what else you receive with your subscription:

- Reliable insights from the industry's leading resources, curated and summarized by our expert editorial team.
- Easy to read newsletter designed for a two-minute read.
- A simplified way to keep track of fact-based news.

Industry trends are changing every day, but as part of your ILTA membership, we'll provide you with the news to stay ahead.

SIGN UP FOR OUR ILTA SMARTBRIEF

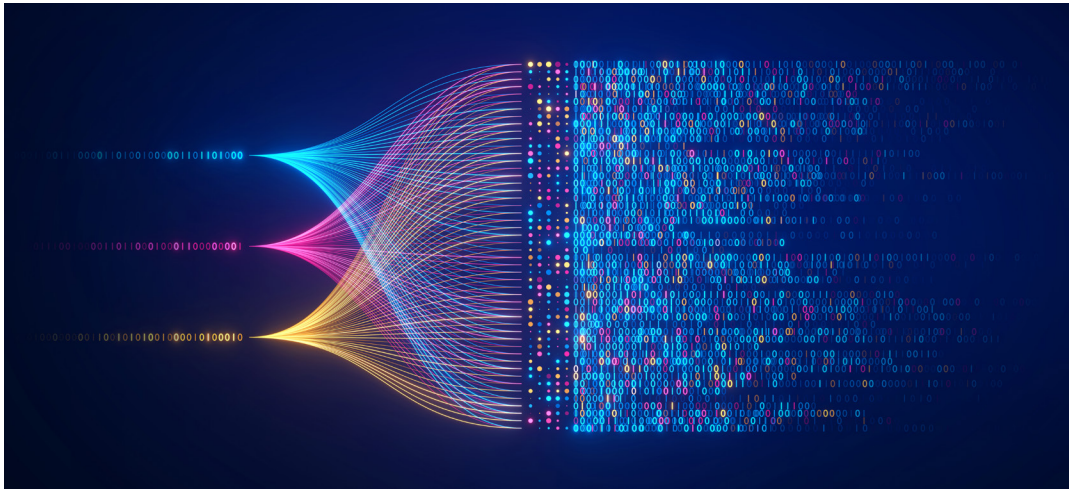


EXPLORING NEW PARADIGMS

A Practical Approach to AI in Modern Legal Practice

BY IMAN BADRI

The legal industry has recently been experiencing a technological transformation driven mainly by advances in artificial intelligence (AI). These innovations and breakthroughs are not just automating routine tasks or making the execution of complicated legal tasks easier but fundamentally changing how legal professionals do their work. From document review and analysis to automating and enhancing the search process and even predicting litigation outcomes, AI is helping lawyers and legal professionals work more efficiently and make better decisions. This article explores some of these developments, current technology usage, advanced applications of large language models, new use cases, technical challenges, ethical considerations, future developments, and some practical guidelines.



THE CURRENT TECHNOLOGY STACK IN LEGAL PRACTICE

Machine learning (ML) models in legal workflows have significantly changed document review processes. Algorithms that can process large amounts of data quickly and accurately are rapidly replacing traditional methods that rely on manual reading.

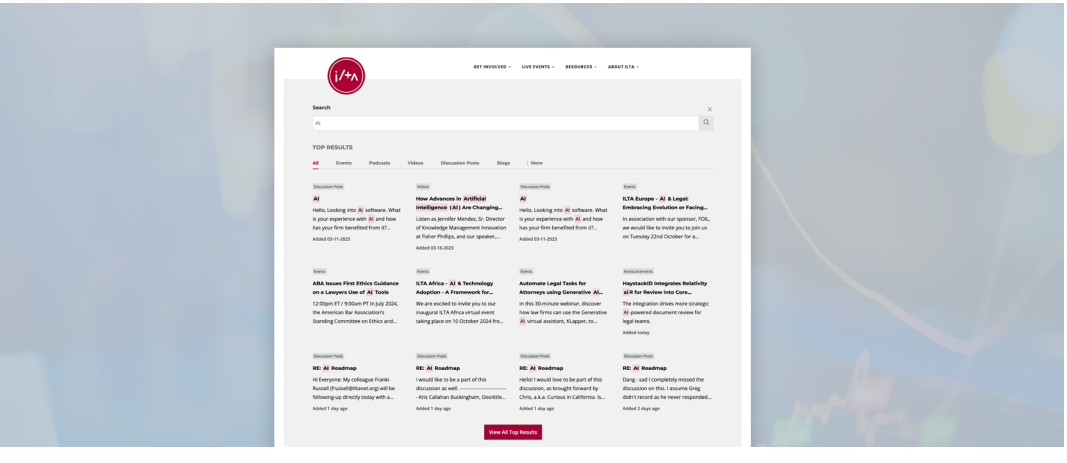
Legal text analysis widely uses ML models like Support Vector Machines (SVMs), Random Forests, and transformer-based models like BERT. Supervised learning methods, which need labeled data, are used for tasks like extracting clauses from contracts. For example, models are trained on annotated contracts to identify and categorize specific clauses, speeding up the review processes.

Unsupervised learning methods, which don't require labeled data, are used for clustering and detecting unusual contract patterns. Algorithms like K-Means clustering group similar contracts based on features from the text, helping identify outliers or unusual terms that might need closer examination.

Performance metrics for these models are necessary to assess how well they work. Transformer-based models, which can understand the context within the text, usually perform better in legal text classification processes. Model evaluation methods include accuracy, precision, recall, and F1-score metrics. Typically, transformer-based models achieve higher scores due to their ability to grasp the complex legal language and its nuances.

MORE ONLINE

Use the new search feature at iltanet.org to find lots of resources from podcasts, session materials, webinar recordings and more!



INTEGRATION ARCHITECTURES CONNECTING DIFFERENT LEGAL TECH SOLUTIONS

For AI to be effective in the context of legal, different software solutions need to work together closely and as smoothly as possible. An API-first approach is common in modern legal software development, focusing on creating robust application programming interfaces (APIs) that allow different applications to communicate and share data whenever necessary. This structure allows for flexibility, scalability, and accessibility so firms can adopt new technologies without overhauling existing systems.

Designing efficient data pipelines is also crucial, especially when processing large volumes of legal documents. Effective data pipelines handle data collection, processing, storage, and retrieval as part of a smooth process. Using technologies like Apache Kafka for real-time data streaming and Apache Spark for large-scale data processing ensures that legal professionals have timely access to relevant information, improving the accuracy and quality of the decision-making process.

ADVANCED APPLICATIONS OF LARGE LANGUAGE MODELS

Large language models (LLMs) have created new possibilities for AI applications in the legal field. These models, trained on vast amounts of data, can perform various language tasks with minimal human intervention.

PROMPT ENGINEERING FOR LEGAL APPLICATIONS

Prompt engineering and different methodologies and approaches in prompt engineering are essential for using LLMs effectively. By carefully crafting efficient prompts, lawyers and legal professionals can guide the model to produce more accurate and relevant outputs.

For example, zero-shot learning allows LLMs to perform tasks without specific training examples by relying on knowledge gained during the pre-training process. Zero-shot learning may be helpful for general legal inquiries. For more specific tasks, few-shot learning can significantly improve performance by giving the model a few examples. Providing examples of desired outputs helps the model understand the context and produce more precise and relevant results.

A common challenge with LLMs is that they can sometimes generate incorrect or made-up information, also known as hallucinations.

Techniques like incorporating retrieval-augmented generation (RAG), where the model accesses external databases to verify facts before generating the final responses, assist with reducing hallucinations. Also, adjusting how the model creates text can help reduce inaccuracies in some cases. Several other prompt engineering techniques are useful for addressing different types of legal tasks beyond this article's scope.

LEGAL-SPECIFIC FINE-TUNING STRATEGIES

Customizing large language models through fine-tuning legal data enhances their capabilities to meet the legal field's specific linguistic and practical requirements. Fine-tuning legal data to customize LLMs involves further training the model on domain-specific data like laws, case judgments, and legal documents from particular legal practice areas.

For the best outcome, the training data needs to be substantial and represent the variety of language used in legal documents. It's essential that the data is current and reflects the latest laws and regulations, especially for models focused on specific jurisdictions.

Managing model drift is an ongoing concern, especially in regulatory compliance, where laws change more frequently. Regularly checking the model's outputs and retraining it with new data helps maintain accuracy for more extended time frames. Using recent cases or laws to validate the model's predictions can detect changes in performance, prompting necessary updates.

EMERGING APPLICATIONS

The potential of artificial intelligence in the legal field extends beyond document review, encompassing predictive analytics and advanced research tools in addition to other legal practice areas.

PREDICTIVE ANALYTICS IN LITIGATION

Predictive analytics uses statistical methods and machine learning algorithms (ML) to forecast outcomes based on historical data. Predictive analytics can help lawyers develop strategies by providing insights into the chances of winning a case by adopting specific strategies.

Feature engineering is a key step involving selecting and transforming variables the model will use to make predictions. Important features might include the judge's ruling history, tendencies of the jurisdiction, legal precedents, and case details. Natural language processing (NLP) techniques extract meaningful patterns from the text in case filings and legal arguments.

Handling class imbalance in datasets of judicial decisions is essential because there might be an uneven number of cases won versus lost. Techniques like resampling the data, using algorithms that handle imbalance well (like ensemble methods), and using appropriate

evaluation metrics (like the area under the ROC curve) help ensure the model's predictions are reliable and relevant.

AUTOMATED LEGAL RESEARCH SYSTEMS

Advanced AI systems are transforming legal research by offering more precise and efficient ways to find relevant precedents and legal information.

Vector embeddings turn words, phrases, or even whole documents into numerical vectors in a high-dimensional space. By capturing the meanings of words, these embeddings allow for matching legal precedents based on concepts and context rather than just keywords. Models like Word2Vec or Doc2Vec are commonly used to create these types of embeddings.

Hybrid retrieval-augmented generation approaches combine traditional search methods with generative models. The retrieval part finds relevant documents, while the generative model summarizes and synthesizes the information, providing concise answers to complex legal questions. This combination improves the depth and quality of legal research results.

TECHNICAL CHALLENGES AND SOLUTIONS

Integrating AI in legal practice presents unique technical challenges, particularly concerning the models' ability to explain their decisions and data security.

MANAGING MODEL EXPLAINABILITY FOR LEGAL REQUIREMENTS

Explaining how AI models make decisions is crucial in legal applications because transparency is necessary, and AI-generated choices may be subject to fact-checking and legal examination.

LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (SHapley Additive exPlanations) help communicate how models reach specific predictions. These methods show which features are essential by computing the contribution of different features to the prediction outcome and how different inputs affect the model's outputs.

Creating acceptable explanations in court means AI systems must provide accurate predictions and reasoning that align with legal standards and the context of the legal case. Incorporating legal reasoning into the models and ensuring that AI-generated explanations match established legal principles increases their credibility in legal proceedings.

DATA PRIVACY AND SECURITY ARCHITECTURE

Protecting client confidentiality and following data protection laws are critical concerns.

Federated learning is a solution for practices operating in multiple jurisdictions with sensitive data. Models are trained locally on data from each specific area without sending it to a central server, sharing only model updates that are then combined. This keeps data private while still benefiting from shared learning.

Homomorphic encryption allows calculations on encrypted data, so sensitive information stays secure even during processing. Although this method can require substantial resources, technological advances are making it more feasible for legal applications where data privacy is essential.

ETHICAL AND PROFESSIONAL CONSIDERATIONS

Ethical considerations such as bias, professional responsibility, and adherence to ethics are crucial in any AI application. These considerations

ensure integrity, fairness, and accountability across various domains of AI technologies.

When applied to legal practice, these ethical concerns become even more significant. The potential consequences of biased AI decisions in the legal field can impact the justice system's credibility and the lives of individuals. Therefore, maintaining high ethical standards and putting strong frameworks in place to address these challenges is crucial.

MODEL BIAS DETECTION AND MITIGATION STRATEGIES

Bias in AI models can lead to unfair outcomes and damage the justice system's integrity.

Metrics for measuring bias include statistical parity difference, equal opportunity difference, and disparate impact ratio. These help identify whether certain groups are unfairly affected by the model's predictions.

Mitigation strategies involve methods before, during, and after model training. Pre-processing techniques adjust the data to balance representation, in-processing methods modify the learning algorithm to reduce bias, and post-processing adjusts the model's outputs. Ensuring these techniques don't reduce the model's accuracy in legal applications is vital.

PROFESSIONAL RESPONSIBILITY IMPLICATIONS

Maintaining attorney-client privilege and following professional standards requires careful use of AI systems. Some technical solutions that support confidentiality include strong encryption, secure authentication, and strict access controls. Implementing detailed audit trails and version control systems ensures that all actions within the AI system are recorded and traceable, which is vital for accountability and compliance.

To ensure that AI systems comply with legal and ethical obligations, conducting regular oversight, performing audits, and adhering to guidelines established by legal authorities and professional organizations is essential.

FUTURE DEVELOPMENTS

Various advancements in AI and model development are expected to push the boundaries of AI in legal practice even further. These developments will likely include more sophisticated algorithms capable of understanding and processing legal language with greater accuracy and nuance. Additionally, integrating AI with other emerging technologies, such as blockchain and advanced data analytics, will open up new possibilities for enhancing the efficiency and effectiveness of legal services. As these technologies evolve, they will provide legal professionals with powerful tools to navigate complex legal landscapes, improve decision-making, and ultimately better serve their clients.

EMERGING ARCHITECTURES FOR LEGAL LANGUAGE MODELS

Developments in NLP architectures designed for legal texts are expected to improve AI systems' capabilities and potentially open up new horizons.

Domain-specific pre-training involves training language models on extensive collections of legal texts, capturing the specific language and reasoning used in a legal context. This specialization leads to better models for tasks like summarizing legal documents, answering legal questions, extracting arguments, or proposing effective strategies.

Multimodal models that combine text with other forms of data, such as images, graphs, or even multimedia content, are emerging. These models can process complex legal documents with more complicated visual elements, providing a more comprehensive understanding of reality and facts.

IMPLEMENTATION GUIDELINES

Integrating AI into legal practice requires practical considerations. This section will explore a few guidelines for legal firms adopting these innovative solutions to ensure a smooth and effective implementation process.

CONCLUSION

Integrating AI into modern legal practice offers excellent opportunities to improve efficiency, accuracy, and innovation. By embracing these advancements, legal professionals can provide better services, make more informed decisions, and stay competitive. However, it's crucial to address the technical challenges, ethical considerations, and professional responsibilities that come with these technologies. A well-planned and strategic approach to implementation will help the legal industry successfully navigate these new changes, ultimately benefiting both practitioners and clients. **ILTA**

TECHNICAL REQUIREMENTS FOR DIFFERENT FIRM SIZES

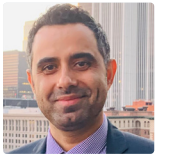
Scalability and efficient resource utilization are critical factors. Smaller firms may benefit from adopting cloud-based solutions and software-as-a-service (SaaS) models, which minimize the need for substantial upfront investments in technology. Conversely, larger firms might opt for customized systems utilizing microservices architectures and containerization technologies such as Docker and Kubernetes to manage complex workloads.

Resource optimization strategies include leveraging serverless computing for cost-effective scaling and utilizing performance monitoring tools to ensure efficient resource usage.

INTEGRATION WITH EXISTING LEGAL TECH STACK

Ensuring compatibility with current systems minimizes disruption. Standardizing APIs helps with integration, allowing different software components to communicate effectively. Following industry standards and protocols enhances compatibility and protects the technology investment for the future.

Data migration and transformation need careful planning. Implementing solid extract, transform, load (ETL) processes ensures data preservation as it transitions to new systems. Thorough testing and validation are necessary to maintain data integrity and keep operations running smoothly.



IMAN BADRI

holds a bachelor's degree in software engineering and two specialized master's degrees in information technology, focusing on AI system development, and innovation. His 20-year career spans various global industries like technology, finance, oil and gas, hospitality, and legal. Dedicated to bridging the gap between theoretical concepts and their practical applications, Iman consistently solves real-world problems by transforming complex concepts into tangible solutions. Iman shares his expertise in data science and AI through articles, book chapters, and presentations, aiming to contribute to the broader conversation within the professional community.

LEGAL LITIGATION IN THE AGE OF DATA PROLIFERATION

Ediscovery Strategies for Success

BY MARK WOODS

Not long ago, discovery in litigation was not as diverse as we see today. It included emails, physical documents, essential computer files, and phone records. Today, discovery within litigation encompasses these sources and data from communication apps like Slack, Teams, Zoom, social media, texts, cloud storage, accounting and financial systems, document management systems, and more. As the body of potentially discoverable data balloons in volume and diversity, firms must sift through an ever-growing collection of extraneous and duplicative data to present the most straightforward and compelling case. Yet, the facts and legal elements of these cases—and how those facts are presented—have not radically changed.

Recently, legal industry analyst Ari Kaplan, principal of Ari Kaplan Advisors, researched to understand how firms are handling the influx of discovery data. Kaplan interviewed litigation support directors about their caseloads, data volumes, and strategies. While 93%

of the respondents reported that the volume of data they manage for average litigation is increasing, half reported that the number of documents, records, or exhibits they use at trial has remained about the same. In other words, the needle is still the same size, but the haystack keeps expanding.

With the increasing velocity of tech adoption, wealth of communication platforms, ease of sharing information digitally, and increasing variety of media types, the exponential growth of discovery data sets is not just a limited trend. It is the state of play in litigation for the foreseeable future. This article will explore the exponential growth of data, what challenges this creates for litigation teams, and how teams can address these changes.

GROWING VOLUME AND VARIETY OF ELECTRONIC DISCOVERY DATA

The use of traditional electronic communications like email and text messages and the ever-growing host of ephemeral messaging apps, social media platforms, and collaboration software is nearly unavoidable at work and outside the office. The impact of a significant increase in electronic communication is present within legal litigation processes as well. For example, earlier in 2024, the United States Department of Justice and Federal Trade Commission issued a joint statement that they were updating language in their standard preservation letters and other notifications in government investigations and litigation "to address the increased use of collaboration tools and ephemeral messaging platforms in the modern workplace." The press release on the change warned of spoliation sanctions for abuse of ephemeral communication technologies.

According to a report from the World Bank, the amount of data stored, transferred, and used globally is growing from one zettabyte (one zettabyte = one billion terabytes) in 2010 to an anticipated 180+ zettabytes in 2025. A report from the National Telecommunications and Information Administration stated that 83% of people in the U.S. aged three and up used the internet in some way in 2023. That is 83% of the population generating some electronically stored information (ESI) that could be discoverable data during litigation.

This massive influx of information is impacting litigation. Increasing data types and volumes are two of the top three issues affecting the eDiscovery industry, according to a 2023 survey of ediscovery professionals. The importance—and burden—of ediscovery will likely continue to increase alongside the increasing adoption of various technologies. Adding to the overwhelming volume of data law firms must handle, the Ari Kaplan Advisors research reports

that 83% of respondents expect their caseloads to increase in the next 12 to 18 months.

STRATEGIES TO EASE THE BURDEN

Ediscovery can provide a powerful tool to tell the story of your case to the judge and jury when properly handled. But it will now take more time and effort to craft that story. As you face ediscovery, consider the following:

Begin with the trial in mind.

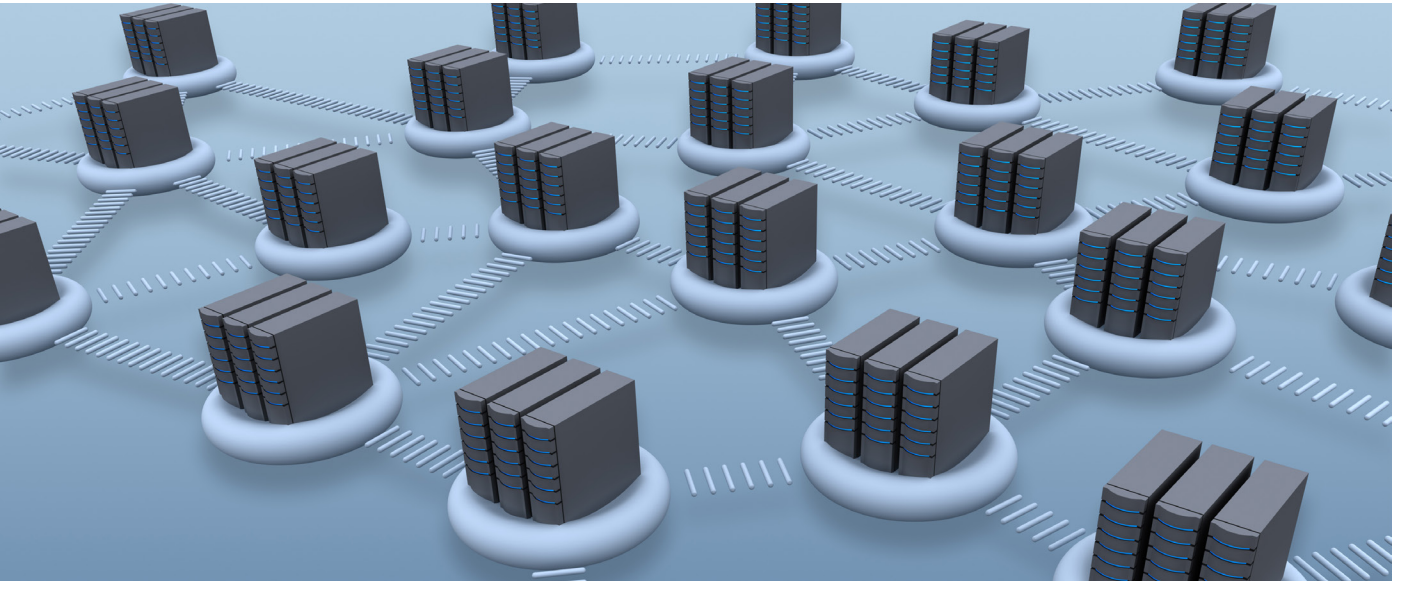
Of course, this concept is not new to discovery and litigation strategy. Whether considering a banker's box of paper documents or a trove of electronic data, it is critical to consider the ultimate objectives of a case before reviewing a single document. What are the key components to build your case? How do you match the legal elements to the evidence? What are the "bad facts," and what evidence will your opponent use to illustrate them?

Collecting and producing increasing amounts of data is not the only challenge. Litigation teams must also consider how to present these new types of media at trial effectively. At the beginning of the matter and during discovery, think about how you will effectively present your evidence at trial. Just one recent example comes from the trial of Sam Bankman-Fried for fraud related to the cryptocurrency company FTX. In that case, screenshots from encrypted messaging app Signal played a prominent role in the prosecution's case.

From the beginning, think about your case. Can those texts be presented via screenshots or print-offs? Or would it be more impactful to create a presentation to give the trier a real-time view of the messages exchanged? There will be unexpected wrinkles along the way, but starting with the end game will help you shape your case and ethically address questions about proportionality, privilege, and more.

Prepare proactively and utilize written plans.
Taking time to formulate a strategy early will give you a playbook in the future. From there, identify areas of impact and where the deluge of ediscovery will be either an effective tool, a challenge, or both.

For example, in a breach of contract case before the current era of ediscovery, the parties would likely present the contract, perhaps some related correspondence, and witnesses for both sides. Today,



probable evidence will include the contract, witnesses, multiple electronic message threads such as emails, texts, Slack messaging, video calls, and more. The messages could contain multiple attachments and may be cumbersome to produce coherently. And the language of all of these is often far less formal than the faxes and even emails of disputes in the past. Will you need a witness to testify to whether the thumbs-up emoji meant there was mutual assent?

Early conversations with clients are necessary to ensure that they understand discovery requirements and the potential consequences of failing to produce discoverable evidence. You also need to know how the people within an organization communicate, such as through chat, on their own or employee-owned devices (likely a combination of both), and where information is stored. Communication with opposing counsel and ediscovery vendors will also be essential to managing the scope, process, and timeline.

A best practice would be to have a single source of truth for storing written trial strategies and project management plans. Include descriptions of how your evidence will be presented and authenticated at trial. Having a written plan that is available to all trial team members prevents information silos. In general, over-communicating when it comes to trial preparation is preferred.

Use technology wisely.

Anything that simplifies life for the litigation team and makes them more efficient will help deal with the

increased data. Of course, ediscovery and litigation support teams are already well-acquainted with using technology to improve efficiency. Ediscovery platforms (and technology-assisted review) have streamlined processing and review for the last decade.

Modern case management solutions that integrate with popular ediscovery software further streamline processes, feeding responsive documents directly into a cloud-based collaborative workspace that allows the litigation team to build their case during data production. These case management solutions enable the creation of chronologies, events, and facts and manage transcripts and designations on the same platform, eliminating the need to hop from tool to tool. And now, the next wave of efficiency-boosting technology is generative AI (Gen AI).

There are two main approaches to implementing Gen AI. The first approach involves adopting a broad, standalone Gen AI tool that may require customization or prompt engineering to work in legal contexts. Once adapted to your needs, users must add steps to their workflow to import data into the Gen AI tool, conduct their analysis, and then move outputs to usable platforms.

The second approach utilizes Gen AI features embedded within case management and preparation solutions. This enables your team to use Gen AI designed explicitly for litigation to help manage case data while keeping work centralized. For many law firms still in the early stages of Gen AI implementation, this shows value quickly and encourages strong user adoption.

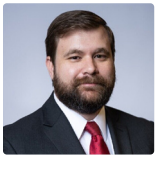
For example, if your firm already has a case management tool, find out if it offers Gen AI to analyze case documents. Instead of manually building a case chronology, using timekeepers to seek relevant events from thousands of case documents, integrated Gen AI can quickly identify key events and the connections between the characters and issues in your case. Then,

you can use it to create and manage your case chronology and character profiles without adding any new steps to your workflow.

On the other hand, if your current toolset does not have integrated Gen AI, you may still use standalone Gen AI to assist with building your chronology and character profiles. However, you must understand the tool's security and data privacy policies and weigh whether the tool saves time and adds value despite the steps needed to add to your workflow. Ensuring that Gen AI outputs are verifiable and accurate is also essential. It may be more challenging to use a broad tool versus a case management solution designed to offer citations linked to underlying documents.

CONCLUSION

While the haystack is not getting any smaller, the right tech tools can help law firms find the needle. Careful consideration of your firm's and client's needs can help overcome the challenges of expanding discovery data and enable you to advocate more effectively. **ILTA**



MARK WOODS

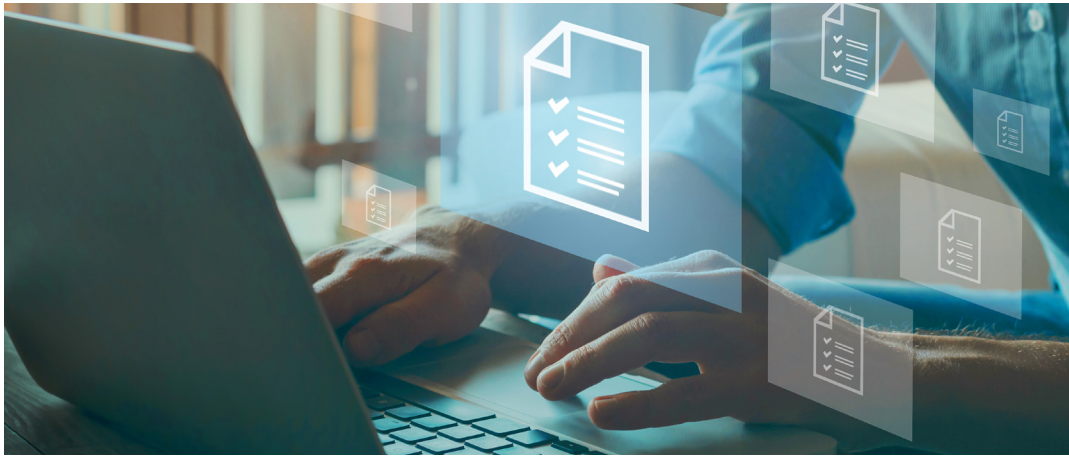
is a sales consultant at Opus 2, a leading legal software and services provider. Before joining Opus 2, Mark practiced as a litigation attorney for a decade, focusing on insurance, premises liability, and other civil disputes.

FROM POLICY TO PRACTICE

Effective AI Governance in Legal Practices

BY KANDACE DONOVAN

Artificial Intelligence (AI) has burst onto the legal scene, generating much excitement about its potential within the industry. It is not a passing trend; it is here to stay, shaking things up from contract analysis to predictive litigation. AI has become as essential as comprehensive legal research databases. The key challenge is harnessing AI's benefits effectively while mitigating associated risks.



Enter AI governance – an essential element of this rapidly changing technology. AI governance involves going beyond managing new technology to handle unfamiliar challenges confidently. So, buckle up, legal eagles! Prepare to transform your practice with effective AI governance. Here are a few ways to become a governance expert.

AI in the legal world is like a double-edged sword. AI can increase efficiency, faster research, and unparalleled data analysis capabilities. However, it also presents ethical concerns, potential biases, and regulatory challenges. AI governance is crucial for protecting your firm from these pitfalls.

Without proper governance, you are operating without direction—much like appearing unprepared in court for a high-stakes litigation case, risking severe consequences. Effective AI governance ensures that your firm leverages AI's benefits while staying on the right side of ethics and regulations. AI governance supports compliance and builds trust with clients, courts, and the public.

DEVELOPING EFFECTIVE POLICIES AND PROCEDURES

So, you recognize the importance of AI governance. Now what? Understanding the importance of AI governance is just the beginning. Now, it is time to develop effective policies and procedures. Your firm's unique needs and values should guide your policy development, as there is no one-size-fits-all solution.

Start by identifying the key areas where AI intersects with your practice. Are you using AI for document review? Predictive analytics? Client intake? Each application needs its own set of guidelines. Your policies should cover everything from data privacy and security to decision-making processes and accountability measures.

Remember, these policies should be living documents, flexible enough to evolve with the AI landscape. Ensure the policies are understandable and accessible, avoiding overly complex language.

ELEMENTS OF AN EFFECTIVE AI POLICY

An effective AI governance policy for law firms addresses key operational, ethical, and compliance issues associated with AI usage. Typically, such policies define what data is used for AI training, which often means setting clear boundaries on client information, ensuring that any data used is either anonymized or explicitly approved for AI purposes through special agreements with clients. Policies also cover procedures for addressing bias or unexpected AI hallucinations—instances where the system generates inaccurate or misleading content—outlining a formal process for assessment, mitigation, and continuous monitoring. Additionally, the policy should establish a method for purging any erroneous or inappropriate content from the AI's knowledge base, ensuring the system constantly evolves to meet firm standards.

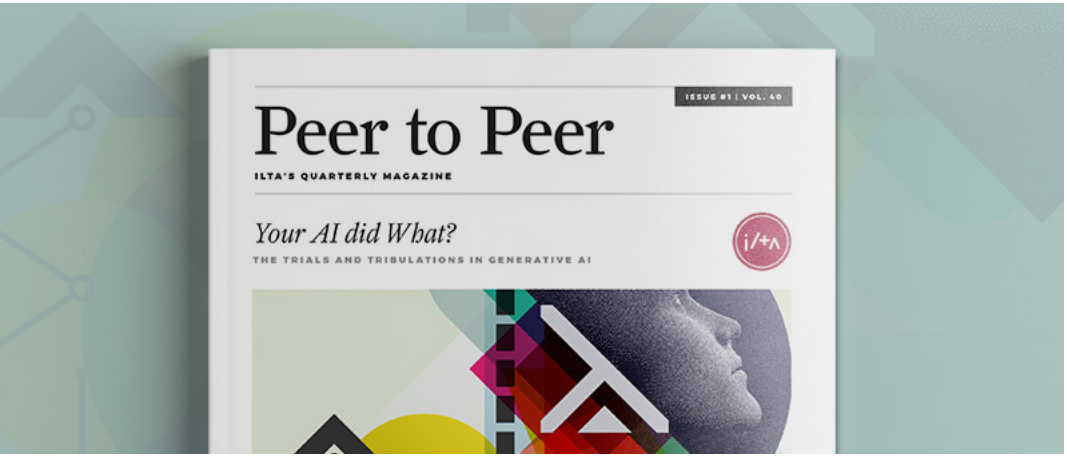
A strong governance framework also includes rules for version control and mechanisms for auditing AI decisions to maintain transparency and accountability across legal practices. Access control and role management are critical. Defining who has access to AI tools and training data helps prevent unauthorized use and ensures proper oversight. Human-in-the-loop oversight should ensure that human experts regularly review AI outputs, particularly for critical legal tasks, helping mitigate the risks of autonomous AI decision-making.

An effective policy must also include an incident response plan detailing procedures for managing AI-related incidents like data breaches or algorithmic errors. This plan should outline how incidents are reported, resolved, and documented to ensure accountability and continuous improvement. Compliance mechanisms are needed to keep up with evolving AI regulations, with a designated team or governance officer responsible for tracking regulatory changes and integrating necessary updates.

Your team should conduct regular ethical impact assessments to evaluate the broader implications of AI systems, focusing on client rights, potential social impacts, and the alignment with the ethical obligations of the legal profession. Data minimization and retention protocols are also essential, as they set rules for collecting only the necessary data and establishing retention timelines to minimize privacy risks and comply with data protection laws.

MORE AI RESOURCES ONLINE

Browse the Spring Issue for more AI resources.



Law firms can proactively address these elements to maintain ethical AI practices supporting regulatory compliance and client trust.

CHOOSING AND ADAPTING AI GOVERNANCE FRAMEWORKS

You might think, "Do I need to reinvent the wheel here?" Good news – you do not! Many existing AI governance frameworks can be adapted to fit your firm's needs. Some popular frameworks include the EU's Ethics Guidelines for Trustworthy AI, which emphasize fairness, transparency, and accountability; the OECD AI Principles, which focus on ensuring AI is innovative and trustworthy while respecting human rights and democratic values; and the IEEE Ethically Aligned Design, which guides aligning AI technologies with ethical standards. These frameworks offer foundational guidance that can be tailored to meet the specific needs of law firms. These frameworks provide a solid foundation, but remember – they are just starting points. You will need to customize them to address the specific challenges and risks in the legal sector.

When choosing a framework, consider factors like your firm's size, practice areas, and the types of AI tools you use. You can mix and match tools and services– taking the best bits from different frameworks can help create a governance structure as unique as your firm's culture.

CONTINUOUS MONITORING AND ASSESSMENT

Implementing AI governance is not a "set it and forget it" model. It is like tending a garden – it needs constant care and attention to flourish. Continuous monitoring and assessment are crucial to ensure your AI systems perform as intended and avoid undesirable behaviors such as bias in predictive outcomes or the mishandling of sensitive client information. Set up regular audits of your AI tools. Are they producing consistent results? Are there any signs of bias creeping in? Are they complying with the latest regulations? These check-ups should be as routine as your annual physical – and potentially just as revealing.

Keep an eye on how your team is interacting with the AI tools. Are they using them correctly? Are there any areas where users need additional training? Remember, the human element is as crucial as the algorithmic one for effective AI governance.

TRAINING AND AWARENESS PROGRAMS

Develop comprehensive training programs that address the technical use of AI tools, ethical considerations, and governance policies. Make it engaging – no one wants to sit through another PowerPoint presentation. Use real-world scenarios, interactive workshops, and gamification to emphasize key points. Training

> Some lawyers might see AI governance as another bureaucratic hurdle, slowing their work. Others might fear that it's the first step towards AI replacing them altogether.

should be continuous to keep pace with industry and technology changes. Incorporate regular refresher courses and updates on new AI developments into your firm's routine training agenda.

CREATING AI ETHICS COMMITTEES

Establishing an AI ethics committee is becoming crucial to AI governance in legal practices. This committee should be a diverse group of individuals from different backgrounds – lawyers, tech experts, ethicists, and even representatives from your client base. Their job? To grapple with the thorny ethical issues that crop up when AI and law intersect. Questions like "Should AI make decisions on case strategy?" and "How do we ensure AI-generated research is unbiased?" are key concerns for the ethics committee. The committee should regularly review AI implementations, discuss ethical considerations, and recommend policy updates. They act as the conscience of your firm's AI operations.

COMMON IMPLEMENTATION CHALLENGES

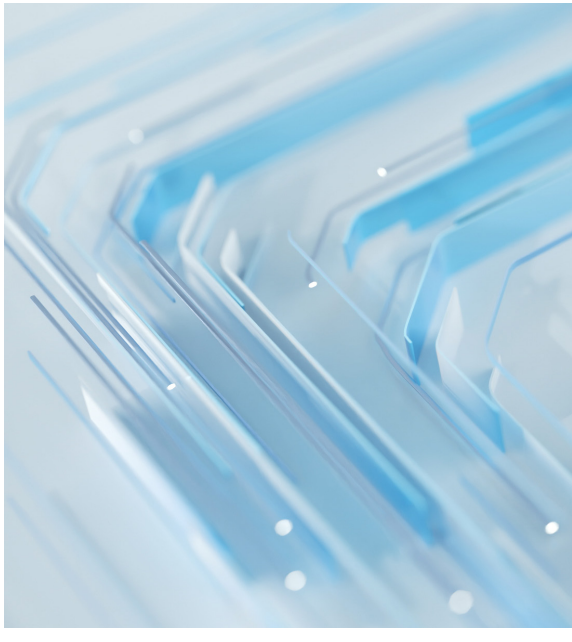
Resistance to Change

Implementing AI governance can be challenging due to resistance, skepticism, and technophobia among lawyers. Some lawyers might see AI governance as another bureaucratic hurdle, slowing their work. Others might fear that it's the first step towards AI replacing them altogether.

To overcome this challenge, emphasize transparent communication about the necessity and benefits of AI governance for both the firm and individual lawyers. Share success stories from other firms or industries. And most importantly, involve your team in the process. When people feel they have a say in shaping the changes, they are much more likely to get on board.

Technical Complexities

Most lawyers did not go to law school to study algorithms and machine learning. The technical aspects of AI can be highly confusing. This complexity can make it challenging to develop and implement effective governance frameworks. How can you govern something you do not fully understand?



The solution? Bridge the knowledge gap. Bring in experts who can explain AI concepts in terms lawyers can understand. Develop AI literacy programs within your firm. And when it comes to governance policies, focus on the outcomes and ethical considerations rather than getting bogged down in technical jargon.

Balancing Innovation and Control

This is a tricky tightrope to walk – you want to encourage innovation and leverage the full potential of AI, but you also need to maintain control and mitigate risks. Lean too far one way, and you might stifle progress. Lean too far the other, and you could be opening Pandora's box of ethical and legal issues.

The key is to create a governance framework that is flexible and adaptable. Set clear boundaries and guidelines while leaving room for experimentation. Encourage a culture of responsible innovation where new ideas are welcome but always subject to ethical scrutiny. Governance is not about saying no to everything but finding ways to say yes safely and responsibly.

BUILDING A STRONG BUSINESS CASE FOR AI GOVERNANCE

Cost-Benefit Analysis

Implementing robust AI governance requires time, resources, and financial investment. Partners might question if it is worth the cost. Conduct a thorough cost-benefit analysis. On the cost side, factor in training

programs, potential new hires (like AI ethics officers), and the time spent developing and implementing policies. On the benefit side, consider the reduced risk of ethical breaches or regulatory fines, improved client trust and satisfaction, and potential competitive advantage in the market. There are also less tangible benefits, like improved decision-making processes and a more ethically aware firm culture. These might be harder to quantify, but they are no less valuable.

Risk Mitigation

Consider this: a single AI-related ethical breach could cost your firm dearly, not only in potential lawsuits or regulatory fines but also in reputation damage that could take years to recover from. AI governance acts as an insurance policy against these risks. Like a seatbelt, it's essential for protection, even if you hope never to need it. Highlight specific risks that AI governance can mitigate, such as data privacy breaches, biased decision-making, and non-compliance with regulations. Quantify these risks where possible. What would be the potential cost of a major data breach? What about a class-action lawsuit alleging AI bias in your firm's practices? Spend time estimating the likely costs of AI-related dilemmas. Framing AI governance as a critical risk management tool shows that it is an essential, rather than optional, part of AI adoption.

Competitive Advantage

In an environment where many law firms are jumping on the AI bandwagon, robust AI governance could be your competitive advantage, like having a high-performance vehicle in a crowded market. Clients are becoming increasingly savvy about AI and its implications. Many are starting to ask tough questions about their data usage and what ethical safeguards are in place. A firm with strong AI governance can confidently answer these questions, building trust and potentially winning more business.

Moreover, as regulations around AI in legal practice inevitably tighten, firms with established governance frameworks will be ahead of the curve. Instead of scrambling to comply – your firm will already be there, maybe even helping to shape the regulations themselves.

EMBRACING THE FUTURE OF LEGAL PRACTICE

As we conclude our exploration of AI governance in legal practices, it is helpful to take a moment to zoom out and see the big picture. AI redefines what it means to be a lawyer in the 21st century. Effective AI governance does more than tick boxes or ensure your firm avoids potential risks. Its adoption is about embracing this new era of legal practice with confidence and integrity. Implementing effective AI governance is about ensuring that we harness AI's power in a way that upholds the core values of justice, fairness, and ethical conduct that are the bedrock of the legal profession.

The road ahead may be challenging, but it offers the opportunity to enhance efficiency, uphold ethical standards, and lead the transformation of legal services. By establishing robust AI governance frameworks, we are protecting our firms and clients and shaping the future of law itself. We are ensuring that as AI becomes an increasingly integral part of legal practice, it does so under the guidance of human wisdom and ethical considerations.

As you embark on your AI governance journey, remember that you are pioneering a new frontier in legal practice. You set the standard for how AI and law can work together to create a more efficient, fair, and just legal system. The future of law is here, and with effective AI governance, we can ensure it is a future we are proud to be part of. **ILTA**

KANDACE DONOVAN

is VP of North American Operations for LegalRM, the makers of iCompli. LegalRM have hosted many educational webinars over the last few months. These webinars are all available free of charge and on demand by clicking here. If you would like to find out how iCompli, from LegalRM can help your firm manage and govern data more efficiently and compliantly then do not hesitate to get in touch.

THE FORGOTTEN AI USE CASE

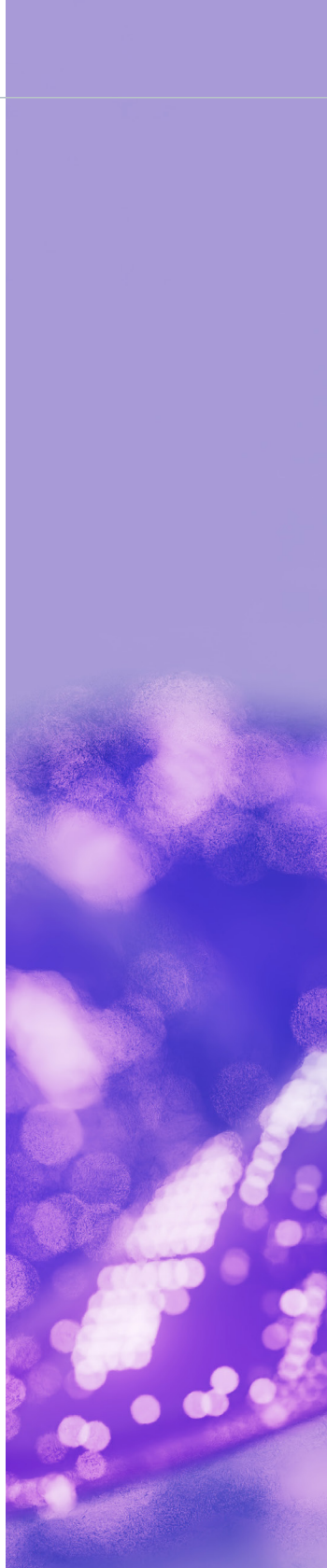
Automating Billable Time-Tracking

BY NANCY JENG

The rise of AI in legal tech has spurred a wave of excitement, primarily focused on automating law practice. Many law firms are exploring AI to streamline core legal work like legal research, summarizing complex information, and drafting initial versions of documents. In fact, according to ILTA's Legal Tech Trends 2024 survey, these three activities top the list of projected AI applications in law firms. The catch? These are all billable activities.

While these AI advancements are impressive, they come with an inherent challenge: AI's impact on billable tasks directly and negatively affects revenue. The more core work that AI handles, the fewer billable hours accrued. Decreasing billable hours can create a conflict of interest for firms operating under the billable hour model, leading to a cautious approach in adopting AI for client-facing, revenue-driving tasks.

Instead, the most transformative AI application may not replace legal work but optimize the business of law—starting with the persistent headache of time-tracking. Time tracking is essential to legal billing but is notorious for being repetitive, time-consuming, and ultimately unbillable.



Automating these back-office tasks can free attorneys from administrative burdens while ensuring no billable time slips through the cracks. Frankly, this is the work that lawyers despise. After all, as fans of Suits might recall, we never saw Harvey and Mike painstakingly logging their hours. And while it might not make for compelling television, the hidden potential in AI-driven time-tracking is substantial.

THE PROBLEM WITH BILLING TODAY

Time-tracking is a cornerstone of legal billing. Yet, it remains one of the profession's most manual, tedious tasks, which makes billing an ideal candidate for AI-driven efficiency gains. Some of the key problems that exist in today's billing processes:

Humans are not timers, nor should they be.
Manual timekeeping places an unnecessary burden on professionals. As expected, people often forget to bill time - especially if work happens outside the office or in ad-hoc instances while context switching. Employing start-and-stop timers also requires continuous user action. This constant interruption undermines the deep focus necessary for strategic, high-impact work. Studies estimate that interrupted work costs businesses over $500 billion annually in lost productivity—a clear signal that time-tracking methods need a transformation.

Rinse and repeat: good for bathing, bad for your sanity.
Billing practices often rely on repetitive, manual processes, from assigning entries to specific clients and matters to following UTMBS standards. This becomes time-consuming and frustratingly inefficient for professionals managing high volumes like insurance work or transactions. Re-assigning the same type of work with the same kinds of codes over and over again may be the new definition of insanity. Simplifying these steps is critical to freeing space for more valuable, billable tasks.

Good billers are not born overnight.
Billing is a mix of art and science. The best billers know it is not just about time captured but how the time entries are bundled, categorized, and described to clients to achieve high realization rates. Good billing behaviors require a lot of word-of-mouth knowledge transfer or tacit learning - both things that take a lot of time and energy from billing partners. While nice in theory, outside counsel guidelines are often ignored in practice. Associates and lawyers need a way to bring these best practices directly into their billing process smartly and seamlessly.

AI-DRIVEN BILLING AUTOMATION

Traditional billing software has often fallen short of true automation. However, AI has now reached a sophistication level that allows it to handle many of the tasks previously managed by human billers. AI can interpret nuanced tasks, make context-based decisions, and adapt to billing preferences over time.

AI can learn to discern billable activities from non-billable ones, reducing time spent reviewing and filtering entries.

Detecting Billable Work

Not every task a lawyer performs is billable, nor is every digital action relevant to a client's bill. AI can learn to discern billable activities from non-billable ones, reducing time spent reviewing and filtering entries.

Assigning Activities to Relevant Clients and Matters

AI can automatically categorize and assign activities to the appropriate client and matter based on contextual cues, whether an email, a call, or a document edit. AI can also detect patterns in activity and group related tasks into a single entry to create more precise and compelling bills.

Generating Detailed Narratives

AI can generate narrative descriptions that accurately capture work activity and adhere to best practices for active language and specificity. AI can detect trends in punctuation style, verb tenses, and more to generate personalized narratives that save lawyers significant time.

Learning Client-Specific Requirements and Preferences

Advanced AI models can adapt to client-specific billing requirements, continually refining how they categorize, narrate, and present activities to meet those standards. By analyzing past billable entries and identifying client preferences, AI can produce bills that reflect the style and detail level expected by each client.

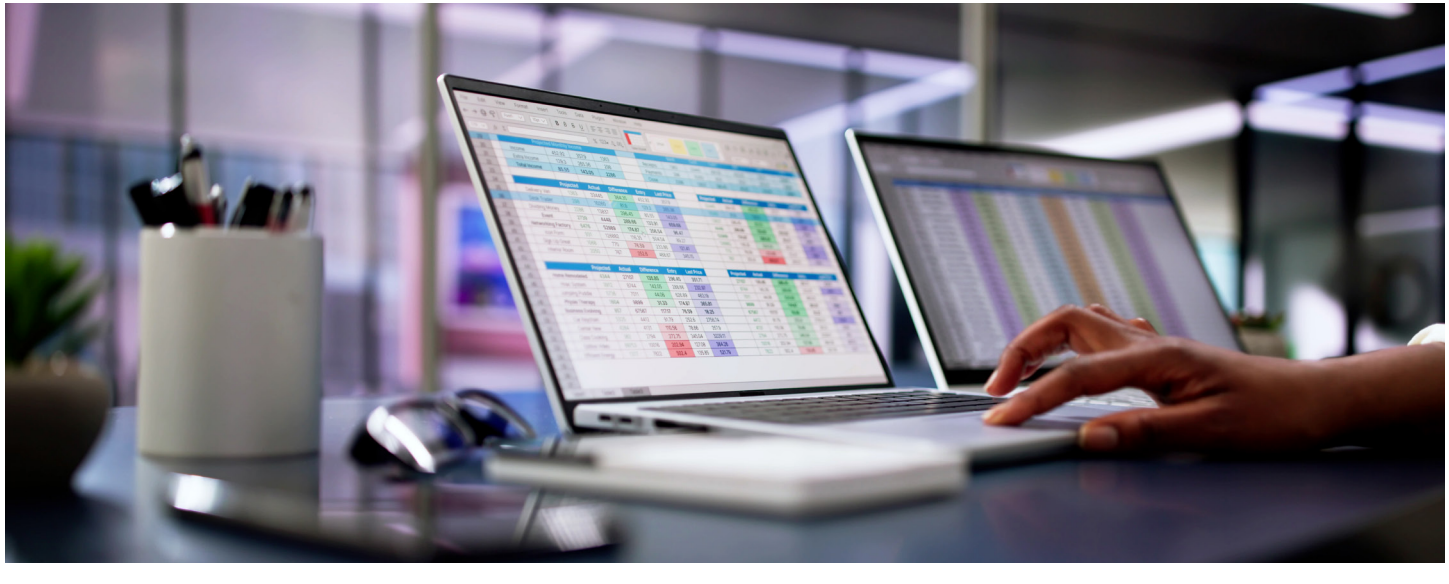
By analyzing past billable entries and identifying client preferences, AI can produce bills that reflect the style and detail level expected by each client.

Automatically Assigning Task or Activity Codes
By learning common coding practices and following ABA standards, AI can automate task or activity code assignments based on work type and client-specific needs, minimizing repetitive billing steps.

In many ways, AI is becoming the ideal billing assistant—attentive to preferences, quick to learn, and diligent about applying new insights. In the past, billing practices were passed down manually through training, settings, and manual reviews. Today, AI can process explicit and implicit signals, building a nuanced understanding of each lawyer's approach to time-tracking.

The complexity of legal billing and time-tracking has made it a complex area for AI to address until recently. Previous solutions could digitize time entries but couldn't capture the entire workflow needed to support legal billing. What makes recent AI developments stand out is their ability to integrate multiple context-aware steps—essentially handling the process end-to-end.

Only in the last few years have AI models become capable of nuanced judgment and complex decision-making, which are crucial for time-tracking automation. From discerning billable work to customizing narratives, these AI capabilities create an "intelligent" billing assistant capable of adapting to each law firm's—and each client's—unique needs.



LESSONS LEARNED IN DEVELOPING AI FOR LEGAL BILLING

Working with billers has revealed that no two are the same; each brings a unique blend of precision, creativity, and client knowledge to their work. Billing is as much an art as a science, involving complex rules, flexible interpretations, and a deep understanding of client expectations.

Accurate billing is not just about creating a precise log of time spent; it is about telling a story that demonstrates value to the client. The best billers use their narratives to affirm the work's importance, building trust and justifying costs. This means that any AI solution must be accurate and capable of persuasively and insightfully presenting billing information. Understanding this complexity has informed our approach to designing AI for time-tracking: creating a solution that values accuracy and client relationship management.

While billable time-tracking is a primary focus, the AI capabilities developed in this area have applications across the broader billing lifecycle. Potential applications include:

Pre-bill Review

AI can help billing partners conduct preliminary reviews of client bills, automatically identifying discrepancies, unnecessary details, or issues that might lead to a rejected invoice.

Outside Counsel Bill Review

For in-house counsel, AI can assist in reviewing bills received from firms, ensuring compliance with billing guidelines, and identifying areas that may require adjustment or clarification.

Fee Estimation

For firms that use flat fees or retainer models, AI can analyze past billing patterns and time spent to estimate better fees based on actual work data (not just reported data), providing a clearer picture of resource allocation by client or matter.

THE FUTURE OF LEGAL BILLING AUTOMATION

AI's potential for transforming legal billing is immense, particularly in areas that do not directly affect billable work. Automating time-tracking frees lawyers to focus on high-value, client-facing tasks while improving the accuracy and transparency of bills. By shifting the focus from billable activities to the business of law, law firms can benefit from AI without eroding their bottom line.

Legal billing is poised for transformation in a field that has historically resisted change. As AI technology continues to evolve, the "forgotten" task of time-tracking may prove to be one of the most valuable applications of AI in legal tech. Through automation, law firms can streamline operations, enhance client relationships, and reimagine the way they manage the business of law. **ILTA**

NANCY JENG

is the co-founder of Billables AI, an automated time-tracking platform for lawyers and other professional service providers. Nancy spent the first decade of her career in the professional service industry, running communication strategies at several national advertising firms. After leaving the ad agency world (and time sheets) behind, Nancy established her career in tech, starting at Google where she oversaw SMB content marketing. Prior to Billables AI, Nancy was leading global product marketing at Pinterest where she helped build and launch productivity and monetization products for advertisers, consumers and creators.

MODERNIZING LEGAL WORKFLOWS

The Role of AI, Automation, and Strategic Partnerships

BY SCOTT ANGELO, JARED GULLBERGH, NANCY GRIFFING & MICHAEL OWEN HILL

As law firms transition from exploring artificial intelligence (AI) to fully adopting and implementing emerging technologies, developing a deep understanding of data and processes is imperative. Buchanan Ingersoll & Rooney PC recognized the value of partnering with industry experts, like NetDocuments and 3545 Consulting, to integrate AI-powered tools into the firm's day-to-day operations. Through this collaborative process, Buchanan implemented NetDocuments' PatternBuilder and App Builder (powered by PatternBuilder MAX) to integrate advanced document automation and AI capabilities to streamline processes and create measurable efficiency improvements.

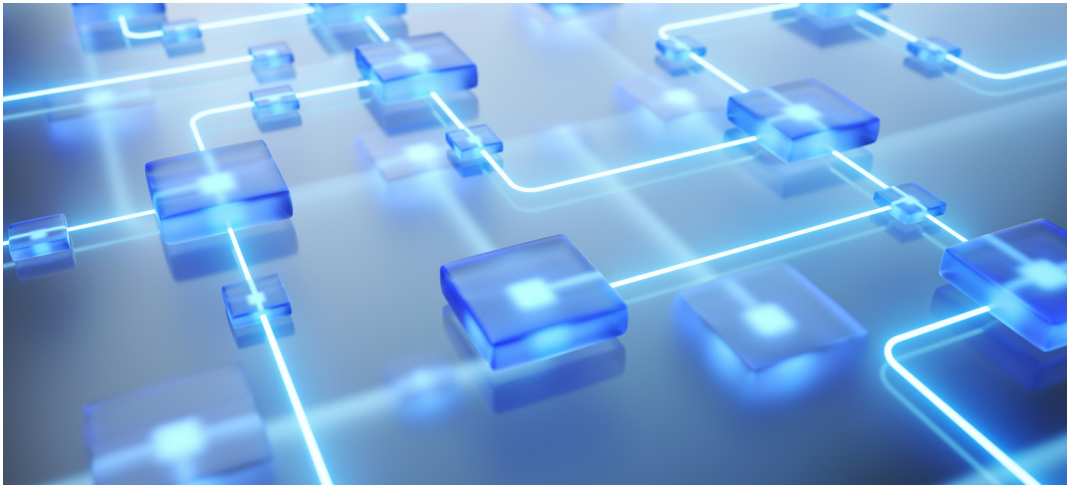
TURNING AMBITION INTO ACTION

In 2023, Buchanan Ingersoll & Rooney PC faced the challenge of staying ahead in a rapidly evolving legal industry. Recognizing the increasing importance of AI, automation, and data analytics, the firm launched a bold five-year strategic plan to transform its operations and better serve client needs. This plan aimed

to align cutting-edge technologies with the firm's workflows, ensuring seamless integration into daily practice.

Scott Angelo, the firm's Chief Information Officer, emphasized the importance of a holistic approach: "We felt that it was very important to build out centers of excellence that covered all three areas [AI, automation, data & analytics]. Because we realized that in order to come up with comprehensive solutions, we were going to have to use ... two to three of those working together to provide the best solution possible." Angelo added, "We really doubled down on AI because it was just so new—not just to the legal industry, but to the world." Under his leadership, Buchanan's efforts to embrace AI have garnered significant attention, earning the firm recognition as one of the "Best of the Best for Generative AI" in the 2024 BTI "Leading Edge Law Firms" survey.

This acknowledgment reflects more than ambition; it highlights the firm's ability to translate innovative ideas into actionable results. By focusing on collaboration and leveraging technology to address client demands, Buchanan has set a benchmark for what is possible in legal technology innovation.



The NetDocuments App Builder quickly became integral to the firm's strategy, enabling them to automate complex legal workflows with AI. By generating documents, extracting data, and streamlining legal processes, the tool delivered immediate efficiency gains and allowed the firm to focus on higher-value work.

A key advantage of the App Builder and PatternBuilder lies in their secure operation within the firm's existing NetDocuments environment. This integration ensures that all workflows and data comply with the robust security standards already established in the document management system. Leveraging this trusted infrastructure, firms like Buchanan can confidently explore AI-driven automation without risking sensitive client or case information.

App Builder offers an approachable starting point for firms new to AI. Its low-code interface and integration with familiar platforms reduce the learning curve, making the transition to AI and automation more accessible. Its unique combination of security, usability, and seamless integration positions App Builder as an ideal tool for incorporating AI-driven technology into legal operations.

To ensure they fully utilized App Builder's capabilities, Buchanan partnered with 3545 Consulting, a firm known for its expertise in legal technology, to bridge the gap between vision and execution. Together, the team set out to modernize workflows, reduce inefficiencies, and ultimately enhance client service—all while maintaining the highest data security standards.

TIME SAVINGS AND CONSISTENT WORK PRODUCT

Buchanan's journey started with a focus on small, high-impact initiatives. Collaborating with 3545 Consulting, they identified key use cases where AI and automation could deliver immediate value, utilizing NetDocuments' App Builder. One of the first successes was using the Summarize app, which automated the process of summarizing legal documents. This tool quickly became a favorite among the firm's attorneys, saving hours of manual work.

Buchanan shifted from built-in apps to customizing a Litigation Master Caption File app that automates the creation of master caption files for litigation matters. The app leverages AI to extract caption data points from an initial court filing and automation to generate a master caption file with jurisdiction-specific formatting. The master caption file can then be used as the starting point for all court filings, ensuring consistency, accuracy, and proper formatting, all while streamlining the drafting process.

Outcomes:

- 1500+ practice assistant and paralegal hours saved annually (approx. $80,000 value)
- Risk avoidance and reputational benefit via accurate, consistent automation
- Improved consistency for 10,000+ court filings annually

Buchanan also created a Durable Power of Attorney App to streamline the drafting process for durable powers of attorney. The app guides users through client-specific questions, such as selecting appointees, defining their powers, and determining whether the document should take effect immediately or upon incapacity. By automating this process, the app reduces drafting time from hours to minutes, enabling attorneys to focus on each case's unique complexities.

Ensuring that measurements were in place to access each app's value was critical to Buchanan's approach. The firm worked with 3545 Consulting to develop a simple yet effective solution: by integrating a counter in each app's data table, the firm could track the app's usage. This tracking system allowed them to quantify the time savings and other benefits generated by each app.

For example, one app saved the firm five hours of attorney time each time it ran. Its utilization increased productivity and improved the consistency and quality of the final documents. Other apps provide less tangible benefits, such as reducing errors and ensuring compliance with standard formats across all outputs.

It is also easier for attorneys to adopt since the apps are built within the familiar NetDocuments platform. Attorneys' familiarity with the existing platform reduced the learning curve, making AI and automation feel less intimidating and encouraging greater adoption across the firm's practice areas.

TURNING USE CASES INTO WORKING APPS

Buchanan's success was not an accident – it stemmed from a well-defined process critical to the success of implementation and execution. The collective team followed these essential steps for app development:

1. IDENTIFY AND PRIORITIZE USE CASES

Workshop with attorneys and staff to identify tasks ripe for automation, such as repetitive document drafting. "We started with smaller groups to get quick wins and build excitement," said Jared Gullbergh, Director of Practice Solutions and IG at Buchanan.

2. DEFINE APP REQUIREMENTS

Collaboratively create clear goals and workflows. PatternBuilder's low-code interface allowed for quick iterations and user involvement.

3. LEVERAGE PRE-BUILT STUDIO APPS AND TEMPLATES

Customize templates to match the firm's needs. This step reduced development time and increased overall value.

4. INCORPORATE AI AND AUTOMATION

Design apps to output consistent, high-quality documents by combining AI-driven data extraction with automation.

5. TEST AND ITERATE

Involve end users early to refine solutions.

6. DEPLOY AND TRAIN

To foster adoption, create training sessions for end users to ensure they understand the apps' value and functionality.

7. MEASURE SUCCESS

Track usage and time savings to quantify impact and identify additional opportunities for improvement.

BUILDING A LONG-TERM STRATEGY THROUGH COLLABORATIVE THOUGHT LEADERSHIP

Buchanan Ingersoll & Rooney's thoughtful, collaborative approach to AI can deliver lasting results.

"We've only scratched the surface of what's possible," said Gullbergh. "With NetDocuments' AI capabilities, we're not just keeping up — we're leading."

Beyond the immediate wins, Buchanan's collaboration with NetDocuments and 3545 Consulting set the stage for sustained success. By equipping Buchanan's IT team to develop and maintain apps independently, the firm ensured it could scale its automation capabilities while still having access to external support from 3545 when needed.

The knowledge gained from this collaborative project is incredibly valuable for other firms transitioning from the theoretical ideation phase to the practical implementation stage of AI-powered technology deployment. The right technology, a strategic approach, and a focus on measurable outcomes can transform legal workflows. Buchanan's journey is an inspiring example of embracing the future while prioritizing client needs and upholding firm values.

As legal technology evolves, the firm's proactive approach is a valuable case study for other firms looking to implement AI and automation. The lessons learned from this collaboration — such as the importance of having an intelligent DMS, the proper training, measuring success, and building internal and external capacity — offer a roadmap for firms seeking to modernize their workflows while delivering value to their clients. **ILTA**

SCOTT ANGELO

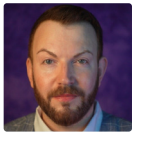
is Buchanan Ingersoll & Rooney PC's Chief Information Officer. He handles the design, implementation, and management of information technology and risk management platforms. Scott manages the firm's IT team to further develop technologies to drive creative solutions for the firm and its clients. He brings more than 30 years of experience to the firm's high-performing culture and efficient client service through an innovative digital infrastructure. Scott was recognized by the Pittsburgh Technology Council as the winner in the 2024 CIO of the Year Awards and by The Legal Intelligencer as a Law Firm Innovator Finalist for the 2024 PA Legal Awards.

JARED GULLBERGH

serves as Buchanan Ingersoll & Rooney PC's Director of Practice Solutions and Information Governance. Jared partners closely with the firm's legal and government relations practices to deliver technology and data-driven solutions to maximize process efficiencies, profitability and to deliver unique client-value. He also directs and oversees the Records and Information Governance functions of the firm. Jared brings to bear more than 19 years of IT and information governance experience in the full-service legal industry.

NANCY GRIFFING

is a seasoned legal technology consultant and a driving force at 3545 Consulting, where she combines decades of industry experience with a forward-thinking approach to innovation. With a deep understanding of the legal sector's unique challenges, Nancy has become a thought leader in the adoption and integration of artificial intelligence. She specializes in helping law firms leverage AI to enhance existing tools, streamline workflows, and drive efficiency at scale. As a Partner at 3545 Consulting, Nancy empowers firms to achieve transformative results while navigating the complexities of modern legal practice.



MICHAEL OWEN HILL

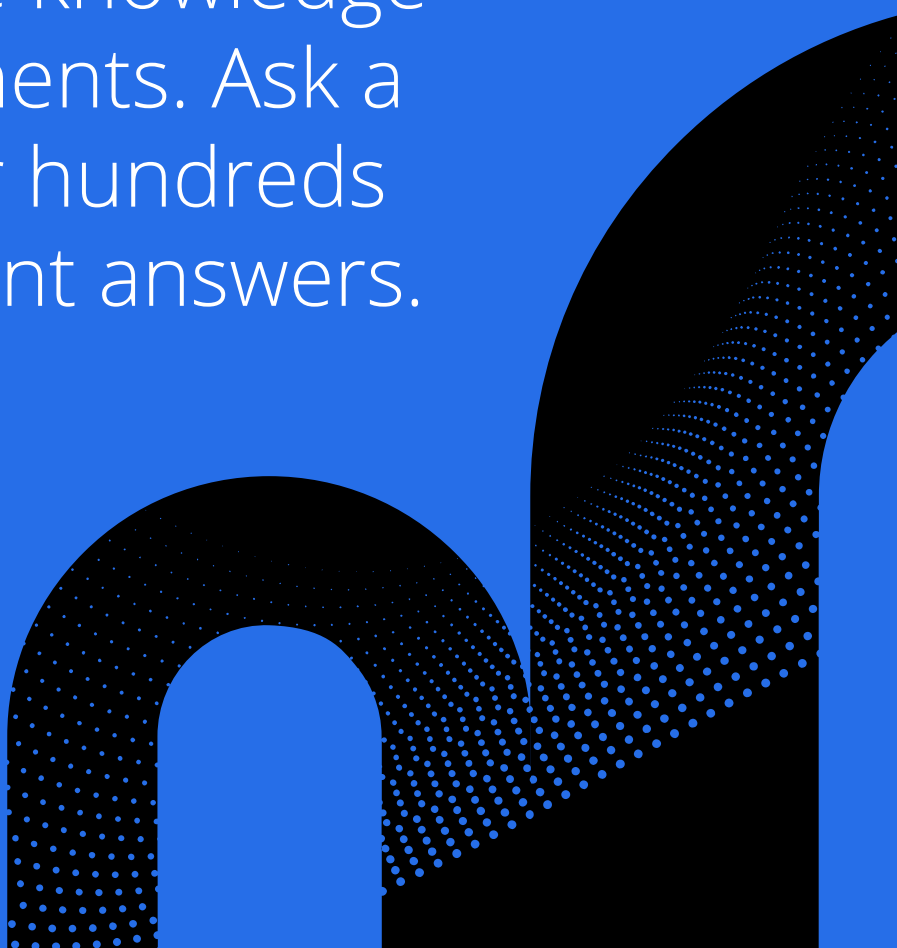
has nearly two decades of experience in the legal tech industry in roles ranging from product management to product marketing to product portfolio strategy. He has worked with small firms, corporate legal departments and the world's largest law firms to advance client development, legal research and know-how, financial and practice management and legal document management goals. As Director of Product Marketing at NetDocuments, Michael focuses on helping customers, partners and employees navigate the complexities of the rapidly evolving legal tech landscape.

netdocuments®

Simply Intelligent

Seamless AI
Powerful Workflows
Smarter Experiences

Introducing the
ndMAX Legal AI Assistant
for an easy to use conversational way to tap into the knowledge within your documents. Ask a question of one or hundreds of files — get instant answers.



BOOK A DEMO >

THE LEGAL IMPACT OF DATA DISPOSAL

Minimizing Risks Through Data Disposition

BY FAHAD DIWAN & ANAMIKA KUMAR

Data is the backbone of modern businesses, underpinning transactions, decisions, and interactions. While organizations invest significantly in data collection, processing, and storage, the final stage of its lifecycle—disposal—is often overlooked. Yet, improperly handling obsolete data can result in legal disputes, regulatory penalties, reputational damage, and erosion of customer trust.

Even if your organization employs robust cybersecurity technologies, adheres to privacy and security policies, and enforces other compliance processes and procedures, failing to dispose of unnecessary or outdated data securely could undermine all these efforts and expose your organization to significant risks. The consequences of such oversights are far-reaching and present in numerous high-profile incidents outlined below.

This article discusses the reasons that necessitate the secure disposition of unnecessary data, including regulatory requirements and the risks of non-compliance, as well as the practical strategies for effectively implementing data disposition within an organization.

THE NECESSITY OF DATA DISPOSAL

The early stages of the data lifecycle—collection, analysis, and use—often receive significant attention and investment from organizations. However, retaining data beyond its purpose introduces substantial risks, turning an asset into a liability. Over-retaining unnecessary data can expose organizations to the following threats:

1

DATA BREACHES

The larger an organization's data holdings are, the greater its attack surface for potential breaches. Retaining unnecessary or outdated data increases exposure risks for sensitive information within an organization's data holdings.

2

LEGAL VIOLATIONS

Many privacy and cybersecurity laws, including those exampled below, explicitly require data disposition once it has outlived its purpose. Organizations must demonstrate that they only retain data if it is legally necessary or operationally justifiable. Failure to comply can lead to regulatory fines and penalties.

3

REPUTATIONAL DAMAGE

Today's consumers expect companies to respect their privacy, including the timely disposition of their personal data. When organizations fail to meet this expectation, they risk eroding customer trust. This reputational harm can lead to customer churn, lost business opportunities, and long-term brand damage.

LEGAL REQUIREMENTS FOR DATA DISPOSITION

As alluded to above, data disposition is not merely a best practice but a legal obligation enshrined in many regulations. The following are some of the privacy and cybersecurity regulations requiring data disposition.

The EU's GDPR

The General Data Protection Regulation (GDPR) in the European Union requires companies to periodically dispose of data that they have no legal or business reason to retain:

- **Data Minimization Principle (Article 5(1)(c)):** Organizations must limit the collection and retention of personal data to what is strictly necessary for the purpose for which it was processed. Once the data is no longer needed for the purpose for which it was collected, organizations must dispose of it. Failure to dispose of unnecessary data constitutes a violation of this principle.

- **Right to Erasure (Article 17):** Often referred to as the "Right to Be Forgotten," this provision gives individuals the right to request the disposition of their personal data. Organizations are obligated to erase such data unless legal or operational needs justify retention.

The GDPR has strict penalties for non-compliance. Fines can reach up to €20 million or 4% of annual global turnover.

California's CCPA

The California Consumer Privacy Act (CCPA) similarly emphasizes the importance of data disposal:

- **Data Minimization Principle (Section 1798.100(c)):** While not explicitly termed as such, the CCPA requires businesses to collect, use, and retain personal information only as reasonably necessary and proportionate to achieve the disclosed purposes. Businesses must dispose of data once its purpose has been fulfilled.

- **Right to Deletion (Section 1798.105):** Consumers can request the deletion of their personal information. Businesses must comply unless specific exemptions apply, such as for legal obligations or ongoing business needs.

Organizations risk fines of up to $2,500 for each unintentional violation and $7,500 for each intentional violation under the CCPA. Notably, these fines are calculated per violation, which means they can quickly escalate to substantial amounts. For instance, if an organization fails to dispose of 100,000 records properly, the fine would be the applicable dollar amount multiplied by the number of records.

The USA's HIPAA

In the U.S., the Health Insurance Portability and Accountability Act (HIPAA) establishes precise requirements for the disposal of protected health information (PHI):

- **Secure Disposal Requirements (45 CFR § 164.310(d)(2)(i)):** HIPAA mandates that covered entities implement policies to ensure the secure destruction of PHI when it is no longer needed.

Penalties under HIPAA's enforcement rules range from $100 to $50,000 per violation, with annual maximums reaching $1.5 million depending on the level of negligence. Similar to the CCPA mentioned above, these fines are calculated per violation and can, therefore, escalate quickly.

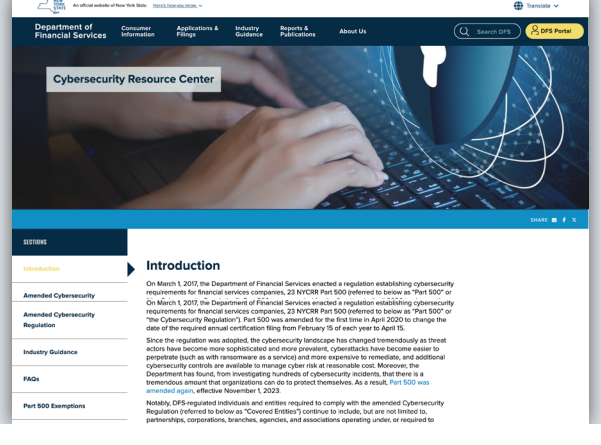
> Penalties under HIPAA's enforcement rules range from $100 to $50,000 per violation.

NY DFS Cybersecurity Regulation

The New York Department of Financial Services (NY DFS) Cybersecurity Regulation (23 NYCRR 500) mandates secure data disposal:

- **Disposal of Nonpublic Information (Section 500.13):** Covered entities are required to implement policies and procedures for securely disposing of nonpublic information once it is no longer needed for business operations or legal purposes.

Non-compliance with this regulation can result in substantial penalties for covered entities. The NY DFS is authorized to impose civil penalties of up to $5,000 per violation daily. Accordingly, these fines can accumulate rapidly, leading to significant financial consequences for organizations that fail to adhere to the regulation's requirements.



NY DFS CYBERSECURITY RESOURCE CENTER

https://www.dfs.ny.gov/industry_guidance/cybersecurity

THE COST OF MISHANDLING DATA DISPOSAL: LEARNING FROM REAL-WORLD CASES

Neglecting to follow these regulatory requirements can lead to a cascade of adverse outcomes, including data breaches, legal disputes, regulatory fines, and reputational damage. Several high-profile incidents underscore the critical importance of disposing of unnecessary data.

VISION CARE COMPANY'S 2020 DATA BREACH

In 2020, a prominent vision care company experienced a data breach involving a shared email account containing six years of customer data. The breach compromised the personal and medical information of approximately 2.1 million individuals. Investigations revealed that the company had retained unnecessary data longer than needed and lacked adequate security measures. As a result, the New York State Department of Financial Services fined the company $4.5 million for violations of its cybersecurity regulation.

ONLINE RETAILER'S 2019 DATA BREACH

In 2019, a popular online retailer suffered a data breach affecting millions of customers and received a $500,000 fine from the U.S. Federal Trade Commission (FTC). The FTC noted that the company had retained personal information indefinitely without a legitimate business need, which increased the breach's impact. The company faced regulatory action partly due to its over-retention practices.

CLOUD COMPUTING PROVIDER'S 2020 DATA BREACH

In 2020, a cloud computing provider experienced a ransomware attack in early 2020. The attackers accessed extensive personal information, including Social Security and bank account numbers. The FTC criticized the company for retaining data longer than necessary, contributing to the breach's severity. As part of a settlement, the regulator required the company to delete unnecessary data and improve its data retention policies. So far, the company has received approximately $60 million in fines from various regulators concerning this breach.

UNDERSTANDING DATA DISPOSAL METHODS

To comply with the data disposition requirements of the above regulations, companies must ensure that data disposal is permanent and renders the disposed data irretrievable. Techniques for ensuring this level of responsible data disposal include:

Cryptographic Erasure

Encrypting data at the time of storage and securely deleting the encryption keys when the data is no longer needed, rendering the data mathematically irretrievable.

Physical Destruction

This describes incinerating hard drives or degaussing magnetic media.

Certified Digital Erasure

Using specialized software to overwrite data on storage devices multiple times, eliminating all traces of the original data.

These methods comply with regulatory requirements for secure and irreversible data disposition.

OTHER METHODS USED BY ORGANIZATIONS

Many organizations, however, use other techniques to remove personal and sensitive data from their databases. Organizations use these methods because they help balance the need to mitigate data retention risks while still extracting value or retrieving the data if needed. Below, we define these techniques, highlight their strengths, and discuss their limitations.

ARCHIVING

Definition: Archiving involves transferring inactive data to a dedicated storage system, potentially managed by a third party for long-term retention. It is not inherently a method of disposal.

Strengths

Organizations can access archived data when necessary, such as for legal or business purposes.

Archiving helps mitigate some risks of retaining data by segregating inactive information from active systems.

Data can be archived even before its retention period has been met since archiving does not constitute permanent disposal.

Weaknesses

Archived data can still be considered retained data, and keeping it beyond the mandated retention period can result in legal violations.

Organizations must still dispose of archived data once it is no longer justifiable to retain it for legal or business purposes. Retaining unnecessary archived data is a source of liability for the organization.

ANONYMIZATION

Definition: Anonymization involves irreversibly removing identifiers from data, ensuring it cannot be traced back to an individual.

Strengths

Can comply with regulatory frameworks' requirements for data disposition as it renders data non-personal.

Allows organizations to continue to derive value from the data (e.g., anonymized data is well-suited for research or statistical purposes where the identity of individuals is not required).

Weaknesses

Certain jurisdictions, including those under the GDPR, mandate robust proof of irreversibility for anonymized data. Failure to meet these standards may result in anonymization being deemed insufficient to satisfy regulatory data disposal requirements.

Improperly anonymized data can sometimes be re-identified through advanced techniques or by cross-referencing it with other datasets. As a result, anonymized data compromised in a data breach can still pose significant liability risks for the organization.

PSEUDONYMIZATION

Definition: Pseudonymization replaces identifiable information with placeholders (e.g., tokens or codes) that allow controlled re-identification if necessary.

Strengths

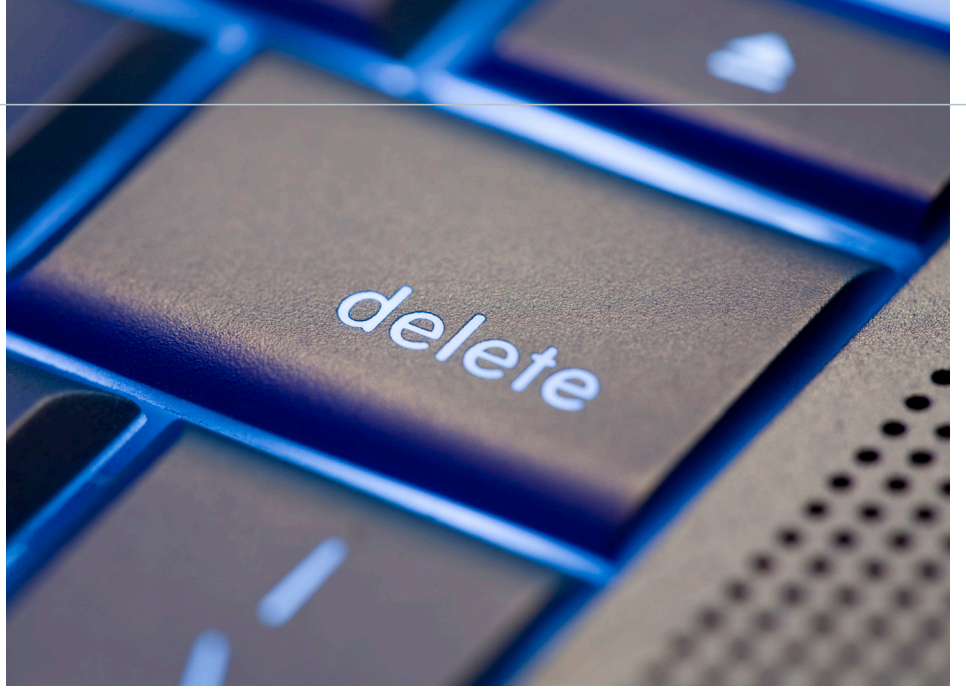
Enhances data security by reducing exposure of identifiable information.

Allows continued utility of the data for specific purposes while protecting privacy.

Weaknesses

Pseudonymized data is still considered personal data under GDPR and other regulations, requiring continued safeguards and eventual secure disposal.

Organizations retaining pseudonymized data that is not legally necessary or operationally justifiable may still be found liable for non-compliance with data disposition requirements.



A GUIDE TO SECURE AND COMPLIANT DATA DISPOSAL

Navigating the complexities of data disposal requires a structured and thoughtful approach. Organizations must move beyond ad hoc efforts and adopt a systematic strategy to ensure compliance, reduce risks, and protect sensitive information. Here is a step-by-step guide to getting it right:

STEP 1

Build a Comprehensive Data Catalog

The cornerstone of effective data disposal is a complete data catalog. A data catalog identifies the databases containing personal and sensitive data, the data types stored within them, and how the organization uses them.

Organizations can create a data catalog manually or through automated data discovery solutions:

- **Manual Method:** This involves interviewing employees to identify which databases contain personal and sensitive data, the types of data within them, and how they use it. This process often requires significant time and effort from employees. Moreover, it is prone to errors because employees may provide incomplete or inaccurate information about the data and its usage.
- **Automated Method:** Modern software solutions allow organizations to build a data catalog using software solutions efficiently. Leading data discovery software can automatically identify databases with personal and sensitive data, classify the data within them, and even determine usage to a certain extent.

Regardless of the method used, organizations may want to build a data catalog incrementally to reduce the burden on organizational resources. Organizations can begin with high-risk databases—those most likely to contain significant volumes of sensitive data, such as Social Security numbers or credit card information. Once these are cataloged, organizations can expand to lower-risk systems, gradually building a comprehensive catalog. This approach allows organizations to significantly mitigate risk without overburdening organizational resources.

STEP 2

Identify Data Unnecessary for Legal Reasons

A standalone data catalog lays a solid foundation, but organizations must identify what they can dispose of. Organizations can defensibly dispose of data they no longer have a legal obligation to retain. There are two primary reasons for retaining data for legal purposes: Retention Regulations and Legal Holds.

- **Retention Regulations:** Many jurisdictions require organizations to

retain certain types of information for a specific period (e.g., payroll records must often be retained for seven years under applicable laws). Organizations typically create data retention schedules to track retention requirements. A data retention schedule lists all the different categories of records an organization holds and links these records to the applicable retention requirements in various regulations.

- **Legal Holds:** Legal holds apply when an organization must retain information relevant to ongoing litigation involving the organization.

Organizations should link their data catalog to their data retention schedule and legal hold list to address these requirements. Organizations can create these links manually or through automated means:

- **Manual Process:** This involves stakeholders collaborating across departments, including records governance, legal, and IT, to establish these links. Similar to the manual process for building a data catalog, the manual process can be resource-intensive and fraught with errors.

- **Automated Solutions:** Certain software solutions can automate this process. Some providers offer data discovery solutions integrated with legal hold and data retention solutions. This integrated offering allows organizations to automatically link information from the data catalog to retention schedules and legal hold lists, enabling faster and more accurate identification of data that no longer needs to be retained for legal reasons.

STEP 3

Dispose of Unnecessary High-Risk Data

Once unnecessary data is identified, organizations can begin to dispose of it. Organizations can significantly reduce risk while minimizing the workload on their employees by starting with disposing of high-risk data, such as sensitive personal information. Sensitive personal data is any data likely to cause significant harm to an individual if it were processed illegally. Examples include:

- Credit card numbers
- Social Security numbers
- Information about job performance or criminal records

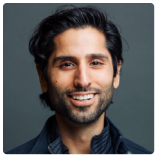
During the disposal process, organizations should document key details, including:

- The rationale for deletion, such as expired retention periods.
- Detailed records of the disposition, including certificates of destruction and logs of the deletion process.

This documentation is essential for audits, litigation, and regulatory compliance. Moreover, to safely comply with the disposition requirements in applicable privacy and cybersecurity laws, organizations should consider disposing of this data using the techniques outlined in the 'Methods that Comply with Data Disposition Requirements' section above.

CONCLUSION

Periodic and secure disposal of unnecessary data is essential for reducing data risk and ensuring compliance with privacy and cybersecurity regulations. Although this may initially appear daunting and resource-intensive, prioritizing high-risk data and using leading software solutions can significantly reduce the time and effort involved while effectively mitigating risks to the organization. ILTA



FAHAD DIWAN

is the Director of Product Management (Privacy and Data Governance) at Exterro. He holds a J.D. and a B.C.L. from McGill's Faculty of Law and an Accounting and Finance degree from Concordia University. Fahad is a Fellow of Information Privacy with CIPP/C and CIPP/M designations from the International Association of Privacy Professionals. Fahad creates software solutions that automate privacy compliance, as well as AI and data governance tasks, like data discovery, assessments manager, data retention, data subject rights request, and consent management.

ANAMIKA KUMAR

is a seasoned content and PR professional at Exterro, overseeing and driving impactful projects across South Asia while contributing to global initiatives. She specializes in thought leadership content and strategic PR campaigns that elevate brand visibility and foster meaningful engagement with diverse audiences. Anamika brings a keen eye for storytelling and a passion for crafting narratives that resonate on a global scale. Outside of work, she enjoys spending time with animals and embarking on new adventures, always seeking inspiration from the world around her.

HOW LAW FIRMS ARE MAXIMIZING GEN AI INVESTMENTS THROUGH ROBUST DATA GOVERNANCE

BY MIN CHEN

Robust data governance is a requisite for business in today's digital age. Data governance might seem overly restrictive or burdensome for some, but it can facilitate and accelerate the enterprise's implementation and adoption of Generative AI (Gen AI). A robust data governance program can give organizations a significant competitive advantage when implementing Gen AI-powered technology successfully.

According to Deloitte, expectations of Gen AI benefits remain high, but many executives feel pressured to realize its value quickly. Given the growing shortage of skilled AI professionals, combined with the fact that many Gen AI

implementations stall at the pilot or proof-of-concept phase, having a robust data governance program can help organizations streamline Gen AI implementation and improve its output, tap its transformational benefits faster while minimizing risk, achieve a much greater return on their investments, and achieve their corporate objectives.

The argument for enterprise organizations to "ground" Gen AI models on internal or industry data vs. the Internet is that it leads to highly focused outputs and potentially fewer hallucinations. The other side of this argument is that the voracious data requirements of today's Gen AI models can increase organizational risk if given keys to the proverbial digital kingdom without already having rock-solid data governance policies and initiatives.

LEVERAGING DATA GOVERNANCE TO IMPROVE GEN AI OUTPUT

How can organizations most effectively leverage Gen AI models while also adhering to data governance best practices? Here are some things to consider:

- **Data Quality:** Ensuring data is clean, accurate, and unbiased is crucial for authoritative Gen AI outputs, but large internal datasets or external industry data can be challenging to clean, validate, and normalize. Using the previous types of large internal or external datasets necessitates additional steps to ensure the data is accurate, reliable, and credible and minimize the chance of hallucinations. For example, if grounding data is biased, Gen AI outputs will reflect those biases, resulting in misleading, discriminatory, or unfair outcomes.

Additionally, it is crucial to continuously monitor and update datasets to maintain accuracy and reliability over time.

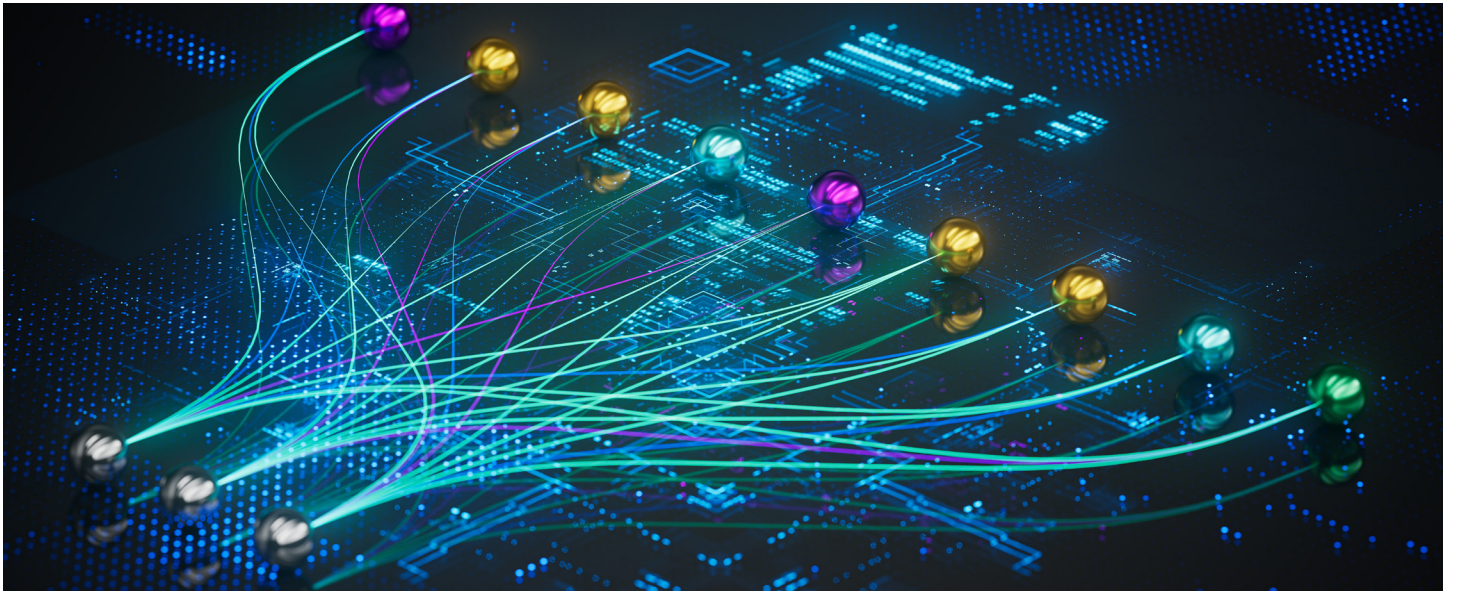
Robust data governance establishes clear data quality standards that ensure consistency, eliminate unnecessary duplication, and streamline data-related processes. It does this by assigning data stewards and domain experts, whose sole responsibility is ensuring data quality, security, and adherence to stated policies. AI tools can help supplement some of this work, but data stewards and data domain experts are needed to provide critical humans-in-the-loop quality control. For example, AI can overhaul the quality assurance process, automating and significantly refining data integrity with minimal direct human involvement.

However, data stewards should always have the final say regarding data governance rules.

- **Data Privacy, Security, and Compliance:** Internal datasets used to train Gen AI models often contain sensitive personal information (PII) or intellectual property (IP), raising significant privacy concerns around data collection, storage, and usage. Concerns around Gen AI-powered technology's impact on PII and IP are present within industries ranging from finance and healthcare to specific functions like marketing and legal – all of which benefit from AI's ability to process and interpret data at scale. Additional concerns include legal exposure from using copyrighted material for model training or grounding or if potentially sensitive organizational data escapes the walled garden and is used to train foundational AI models.

Data governance helps by establishing clear rules and enforcing security measures related to who (or what) can access specific data and when. It also identifies and tags potentially sensitive data and mandates that organizations have the appropriate policies for encryption, storage, retention, and deletion, data access and handling procedures, and other controls (e.g., data anonymization, pseudonymization) in place to mitigate risk and meet legal and global regulatory requirements (e.g., GDPR, HIPAA, and CCPA).

For example, in healthcare, the same data governance policies and procedures that protect patient privacy would also prevent Gen AI models from accessing or compromising their data. In the case of copyrighted material, data governance would ensure they have explicit permission to use the content for Gen AI model training, grounding, or other outputs before the system ingests the data. These data access and handling policies and other legal and IT security practices will help organizations remain nimble in a rapidly changing regulatory environment and prevent unintentional data leakage. They can also shield organizations against AI copyright infringement issues by reviewing all external data – whether licensed or publicly available – for any conditions, restrictions, or possible violations that could stem from its use.



- **Data Scale and Integration:** Datasets are growing exponentially and becoming more complex, thanks partly to duplicate, unstructured, and multi-modal data (e.g., video, images, audio). Training or grounding Gen AI models becomes more challenging, costly, and time-consuming on the back end without proper data preparation on the front end (e.g., tagging, metadata, de-duping, homogenizing formats).

  Data governance facilitates this by establishing universal data definitions and formats, developing standardized processes, and guiding data integration tool selection and use. Such processes help ensure the seamless integration of data across different sources, enabling organizations to effectively utilize large volumes of data from diverse sources for AI applications without extensive data cleansing or transformation processes.

- **Transparency and Explainability:** Aside from hallucinations, one of the biggest concerns about Gen AI is the "black box" obscurity of how large language models (LLMs) source the information they use to generate their results. The mistrust and doubt generated in response to the sourcing ambiguity around many LLMs is a frequent hindrance to Gen AI adoption, leading to decreased efficiency and productivity, the inability to effectively capitalize on institutional or industry knowledge, reduced competitiveness and customer service, and wasted Gen AI investments.

Robust data governance entails documenting the pedigree of its sources, data types, how it was collected and processed, and intended usage. While this level of detail may be invisible to end-users, Gen AI applications can use it to present precise citations and direct links to source materials next to the AI-generated responses.

Explainable AI (XAI) tools can also provide users (or regulators) with insights into the factors influencing a model's output, allowing for a deeper understanding of how the AI system generated its outputs. These tools can identify what specific data or source material had the most significant impact or influence on the model's production. This ability enhances contextual understanding of and confidence in how the model generated its results and could potentially illuminate any existing biases, leading to better-informed decision-making.

- **Responsible AI Use:** Organizations need to fully understand and consider the real-world impact of Gen AI on employees, customers, partners, and end-users to prevent the creation or reinforcement of unfair bias, discriminatory outputs, or other harmful consequences. This process entails a deep understanding of how Gen AI models are created, trained, implemented, and used – by whom and for what purpose – and extensive internal planning.

Data governance policies and procedures ensure that data used to train AI models is accurate, diverse, unbiased, and handled responsibly. This helps mitigate risks like privacy or copyright violations, discrimination, and inaccurate or misleading information.

These are examples of ways organizations can leverage data governance policies to improve Gen AI output and protect their organizations. Other steps they can take to enhance their organization's Gen AI applications include:

- Employing Retrieval-Augmented Generation (RAG) technologies to keep Gen AI focused on specific datasets for specific queries or use cases. Incorporating RAG allows the model to dynamically retrieve relevant information from a designated data source before generating a response, ensuring the output aligns with the specific context of the query and the specified dataset.

- Establishing IT frameworks that enable the easy/rapid switching of AI models without extensive retooling of applications or data, as well as regulatory frameworks to facilitate compliance and streamline reporting.

- Collaborating with trusted partners that leverage proven tools, Gen AI-approved data, and licensed

content, and have a strong track record of data privacy, security, and responsible AI practices.

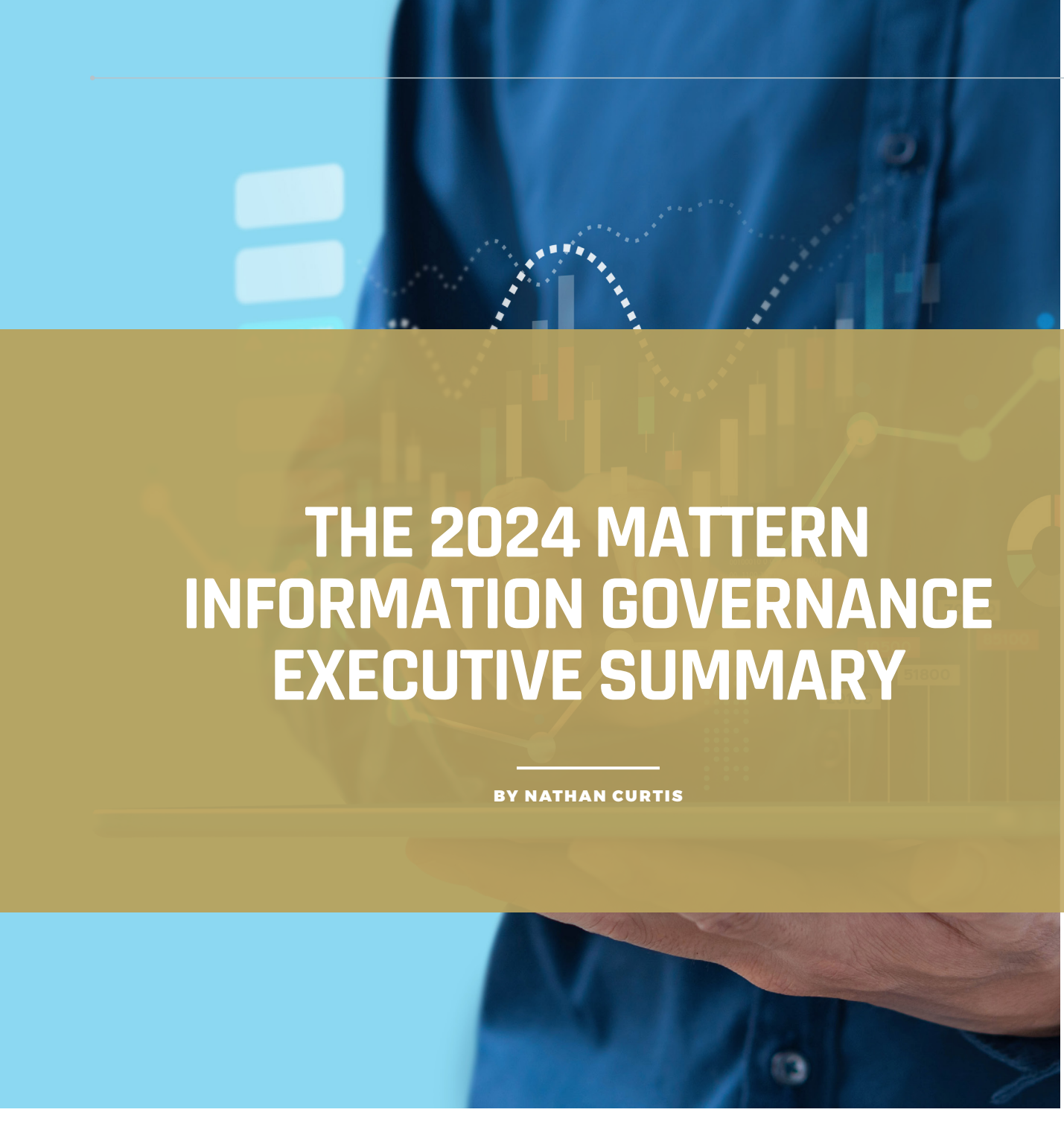
- Employing comprehensive data handling processes, such as data acquisition and preprocessing, data augmentation to enrich and expand datasets, end-to-end data tracking and monitoring, data security and privacy by design, and adherence to responsible AI principles, ensuring data quality, integrity, and ethical practices.

- Utilizing cloud computing platforms to democratize everyone's access to Gen AI technologies. For instance, AWS Bedrock provides a secure and streamlined environment to leverage rapidly evolving LLMs, facilitating quick adoption and saving operational management efforts, enabling the data scientist teams to focus on innovation experiments and delivery.

Organizations planning to train or ground Gen AI models on internal data must have robust, mature data governance policies to implement Gen AI more quickly, leverage it more effectively, mitigate any possible risks, and maximize the return on their investment. The transformation of Gen AI has genuinely raised the bar for building a robust data platform and infrastructure to an unprecedented level. The investments made in strengthening data governance – including ensuring data quality, privacy, security, transparency, and explainability; leveraging data stewards as humans-in-the-loop; and using internal policies to guide responsible AI use – will enhance an organization's ability to successfully tap into AI's transformative capabilities and achieve their business objectives. **ILTA**

MIN CHEN

As VP and CTO of LexisNexis Global Platforms, Asia, Pacific, Min Chen is responsible for driving and delivering customer-centric legal innovations across the globe through artificial intelligence, data analytics, visualization, big data processing, cloud native solutions and other advanced technologies. She leads 600+ engineers and product people in Asia, Pacific and US, with core teams in Shanghai, Raleigh and Gurgaon.



THE 2024 MATTERN INFORMATION GOVERNANCE EXECUTIVE SUMMARY

BY NATHAN CURTIS

Law firms routinely handle large amounts of data, both in hard copy and electronic form, with most of this data being sensitive and/or confidential, driving the need for robust information governance (IG) programs. Chief Legal Officers rank key components of a comprehensive IG program, such as cybersecurity, regulatory compliance, and data privacy, as the most critical issues they face year after year, according to the ACC CLO Survey in 2023.

IG programs are designed for the protection, management, security, and availability of a firm's information, fostering compliance with legal and regulatory requirements while simultaneously improving the firm's overall efficiency and productivity. At the core of all IG is the desire to maximize the value derived from data while mitigating the risks and roadblocks to compliance maintaining it presents.

An effective IG policies and procedures framework guides employees on appropriate data use. It details where to store information and how and when information should be disposed of (returned to the client, destroyed, or converted to vital record status). IG programs further help with regulatory and outside counsel guideline compliance, operational efficiency, and reducing discovery costs.

Mattern's purpose for developing the Information Governance Survey was to answer the primary question most firms pose. What are our peer firms doing in this area?

Throughout Mattern's engagements with firms of all sizes, this question commonly surfaces regarding various topics, including email retention, digitization programs, or how to curb data sprawl across network shares. We intend to keep the survey open to capture evolving IG trends and new survey questions derived from participant feedback.

We are pleased to share the results.

KEY FINDINGS

An overwhelming majority (94%) of respondent firms reported having some IG policy in place. Moreover, many firms (86%) have recognized the importance of IG by having positions within their respective firms dedicated to overseeing records/IG.

Cost considerations underscore the financial implications of effective data management. While a significant portion of firms surveyed noted cost-related factors as drivers for creating and implementing IG policies, cost is more of a concern for small firms than large firms.

Two examples are the cost of hard copy records in off-site storage, which is a concern for 36% of large firms and 57% of small, and the cost of data storage in document management systems, network shares, and email accounts, which is a cost concern for only 9% of large but 43% of small firms.

Beyond the financial implications, several additional challenges manifest themselves in the day-to-day implementation of successful IG programs, most of which are internal to the firm, regardless of size, but extend beyond the firm's walls as well. For example, over half (54%) of firms struggle to comply with outside counsel guidelines (OCG).

Even more telling is the overall substantial non-compliance rate of 44% and endemic cultures of "granting exceptions" reported by both large and small firms (31%). Extracting firms with mature IG programs and removing those without, we see that 47% are compliant with most IG policy elements.

Enforcement/compliance is the biggest challenge at firms of all sizes. Overall, only 4% of all respondent firms reported strict compliance with their IG policies (9% of large firms and 0% of small firms). Although mandating strict adherence to IG policies and procedures may seem like an easy fix, taking a step back reveals that the lack of enforcement/compliance is far more complex and includes other IG-related variables.

DATA ORGANIZATION

A significant percentage (27%) of law firms indicated no formal structure for their network share drive content, characterizing these repositories as the "Wild West." This lack of a meaningful folder taxonomy perpetuates poor IG practices in so many ways that information cannot be associated with specific clients or matters to apply appropriate retention and/or ethical walls.

Remediating information in network shares is a daunting task for any IT or IG professional, resulting in a "kicking of the can" event, in Mattern's experience, further exacerbating the

issue and associated risks. Fortunately, only 17% of firms with mature IG programs report this "Wild West" network share drive situation, with 67% forcing client/matter folder taxonomy at matter opening and another 8% not permitting the creation of these workspaces.

Also, a substantial percentage of firms mention challenges related to unstructured information in network shares, email accounts, and the inability to collaborate effectively (33%) or apply ethical walls (15%) outside the document management system.

These challenges suggest that firms strive to streamline information management for improved collaboration and security compliance. These values are also significantly lower in firms having a hybrid paper/electronic IG program (18% and 6%, respectively), due in all likelihood to their more mature IG program and information appropriately profiled in the DMS, where collaboration can take place while appropriate access rights are applied.

RETENTION

Additionally, regardless of how well a firm's data is structured, retention constantly struggles. Retention is relevant to many data repositories, including document management systems, eDiscovery databases, network shares, extranet file shares, lawyer, administrative, support staff email accounts, email archives, and more.

Likewise, clients communicate their concerns over retention practices by prescribing these topics in outside counsel guidelines. Complicating matters, a significant percentage of large and small firms (42%) indicated they currently have no strategy for limiting data sprawl, and they have not applied retention to electronic files in direct conflict with client wishes in OCGs.

Responses regarding what to keep and for how long differed significantly. The most common retention schedules law firms adopt for lawyer email accounts is no retention or unlimited retention (47%). This extends to maintaining former lawyer email accounts indefinitely (33%), and no data purge applied to support staff and administrative leadership email accounts (34%) even following departure (20%).

However, email retention practices have been improved for firms reporting that mature IG programs are in place. 31% have unlimited retention applied to current attorney email accounts, and the indefinite maintenance of departed attorney email accounts has been reduced slightly to 27%.

It should come as no surprise that the common practice of not deleting emails extends to email archives, where the most common retention schedule adopted by law firms is unlimited.

The most common retention schedule for extranet file share content is unlimited (37%). This means content in extranet file shares is not automatically deleted and is typically indefinitely retained. Interestingly, only 8% of firms reporting to have deployed comprehensive IG programs fail to address automated data purge in extranet file shares.

Moreover, the most common retention schedule law firms adopt for unstructured, user-assigned repositories is unlimited (57%). However, survey results show that this indefinite retention schedule reduces to only 23% of firms with comprehensive IG programs.

The indefinite retention of network share drive information is no surprise given the previously mentioned "Wild West" situation most firms describe, translating into a significant data mapping project to unravel. Many cases warrant deciding the degree of risk a firm will accept in purging information without structuring and aligning it with clients/matters.

For network share drive content structured in client/matter directories, the most common retention schedule adopted by the industry is once again unlimited. While this over-retention is concerning, the taxonomy quickly aligns content with the rest of the matter file's retention period or profiles this content to the document management system. Further, no firms with comprehensive IG programs reported unlimited retention applied to structured network shares, with 77% aligning these with the rest of the matter file's paper and DMS content.

In contemplating litigation related to retention, 44% of firms expressed concerns about the potential for being

subpoenaed to produce over-retained information that could harm their clients in future disputes. This risk factor highlights the importance of data retention and risk mitigation from another perspective: if you do not have it, you cannot produce it!

The survey also highlighted that a majority of clients (65%) are pushing for OCGs to include clearly defined retention limits regarding client-provided information and firm work product, and this push for retention applies to both large (64%) and small (50%) firms.

In summary, the importance of a compliant IG program cannot be overstated. Clients are requiring outside counsel to advance IG initiatives, namely ethical screens and data disposition, beyond a typical lawyer's comfort zone. However, it's hard to escape the reality that cybercrime continues to escalate, privacy regulations are expanding, and law firms continue to serve as targets for discovery requests. As evidenced by these survey results, the industry has made positive steps in recent years but has a long road ahead to achieving defensible IG programs.

ABOUT THE SURVEY

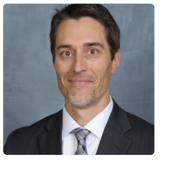
The 2024 Mattern Information Governance Survey is the first of its kind, targeting the legal market and looking deeper into the specific information governance practices within law firms.

50 law firms participated, ranging in size from 21 to 3,000 attorneys. Below is the participant breakdown by firm size:

- Firms over 100 attorneys: 16%
- Firms between 101 -250 lawyers: 18%
- Firms between 251 – 500 lawyers: 26%
- Firms between 501 – 999 lawyers: 18%
- Firms with 1,000 and over lawyers: 22%

ACKNOWLEDGMENTS

Mattern is grateful for the inputs of legal information governance professionals and peer review groups who provided their insights and time to this survey effort and helped deliver this detailed, first-of-kind data that benefits all law firms. ILTA



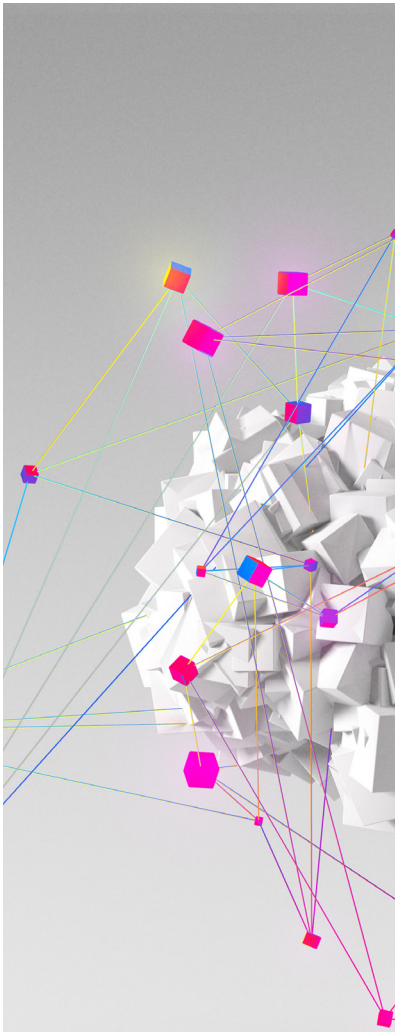
NATHAN CURTIS

a Six Sigma Yellow Belt, brings over 20 years of experience working with law firms in the U.S. and overseas in developing industry-first solutions across Information Governance, Litigation Support, Digital Imaging, and traditional Office Services. As a consultant for Mattern, Nathan is focused on emerging technologies and their application in the legal environment, driving results through Mattern's customized RFP process, and overseeing service, technology and policy implementations.

AI AND DATA-DRIVEN PROJECT MANAGEMENT

Insights for Legal Technologists

BY NILSA MORENO



The legal field generates vast amounts of data through case files, contracts, client communications, and court records. This wealth of information holds significant potential for improving legal practice. Law firms are increasingly adopting advanced technologies like AI to gain insights. A recent McKinsey report highlights this trend: "65 percent of respondents report that their organizations regularly use Gen AI, nearly double the percentage from our previous survey just ten months ago." While this statistic covers various industries, it underscores the growing importance of AI in data-intensive fields like law. Using these technologies to analyze and manage data allows legal professionals to gain valuable insights, make informed decisions, and achieve better client outcomes. Data-driven project management provides a framework to utilize data and AI in legal practice effectively.

Data-driven project management is not just a trendy buzzword but a transformative approach that empowers legal teams to leverage data to plan, execute, and monitor projects more effectively. Legal professionals can optimize resources, mitigate risks, and succeed in their projects by analyzing data, identifying trends, and predicting potential roadblocks. This approach is essential in today's legal landscape, where firms and legal departments face mounting pressure to increase efficiency, reduce costs, and increase value for their clients.

However, the data revolution in law involves more than the expansive influx of artificial intelligence; it is a fundamental shift in how legal work is done and managed. Technology is transforming the legal profession at an unprecedented pace. Cloud computing, mobile technology, and the rise of legal tech startups are changing how lawyers work, collaborate, and serve their clients. Effective data leveraging is essential to maintain a competitive edge in the modern legal field, and robust project management is key to navigating this transformation.

Law firms and legal departments can achieve better client outcomes by embracing data-driven approaches. They can use data to understand client needs better, develop more effective legal strategies, and streamline operations. They can also use data to identify new business opportunities and expand into new markets.

Legal professionals need a structured approach to project management to leverage data's power effectively. This framework guides them in planning,

executing, and monitoring projects in a data-driven way. This article explores key principles and practical applications of data-driven project management in the legal field, providing a roadmap for legal technologists to leverage AI and data analytics, including generative AI. We delve into the types of data analytics and AI transforming legal practice, examine the implementation challenges, and highlight the strategic advantages of embracing this data-centric approach. First, we will consider the widely recognized Project Management Institute (PMI) framework.

THE PMI FRAMEWORK: A FOUNDATION FOR SUCCESS

The Project Management Institute (PMI) framework provides a structured approach to project management, guiding projects through key phases: initiating, planning, executing, monitoring and controlling, and closing. This framework is particularly relevant in the age of AI, as it provides a solid foundation for managing complex projects that involve emerging technologies and data-driven decision-making. A 2023 PMI Annual Global Survey on Project Management says, "21% of respondents say they are using AI always or often in the management of projects."

The PMI framework emphasizes adaptability and flexibility, recognizing that projects rarely follow a linear path. Adaptability is critical for AI-powered tools in the legal sector, where projects may involve iterative development, continuous learning, and adaptation to new data and insights. By adopting the PMI framework, legal teams can ensure that their AI projects are well-managed, aligned with strategic goals, and deliver tangible results.

For example, consider a law firm implementing a new AI-powered ediscovery solution. The PMI framework can guide the project from the initial assessment of needs and selection of the right tool to the implementation, training, and ongoing monitoring of the solution's effectiveness. The framework can also help manage the risks associated with AI projects, such as data security, ethical considerations, and the need for ongoing human oversight.

PMI certification can be valuable for legal professionals managing AI and data-driven projects. It demonstrates their expertise in project management principles and enhances their career prospects by showcasing their ability to apply them in a complex and evolving technological landscape. Furthermore, professionals with PMI certification can play a crucial role in fostering a data-driven culture within their organizations, acting as champions for adopting AI and data analytics in legal

practice. Understanding project management fundamentals provides a solid framework for effectively wielding AI and data analytics tools.

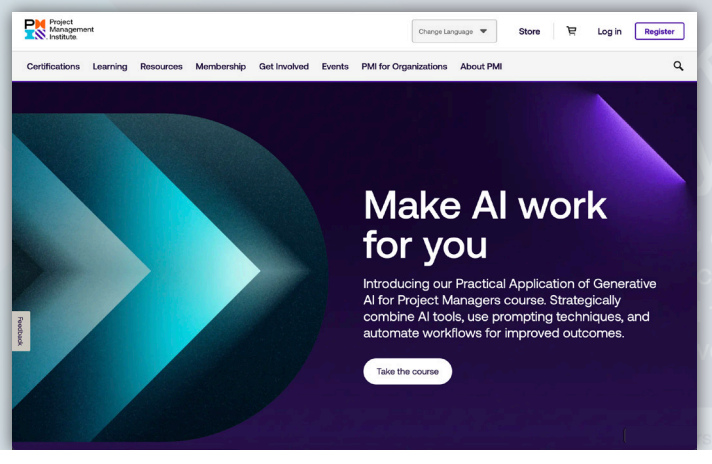
PRACTICAL APPLICATIONS: WHERE AI AND DATA MEET LEGAL PROJECTS

AI and data analytics are advancing in various legal practice areas, transforming how legal projects are managed and executed.

Think of AI tools as your legal tech toolkit. Many AI-driven tools are available, each designed to tackle different tasks and streamline your workflow. Need to sift through mountains of case law? Legal research platforms like Westlaw Edge and Lex Machina can do that quickly, providing you with relevant precedents, statutes, and secondary sources in a fraction of the time it would take to do manual research. Bogged down by contract review? Document automation tools like Contract Express and HotDocs can automate those tedious tasks, freeing up your time for more strategic work. And when it comes to ediscovery, platforms like Relativity and Everlaw use AI to make sense of those massive datasets, identifying key documents and evidence quickly and efficiently. Even your practice management software is getting in on the action, with solutions like Clio and MyCase incorporating AI features to automate tasks like time tracking, billing, and client communication, helping you manage your practice more effectively.

These AI tools are data-hungry! They thrive on all sorts of information to do their magic. Think of it like this: they need the fuel of case data (court filings, pleadings, transcripts), client data (contact info, case details, billing records), financial data (budgets, expenses, revenue), and project management data (timelines, tasks, resources) to power their insights and predictions.

Where does AI fit into your data-driven project management workflow? Well, pretty much everywhere! From the initial case assessment and planning stage, where AI can help you analyze data and predict outcomes, to the nitty-gritty of legal research and document drafting, AI tools can streamline your process and improve accuracy. Even during trial preparation and litigation support, AI can help you analyze



MORE ONLINE

Find more project managment resources online at pmi.org.

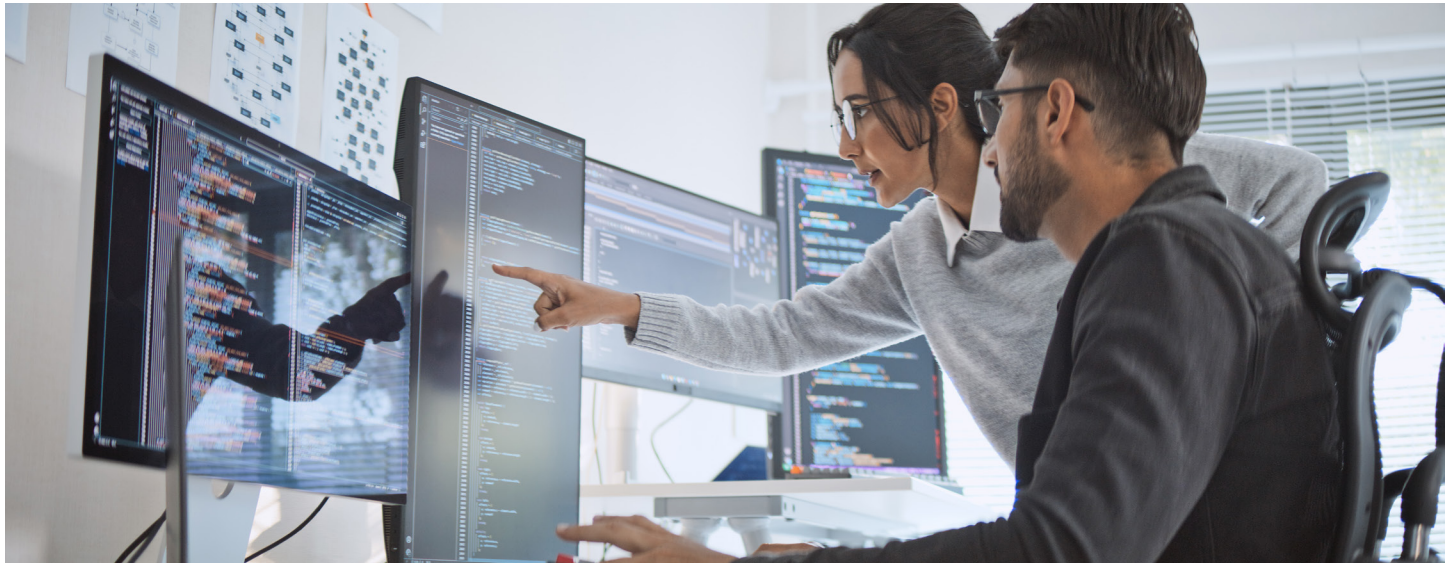
evidence and develop winning arguments. And regarding the less glamorous side of things—billing and time management—AI can automate those tasks and free up your time for more strategic work.

CHALLENGES AND CONSIDERATIONS: THE REALITIES OF THE AI LANDSCAPE

While the potential benefits of AI in data-driven project management are significant, it's essential to acknowledge the challenges and ethical considerations that come with it.

Of course, it is not all smooth sailing. Implementing AI in data-driven project management comes with its own set of challenges. Data security and privacy are paramount, so you must ensure those AI tools handle confidential client information responsibly. This means ensuring compliance with ethical guidelines and data privacy regulations, such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). And these tools can be expensive! You will need to factor in software, hardware, and training costs. Integrating these new tools with your existing systems can be a headache, requiring careful planning and technical expertise.

Beyond the practical challenges, ethical considerations are paramount when implementing AI in legal practice. AI algorithms can perpetuate biases in the data they are trained on, potentially leading to unfair or inaccurate outcomes. For instance, an AI tool trained on a dataset representing a particular legal perspective will likely



provide biased results, overlooking relevant precedents or arguments. Legal professionals must ensure they use AI tools with fairness and transparency. This requires careful selection of training data, ongoing monitoring of AI systems for bias, and a commitment to addressing potential biases. Legal project management professionals can help ensure that AI is used responsibly and objectively by actively engaging in ethical discussions and implementing safeguards.

Similarly, compliance monitoring in the age of AI presents its own set of unique considerations. Imagine a complex legal case involving a multinational corporation. AI tools can automate the arduous task of tracking regulatory changes across multiple jurisdictions and identifying potential risks, but these tools must be carefully implemented and monitored. If the AI system flags a routine international transaction as suspicious based on incomplete or biased data, it could trigger an unnecessary investigation, delaying the case and harming the client's business operations. Legal professionals must ensure these tools adhere to strict ethical guidelines and data privacy regulations. It's crucial to be mindful of the potential for AI to perpetuate biases or make inaccurate predictions, especially in compliance where the stakes are high.

STRATEGIC ADVANTAGES: REAPING THE REWARDS OF DATA-DRIVEN PROJECT MANAGEMENT

While addressing ethical and logistical challenges is crucial, it is equally important to recognize the tangible benefits that data-driven project management can bring to your legal practice. Below are examples of ways leveraging AI and data analytics can translate into real-world advantages, including increased efficiency, reduced risks, and improved client relationships.

- **Enhanced efficiency and productivity:** Streamlining workflows and optimizing resources can significantly reduce the time and effort required to complete legal projects. By automating tasks like document review and legal research, AI frees lawyers to focus on higher-value work, such as client interaction, strategy development, and negotiation.

- **Improved accuracy and reduced risk:** AI can help minimize errors and inconsistencies in legal documents and processes, reducing the risk of costly mistakes. For example, AI-powered contract review tools can identify potential loopholes or inconsistencies that human

reviewers might miss. This can help prevent legal disputes and ensure that contracts are airtight.

- **Increased client satisfaction:** By delivering faster, more accurate, and more cost-effective legal services, law firms and legal departments can improve client satisfaction and build stronger relationships. AI can also help personalize client communication and provide more tailored legal advice. For instance, AI-powered chatbots can answer client questions, provide updates on case progress, and even schedule appointments.
- **Gaining a competitive edge:** Adopting AI and data-driven project management in today's competitive legal market can give law firms and legal departments a significant advantage. By leveraging these technologies, they can offer more innovative and efficient services, attract and retain top talent, and position themselves as leaders in the field.

With a clearer picture of the strategic advantages, we can explore how to cultivate a data-driven culture within your legal team.

BUILDING A DATA-DRIVEN CULTURE: EQUIPPING LEGAL TEAMS

So, how can legal professionals thrive in this new data-driven world? We need to develop new skills in data analysis and technology, and we must remember that PMI certification can be an asset in this journey, helping us stay informed about the latest advancements while prioritizing judicious considerations in all aspects of our work.

The future of law is data-driven, and those who embrace this change will be the ones who lead the way. Change leadership requires a shift in mindset, not just for individual lawyers but for entire legal teams. This change fosters a culture that values data, openly shares insights, and makes decisions from evidence and insights. This cultural shift requires leadership from the top, investment in training and development, and a commitment to collaboration between legal and technology teams.

To successfully implement a data-driven project management culture, law firms and legal departments can take several steps:

- **Provide training on data analysis and AI tools:** Equip legal professionals with the skills, certifications, and knowledge they need to understand and utilize data effectively.
- **Encourage data sharing and collaboration:** Create platforms and processes for sharing data and insights across different teams and departments.

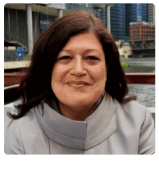
- **Celebrate data-driven successes:** Recognize and reward individuals and teams who use data to achieve positive outcomes.

- **Foster a culture of experimentation and innovation:** Encourage legal professionals to explore new ways of using data and AI to improve their work.

To truly leverage the power of AI and data in legal project management, law firms, and legal departments need to foster a data-driven culture. This shift involves cultivating a mindset that values data analysis and evidence-based decision-making, encouraging collaboration between legal and technology teams, and investing in training and development to equip legal professionals with the necessary skills to thrive in a data-centric environment. By embracing these principles, legal teams can create a foundation for success in the evolving legal landscape, where data is not just an asset but a strategic imperative.

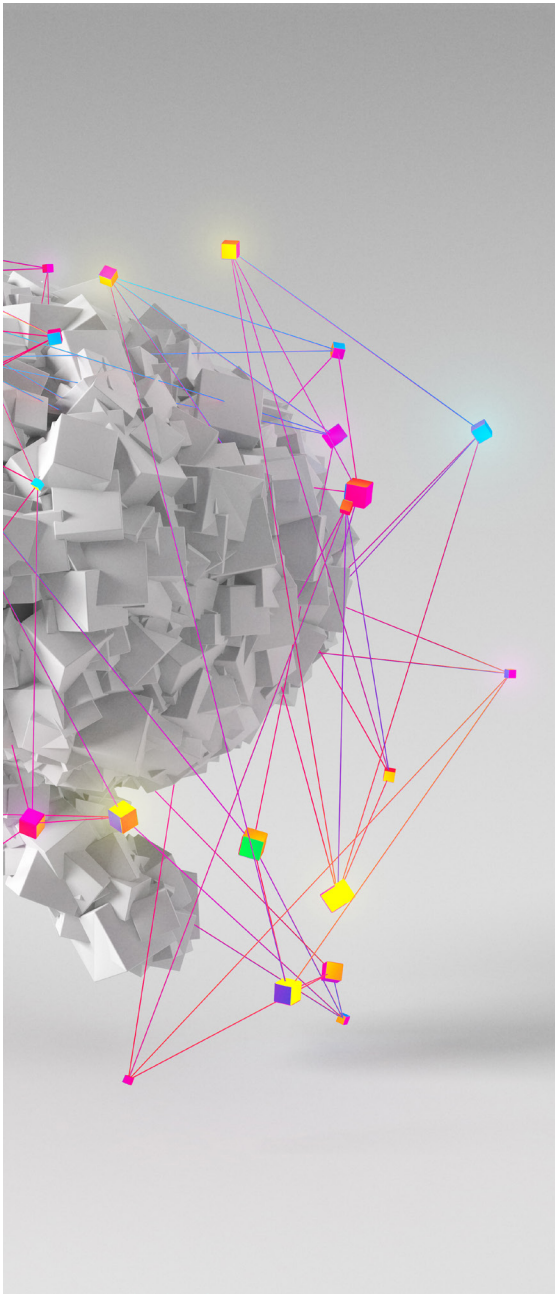
CONCLUSION: THE FUTURE OF LAW IS DATA-DRIVEN

The legal profession stands at the intersection of tradition and innovation, and data is the bridge connecting the two. AI and data analytics are no longer futuristic concepts but essential tools reshaping how legal professionals work. By embracing data-driven project management, legal teams can realize unprecedented efficiency, accuracy, and strategic advantage. The future of law is data-driven, a future where technology enables legal professionals to make more informed decisions, provide better service to clients, and achieve better outcomes. To prepare for this future, legal technologists must be proactive in embracing AI and data-driven approaches, ensuring that AI is used to enhance the practice of law and promote justice. **ILTA**



NILSA MORENO

is a legal technology leader with over 20 years of experience spanning law, technology, information governance, legal operations, and project management. Nilsa has a proven track record of success in providing comprehensive Information Governance, including ediscovery and litigation technology management solutions and strategic services for AM Law 200 clients. Her areas of expertise include consulting, implementation, training, forensics, and best practices in information governance and data management. Nilsa is the founder of CollaborTech, LLC.



FORENSIC DATA COLLECTION

A Bridge Between Digital Forensics, Ediscovery, and Artificial Intelligence

BY THOMAS YOHANNAN

In the summer of 1956, when John McCarthy gathered researchers at Dartmouth College, he did not just coin the term "artificial intelligence" – he sparked a digital revolution that would span generations. While the world is enamored with headlines of the changing face of society due to Large Language Models (LLMs), artificial intelligence (AI) extends into nearly every corner of innovation: self-driving vehicles navigate our streets, computer vision systems diagnose diseases, and neural networks unlock patterns in vast seas of data. Yet beneath the complexity of these systems lies a fundamental truth: the quality of AI is only as reliable as its foundational data. The emphasis on data quality is why data collection from a legal and investigative perspective is crucial for moving forward.

In its core definition, digital forensics deals with recovering and investigating data residing in digital devices and cloud-native storage, generally in the context of cybercrime. Similarly, in the context of the EDRM model, ediscovery manages data as evidence from initial collection to presentation for use in both civil and criminal legal cases. Yet, for both disciplines, the point of origination is through data collection, which is this crucial process of gathering digital evidence or information that is accurate and legally allowable.

Like digital forensics and e-discovery, AI's effectiveness hinges on forensically sound data collection. Why? Forensically sound data is information collected, preserved, and handled in a way that maintains its integrity and authenticity so it can be reliably used as evidence or for analysis. For AI, gathering and validating data isn't just a 'best practice,' but it's essential for building trustworthy AI systems that can identify patterns, form patterns, and produce insights. Illumination of the need for forensically sound data as the foundation of AI occurs when considering the fact that when LLMs memorize errors and biases and create incomplete analyses, there is an audit trail to see where such misgivings originated. The critical role of forensically sound data collection and its parallels with established practices in digital forensics and ediscovery must be examined to seize the opportunity of AI technologies.

THE DATA COLLECTION CHALLENGE

According to research by AI Multiple Research, training data collection has been identified as one of the main barriers to AI adoption. Their analysis highlights six significant data collection challenges: availability issues, bias problems, quality concerns, protection and legal requirements, cost constraints, and data drift prevention. Three challenges—quality concerns, protection and legal requirements, and bias problems—can be effectively addressed through forensic data collection methods because forensically sound data is collected to ensure integrity using legally prescribed standards. When talking about clean data in the context of AI, we generally mean valid, consistent, and uncorrupted data. The large volume, complexity, and rapid data evolution of data within an organization make the task difficult. However, these challenges present an opportunity to leverage established forensic methodologies to ensure data quality.

THE ROLE OF FORENSICALLY SOUND DATA COLLECTION

Artificial intelligence begins with data collection. Every technology starts with data collection. Data collection is not just the first step in the decision-making process; it is the driver of machine learning. The integrity and reliability of AI systems hinge on acquiring meaningful information to build a consistent and complete

dataset for a specific business purpose. This particular purpose can include decision-making, answering research questions, or strategic planning. It's the first and essential stage of data-related activities and projects.

Yet, the integrity and reliability of AI systems depend entirely on data that remains untouched and unaltered from its original state (i.e., forensically sound data). A few critical aspects must be in place when gathering training data for AI, similar to digital forensics.

CRITICAL ASPECTS OF FORENSIC DATA INTEGRITY

Chain of Custody: Tracks every interaction with the data through detailed chronological records of collection, storage, and access, including timestamps and user details for complete accountability.

Cryptographic Hashing: Generates unique digital fingerprints of data files, enabling immediate detection of any modifications or tampering through hash value verification.

Data Acquisition Methods: Utilizes specialized forensic tools to capture data while preserving original file structures and metadata, ensuring authenticity from the point of collection.

Documentation: Maintains transparent records of collection processes, methodologies, transformations, and limitations, establishing clear data provenance.

Metadata Preservation: Retains all contextual information about data sources, providing crucial context for forensic investigations.

Additionally, just as traditional digital forensics requires meticulous documentation and validated tools, organizations using AI need strict protocols to preserve training data, model parameters, and system logs in their original form. This forensic approach to data handling does more than just feed algorithms—it creates an auditable trail that proves your system's decisions are based on reliable, untampered information, building trust and meeting compliance standards.

"For many companies, building a forensically sound data approach feels overwhelming," notes Christian J. Ward, Chief Data Officer of Yext, a corporate knowledge graph and search company. "Here's the reality: your already structured data can integrate seamlessly with AI solutions. Whether custom or off-the-shelf, today's AI models have massive training datasets beyond any single organization. You can merge this AI with forensically

sound data structures through RAG solutions or similar protocols – combining broad language understanding with verified, trusted information. This isn't just about feeding data to machines. It's about ensuring every AI response draws from forensically verified knowledge."

Forensic data collection in AI serves several critical functions. First, it ensures data integrity by implementing strict protocols for gathering and preserving training datasets, similar to evidence handling in criminal investigations. This process includes maintaining detailed documentation of data sources, collection methods, and preprocessing steps. For instance, when collecting employee emails from a corporate server using Rocket, each email is preserved with its complete metadata, including sender, timestamp, and routing information, creating exact copies. It also includes detailed documentation of data sources (whether emails came from Exchange servers or local backups), collection methods (whether extracted using Rocket or Outlook exports), and preprocessing steps (how emails were filtered and redacted). For AI systems, this forensic approach helps track potential biases, data quality issues, or manipulations that could affect model behavior.

The rigorous protocols extend beyond data collection – they encompass recording model parameters, system logs, and decision-making processes to ensure data remains valid and uncorrupted throughout its lifecycle. For example, when an AI system analyzes employee behavior patterns for security threats, forensic documentation would allow investigators to trace the exact sequence of events, from the initial log files captured through the AI's analysis steps to the final alert generation. This level of detail becomes crucial for auditing AI behavior for accuracy and verifying that the underlying data hasn't been tampered with or degraded. By maintaining this detailed chain of custody for data and model decisions, organizations can demonstrate compliance with AI regulations while building trust through transparency – much like how a bank must prove its transaction records are authentic and unaltered for regulatory audits.

Just as a high-performance engine requires clean fuel to run efficiently, AI systems need pristine data to produce reliable outcomes.

BRIDGING TO ARTIFICIAL INTELLIGENCE

Data is the fuel that powers artificial intelligence and machine learning systems. If AI works with premium and structured data, it creates more meaningful and accurate insights. Forensically sound data collection becomes crucial when looking for meaningful and accurate insights. Just as a high-performance engine requires clean fuel to run efficiently, AI systems need pristine data to produce reliable outcomes. When organizations feed their AI models with forensically sound data collected through rigorous digital forensics and ediscovery processes, they create a foundation for success. However, using poor-quality data is like putting cheap fuel in your engine, leading to unreliable performance and questionable results.

As Zach Warren, Technology & Innovation Insights, Thomson Reuters Institute notes, "The idea of 'garbage in, garbage out' might be something that every lawyer has heard at this point, but being repeated so often doesn't make it any less true. In fact, the availability of Gen AI may make this maxim even more pressing: If law firm leaders see technology as a key firm differentiator in the near future, that makes clean data to run these tools not just a nice-to-have tech issue, but a key business problem that has to be solved."

With the surge of digital transformation, organizations may need to establish a solid data foundation before implementing AI. Jumping

to the conclusion of the process, AI activation, without ensuring their data meets the necessary quality standards, will only harm the usage of transformational technologies.

All successful companies do it: constantly collect data. Data holds exceptional importance in fueling AI, as its strength lies in analyzing large amounts of data and making predictions based on its inputs. Data accuracy directly correlates with AI's ability to be intelligent. The data truly is the differentiator. Organizations must realize that foundational data is the first and most crucial step in creating accurate artificial intelligence, not jumping straight to activation. Organizations must prioritize accurate data from the start to maximize AI model performance.

INCREASING AI INTELLIGENCE USING FORENSIC AND EDISCOVERY DATA

Building on this foundation of clean, forensically sound data, organizations can leverage digital forensics and ediscovery principles to provide a rich training ground for AI algorithms. "Generative AI in ediscovery isn't just a tool; it's a force multiplier. Picture this: mountains of data that would take human teams months to review, tackled in hours. And it doesn't stop there—this tech learns and evolves, anticipating needs and uncovering connections you didn't even know to look for. It's not replacing humans; it's unleashing their potential by cutting through the noise and delivering actionable insights faster than you can say data overload," says Cat Casey, Chief Growth Officer, Reveal.

Digital forensics and ediscovery data can offer a rich training ground for AI algorithms. For example, the AI can be presented with recurring incident patterns of cybercrime to predict or identify various occurrences of cybercrime to further assist in their cybersecurity measures. Similarly, AI will use information from an ediscovery process to automate and improve identifying relevant documents in legal cases, saving time and costs.

HOW TO CREATE AI-READY FORENSIC DATA

Creating AI-ready forensic data requires four essential pillars that ensure effective utilization in artificial intelligence and machine learning applications:

- **Data Quality:** The foundation of reliable AI systems demands accurate, complete, and consistent data. This fundamental requirement ensures trustworthy model outputs and dependable results.

- **Governance:** In today's regulatory landscape, data must be trusted, consented adequately to, and fully auditable to maintain compliance with privacy regulations and AI guidelines while protecting organizational interests.

- **Understandability:** Data becomes more valuable when enriched with contextual intelligence, comprehensive metadata, and accurate labels, enabling AI systems to interpret and utilize the information better.

- **Availability:** Ensuring the correct data is accessible at the right time through robust interoperability and real-time delivery capabilities is crucial for practical AI training and activation.

These pillars work together to create a framework that enables organizations to build reliable AI systems while maintaining forensic data integrity.

CHALLENGES AND CONSIDERATIONS

Data collection enhances AI, but the opposite is true - AI enhances data collection efficiency. An AI feedback loop is where AI can further add value by optimizing the processes of collecting data in and of itself. A prime example is predictive coding in ediscovery, where an AI-driven process streamlines document review by prioritizing the most relevant data, creating a more efficient collection process. However, while this convergence of digital forensics, ediscovery, and AI presents opportunities, several critical considerations demand attention.

The success of AI implementations hinges entirely on data quality. As industry experts emphasize, AI models follow the principle of "garbage in, garbage out" without exception. This reality makes the creation of forensically sound AI datasets particularly challenging in three key areas:

- **Accurate Data:** AI's foundational element is ensuring data is solid, correct, and represents what is trying to be studied. It's about being thorough and meticulous in how data is collected and verified.

- **Playing by the Rules:** With all the privacy laws and regulations out there, organizations are expected to adhere more and more to data requirements and legal frameworks. It is critical to balance using valid data and respecting people's privacy.

- **Keeping Secrets Safe:** Protecting sensitive information while maintaining valuable data for AI training is a top priority. Think of it as redacting a document - you want to hide the sensitive bits while keeping the vital context intact.

CONCLUSION

The most fundamental challenge underlying digital forensics, ediscovery, and AI is the issue of data collection. Moving forward, centralizing data architectures of various technology landscapes on forensically sound data collection will lead to an ease in innovation. Making data compliant and secure while attaching to it the principles of integrity and accountability that are the mainstays of digital forensics and ediscovery should be the norm when thinking about the changing landscape of artificial intelligence. **ILTA**

THOMAS YOHANNAN

is Co-Founder of Digital DNA, creators of Rocket - the industry's first cloud-native remote forensic collection platform for Windows & MacOS that operates without installed software, hardware, or on-site personnel. As an attorney merging legal expertise with technical acumen, he specializes in security, data forensics, and cyberinsurance. Thomas excels at bringing innovative solutions to market through strategic analysis of risk and regulatory frameworks in high-touch verticals. His multidisciplinary approach helps enterprises navigate complex digital challenges. thomas@digitaldnagroup.com