Moneyball for Lawyers

How Legal Analytics® are Transforming the Practice of Law

By Owen Byrd, Chief Evangelist & General Counsel, Lex Machina

Imagine you could make a data-driven prediction about how opposing counsel, or a judge, or a party to litigation or a transaction, will behave. What if you could anticipate the results a specific legal strategy or argument will produce? Would you continue to rely exclusively on traditional legal research and reasoning to inform the advice you give clients, the documents you draft, the negotiations you conduct and the arguments you make? Or would you integrate Legal Analytics into your lawyering by practicing Moneyball for lawyers?

The Origins of *Moneyball*

Since the 2003 publication of *Moneyball: The Art of Winning an Unfair Game,* by Michael Lewis, and the 2011 release of a popular film based on the book, "Moneyball" has become shorthand for making data-driven decisions in domains far removed from baseball.

The Moneyball story chronicles how the 2002 Oakland A's built a cost-effective, winning team by analyzing new measures of player performance, such as on-base and slugging percentages. This data more accurately predicted a player's offensive value than did other, more traditional baseball stats. It also provided an objective way to supplement - and sometimes overrule - the subjective opinions of traditionalist scouts, coaches and managers. In the ensuing decade, the Moneyball approach has come to dominate professional baseball.

Moneyball for Lawyers

So what does Moneyball have to do with law? Legal Analytics involves the use of data to predict outcomes and achieve results - in litigation and transactions. It also applies to business development efforts by lawyers and law firms to attract and retain clients, as well as the selection and management of those lawyers and law firms by companies and their in-house attorneys.

The Attorney-Client Mating Game

In order for an attorney to apply data-driven decision-making to a dispute or transaction, that attorney must first have a client. Moneyball Legal Analytics change the two-way dynamics that govern both business development by law firms and outside counsel selection and management by companies.

Performance data about lawyers enables law firms to assemble the most compelling pitch to a new client, or to an existing client about a new matter. Firms can quantify specific experience with the subject matter of the dispute or transaction, as well as with opposing parties and counsel. For new litigation, they can also quantify their prior experience with the judge. When a law firm knows that it is competing with a specific other law firm over a new matter, it can use data to compare its relevant experience and available bandwidth.

Litigation

Legal Analytics can influence every step of the dispute resolution process. The threshold questions begin before a demand letter is ever drafted, sent or received.

Lex Machina

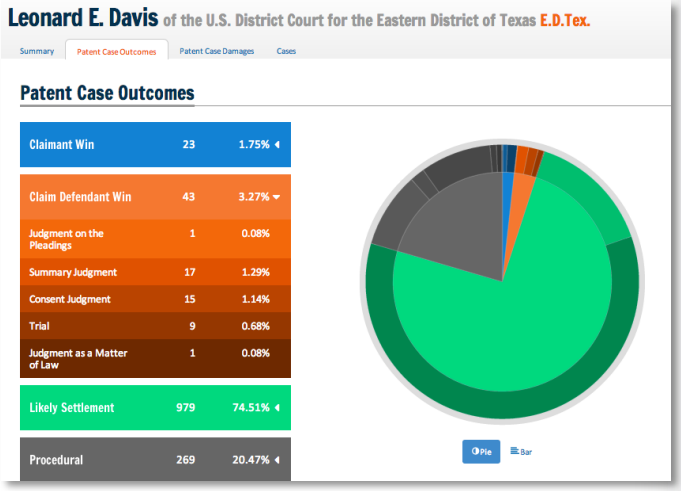
Plaintiffs and their attorneys must analyze not only who has caused the harm and who has the deepest pockets, but also how prospective defendants will react to a claim.

One obvious measure is prior litigation behavior of a defendant. But what if the defendant has never before been sued, at least about the subject matter of the plaintiff's claim? What about the behavior of similar defendants faced with similar claims? For example, a patent holder asserting a claim of infringement about a patent that has never before been litigated, against a party that has never before been sued for infringement, can extract predictive insights out of the litigation behavior of similar parties in similar patent lawsuits.

Similarly, defendants and their attorneys can use Legal Analytics to inform their response to a demand letter or complaint. Has the plaintiff brought other similar lawsuits? If so, how did the plaintiff behave at each stage of litigation? Was the plaintiff successful? How have other defendants responded to the plaintiff, or to similar plaintiffs with similar claims or behaviors?

Consider also the choice of venue by plaintiffs and efforts by defendants to transfer venue. For plaintiffs, establishing jurisdiction by a court with both a measurable track-record favoring plaintiffs with similar claims, as well as one that moves faster and more often to a jury trial, can have more impact on the outcome of litigation than any subsequent tactics. For a defendant moving for transfer of venue, quantifying the arguments that have succeeded in the past with a specific judge can propel the case out of an unfriendly environment and into one where similar defendants more frequently prevail.

Once litigation is under way, every step in the process can be informed by Legal Analytics that improve a party's chances of winning, while at the same time minimizing unnecessary legal spending. For example, in-house attorneys for a large pharmaceutical company had heard anecdotally that the judge presiding over their patent case often ruled on claim construction solely on the briefs, without holding a hearing. Before the company designed and executed on its claim construction strategy, it obtained data that revealed that the judge did rule on claim construction without a hearing over 80 percent of the time. The company then knew that it had to include all of its arguments in its brief, holding nothing back for a hearing that was likely never to happen.



Such granular data about judges can significantly impact litigation strategy. Data about average time to termination and time to trial, for example, can help a party budget its funds for a case and a lawyer budget her time.

Data about the behavior of peer companies in similar litigation can confirm or cause adjustments to a litigant's behavior. One large technology company known for its aggressive responses to patent lawsuits commissioned a comparative study of its patent litigation behavior and the behavior of 15 of its peers.

Lex Machina

Subjects studied included case volumes, case types, products at issue, venue, settlement volumes, case phase at settlement, settlement rates, chances of early settlement, number of cases stayed pending re-examination of the patent at issue by the U.S. Patent and Trademark Office, overall case outcomes, win/loss rates and damages awards.

The results showed a wide variety of behaviors and outcomes, even in response to similar claims by similar plaintiffs. Some companies settled early. Others fought every claim to the bitter end. The data revealed an optimized combination of spending and tactics that would have been impossible to know and deploy without a Moneyball approach.

For common motions, such as motions to transfer, stay or compel production, or to provide an injunction or protective order, or even for summary judgment, Moneyball Legal Analytics can include assembling and using statistics on win/loss rates for parties and counsel, as well as statistics on win/loss rates and time to disposition for judges.

Charles Kramer Verhoeven

Summary | Cases | Client List | Judge Appearances

Judge	District	For Plaintiff	For Defendant	Total Cases
James Rodney Gilstrap	E.D.Tex.	2	12	14
David Folsom	E.D.Tex.	1	7	8
Carol E. Jackson	E.D.Mo.	4	2	6
Leonard E. Davis	E.D.Tex.	0	5	5
Claudia Ann Wilken	N.D.Cal.	2	3	5
Susan Yvonne Illston	N.D.Cal.	2	3	5
Richard G. Seeborg	N.D.Cal.	0	4	4

Finally, zealous representation in litigation can now be informed by statistical insights into the caseloads and behaviors of opposing counsel. If you knew about all the cases that your opposing counsel is currently handling, and you mapped out all the upcoming filing deadlines in those cases, would you make sure to serve your discovery requests right when opposing counsel is busiest with other deadlines? It may be hardball, but Moneyball developed to enable big league teams, not Little Leaguers, to win.

Transactions

Litigation data can also enable data-driven decisions about transactions, such as purchase, sale and license agreements. For example, a lawyer advising a client about whether or not to grant a patent license to a prospective licensee can investigate:

Prior lawsuits:

- By the licensee for a declaratory judgment about IP rights?
- By the licensee against a licensor about an IP license?
- Against the licensee by a licensor for exceeding the scope of a license?

Related patents:

- Have they been litigated?
- Has that litigation included any claims of infringement by licensees?

Subject matter:

- Has the subject matter of the IP generated litigation?

Due diligence investigation performed prior to a merger or acquisition represents another type of transactional lawyering that is well suited for the application of Legal Analytics.

Lex Machina

Advanced Subject Matter Analytics Example: Patent Similarity Engine

Legal Analytics Moneyball can be applied to analyze the subject of a transaction or dispute, in order to maximize the value of an asset and the income it generates, while at the same time minimizing risk and expense.

For example, Lex Machina's Patent Similarity Engine uses hundreds of factors to compare a specific portfolio of patents with over 37,000 patents previously asserted in federal litigation since 2000. It produces a ranked list of asserted patents that are most conceptually similar to the patents in the portfolio. Lex Machina then performs additional expert analysis to develop specific recommendations about how best to monetize the portfolio.

The Patent Similarity Engine combines output from three different analyses - textual similarity, latent semantic analysis and the litigation graph - to produce a specific similarity score for each relevant asserted patent.

1. Textual Similarity - This calculation measures the overlap between specific elements of each patent in the portfolio and all the patents asserted in litigation.

The Patent Similarity Engine constructs a mega-patent by concatenating the textual fields for all portfolio patents. It then generates queries that retrieve similar asserted patents, with specific terms in the query boosted based on their frequency in the mega-patent. Each field - patent title, abstract, description and claims - is treated as a separate strategy. A combined strategy also concatenates text from all other fields.

2. Latent Semantic Analysis - This analysis converts the portfolio and each asserted patent into 300-dimension vectors, which are compared with cosine similarity.

This analysis uses a model trained on over 500,000 randomly selected patents granted from 2005-2012. Certain terms receive greater weight based on their global significance. Re-ranking is used so that the analysis is only applied to the top 2,000 patents retrieved using textual similarity and litigation graph analysis.

3. Litigation Graph - This strategy awards similarity weight to patents asserted in cases in which the portfolio patents were also asserted, as well as to patents asserted by parties which are part of a common larger entity.

Conclusion

We are not far off from the day when stats for the players and teams in the legal game will be as omnipresent and impactful as stats for baseball players and teams. Lawyers and law firms, parties and their law departments and districts and judges will all be subject to statistical analysis and performance rankings. All these players will, in turn, use this data to compete more vigorously with each other.

Moneyball for lawyers is bringing objective rigor to the traditionally subjective practice and business of law. The game will become more economically efficient and transparent, delivering better results for those most in need of legal services. While some traditionalist lawyers may resist the application of Legal Analytics to law, those who embrace it, especially today's early adopters, are likely to gain significant and lasting competitive advantage. While they may not win the Commissioner's Trophy, they will win the client, transaction or case – the ultimate goals of every lawyer.

This article first appeared in *The Bottom Line*, April 2013, the official publication of the State Bar of California

Lex Machina