

# Data Science for the Legal Profession



**With the increase of data that is produced and with no sign of slowing down any time soon, coupled with advancements in machine hardware the legal profession is set for a major data driven overhaul in the coming years.**

**Disclaimer**

The material in this document is intended for information purposes only and should not be taken as a commitment to deliver any product or service, and may not be incorporated into any contract.

# Contents

Summary .....	3
Data Science and the Legal Profession .....	4
Big Data .....	5
Machine Learning.....	6
Predictive Analytics.....	7
Emerging Applications of Data Science for the Legal Profession. ....	8
Case Assessment and Prediction .....	8
Discovery.....	9
Judge Analysis .....	10
Opposition Scrutiny.....	11
Client and Witness Insights.....	12
Case Strategy Development.....	13
Conclusion.....	14

## Summary

This white paper presents an overview of how Data Science related technologies are growing in the legal profession. The paper also highlights where Quanovo believes the main areas are that legal companies can use these technologies to their advantage if they are not already doing so. Finally, in the conclusion of

this document we discuss where these technologies may take the legal profession in the future and the importance of embracing the technology early to gain competitive advantage.

# Data Science and the Legal Profession

Traditionally the legal profession has a fairly conservative outlook when it comes to adopting new technology and integrating it successfully into their businesses. However, the last decade has seen a major shift in the way data is produced and more importantly how it is accessible. Big Data has become more powerful as a result and has been utilised across most industries, including government and healthcare, in order to achieve desired outcomes. Legal professionals have been using hidden facts found in court documents such as witness statements, judge's summaries and court logs to help them win arguments but, like those other industries, the legal system is creating an ever-increasing amount of data. That is where Data Science and technologies such as Machine Learning and Big Data analytics can help. They are able to deal with the volume of data that is produced and return actionable results in real time meaning that legal professionals are able to concentrate more on task that require a human touch rather than the grind of fact gathering and analysis.

Case information is becoming more and more accessible with online resources storing huge amounts of digital documents that in many cases can be accessed for free. The delivery of

such systems is greater in the USA but the UK and the rest of Europe are not far behind with traditional systems being overhauled and government initiatives being put in place to aid with the provision of open data. All of this provides a solid platform for Data Science and its related technologies to really take off in the legal profession in the coming years.

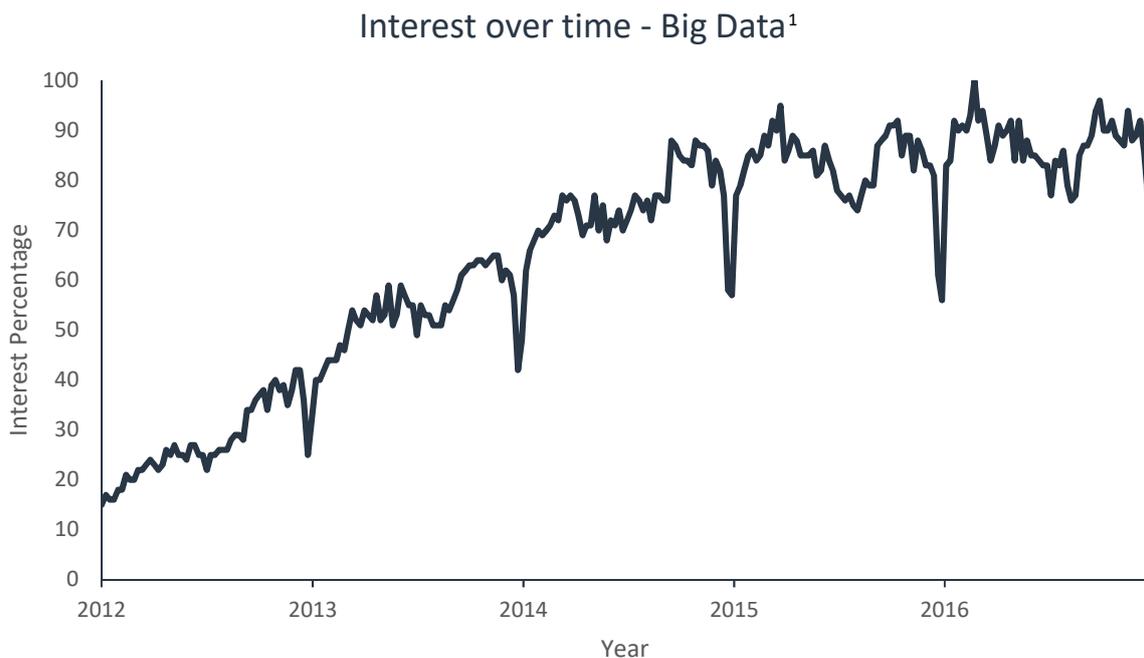
Legal profession clients are no longer willing to pay the huge hourly fees that they were used to a decade ago. A large part of that hourly fee is made up from the teams of associates and paralegals that work on each case. With the help of digital solutions law firms are able to cut down on their overheads per case by employing less people to do the same amount of work through a symbiotic working relationship with machines. Not only does this solution help keep costs down but in many circumstances, it can actually mean finding the right information a lot faster than previously possible.

In this document, we will explore some of the ways that Data Science can be utilised effectively in the new data powered landscape that is emerging in the legal profession.

# Big Data

Big Data is not a new phenomenon, in fact the notion has been around for almost half a century. However, it is a topic that has been growing in exposure across every industry in recent years. There are two main reasons that, although the idea of Big Data has been around for some time, it is now gaining traction and rapidly:

- The first of these two reasons are that globally we are producing greater quantities of data than ever before meaning that how we store and process that data to make constructive use of it has become a key issue.
- The second reason is that the advancement of hardware solutions has reached the point where they are capable of running the underlying algorithms on such huge data sets. Meaning that no longer is the field of Big Data confined to a corner of academia or indeed gigantic organisations as the hardware required has become more affordable and practical to obtain and use.



<sup>1</sup> Data from Google Trends - <https://www.google.com/trends/explore?q=big%20data>

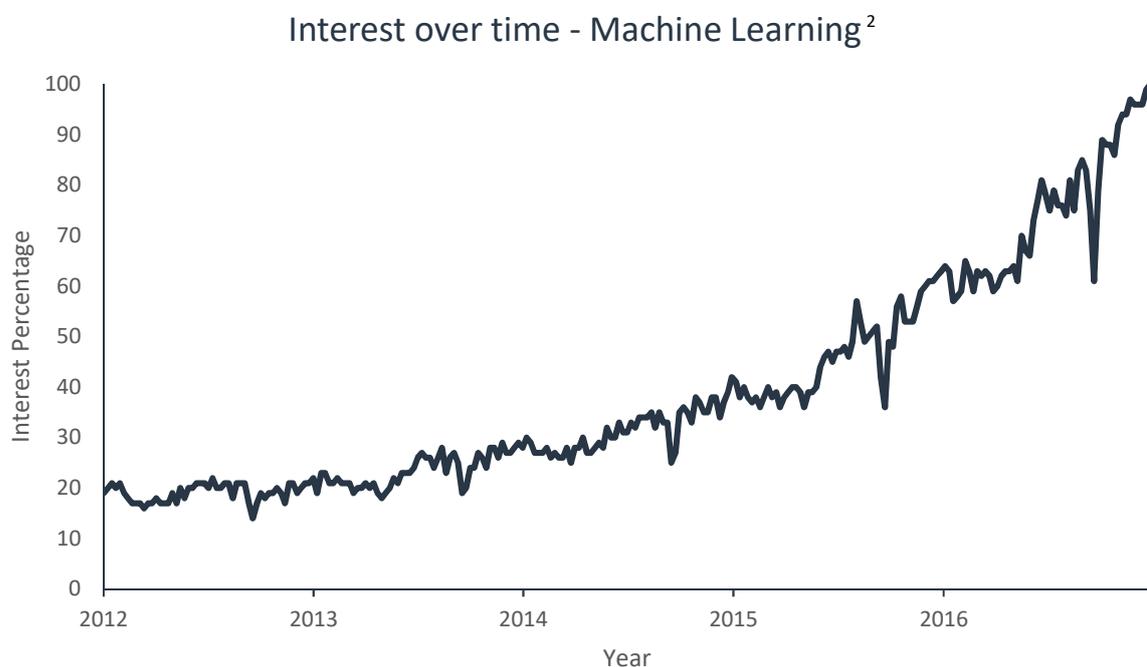
# Machine Learning

Machine Learning is similar to the process of Data Mining and in many ways, can be considered a subset of the field. Both share the concept of looking through data with the goal of identifying patterns and establishing connections.

The algorithms that Machine Learning is built upon can be very useful for extracting qualitative data from datasets that are simply too large for people to sort through. These algorithms are generally split into two main categories, supervised and unsupervised. Supervised Machine Learning algorithms take what has already been learned and apply that knowledge to new data, whereas Unsupervised Machine Learning is the process of sifting through data with the aim of discovering some, as yet unknown meaning, from it.

In the real world, what this means is that we can utilise Machine Learning technology to carry out tasks such as extract sentiment from bodies of text like product reviews and use that information to derive the reasons for negative or positive feedback.

The real power of Machine Learning comes from the fact that it is a technology that can enable the prediction of future events even when certain datasets are absent.

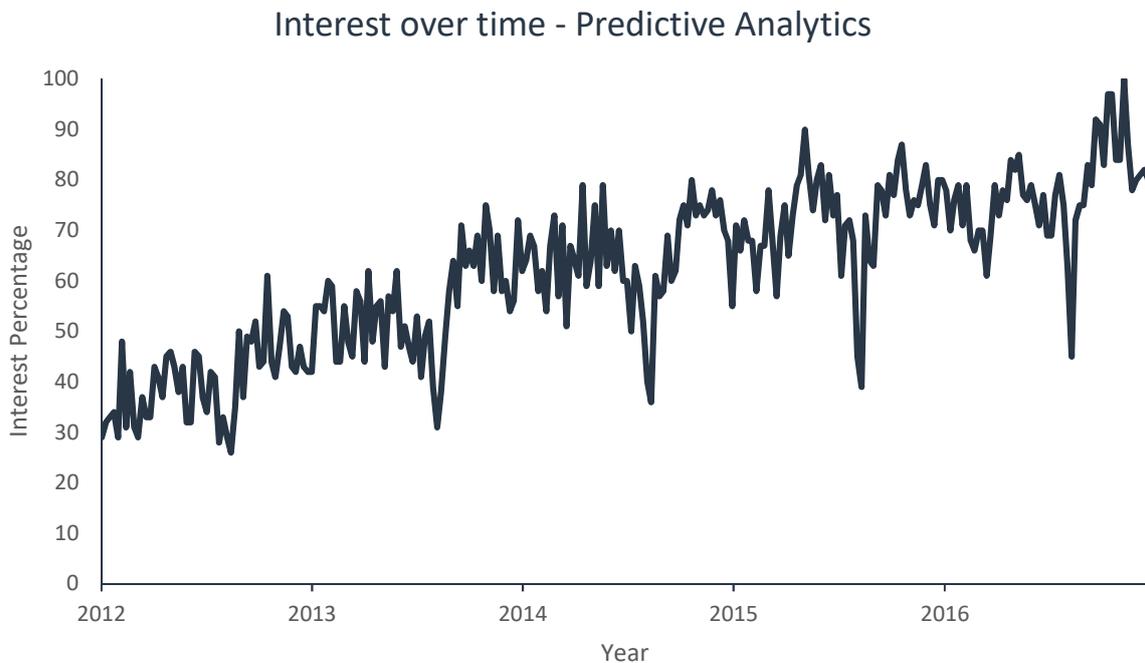


<sup>2</sup> Data from Google Trends - <https://trends.google.com/trends/explore?q=machine%20learning>

## Predictive Analytics

When applied in the realm of Machine Learning the term Predictive Analytics refers to the ability to predict future outcomes using trends and indications that can be derived from the underlying data. By predicting what is likely to occur in the future, Predictive Analytics provide the opportunity for an organisation to gain advantage surrounding any action that should be taken and in doing so transitions that business from reactive to proactive.

The direct applications of Predictive Analytics to the legal profession may take the form of answering questions such as how a specific judge is likely to react to particular legal argument? Or what are the current trends for win/loss rates for a specific type of case? That last one could also lead to Prescriptive Analytical insight in helping to answer the question of What a good case framework might be to follow for that case to maximise chances of a win?



## Emerging Applications of Data Science for the Legal Profession.

“Advise on how a particular type of case will fare under certain circumstances.”

A huge misconception when it comes to Machine Learning is that, in order to produce any useful insights, an organisation must have huge amounts of data. It is not quite as black and white as this and in fact many legal firms already have adequate volumes and quality of data to begin exploring the Data Science techniques available to them.

### Case Assessment and Prediction

More and more, law firms are turning to Big Data to identify which cases are likely to be straight forward and which cases will have hidden hurdles to overcome. The basis of a good lawyer comes not just from their understanding of the law but also from the experience they have navigating its countless nuances and dealing with the numerous moving parts that make up the legal profession. Once a lawyer has that experience they become extremely valuable to a law firm in being able to advise on how a particular type of case will fare under certain circumstances. This sort of insight is invaluable but it takes many years to obtain and also usually is only focused on one area of the law as lawyers tend to specialise as they progress through their careers. Through the power of Data Science technologies, it is possible to analyse the large amounts of case data that is available both publicly and internally to provide similar insights for every lawyer at the company.

The types of insight that can be produced are determined by the data that is available but can range from providing insight into potential case outcomes based on input parameters to assessing litigation cases with regards to a patent history.

## Discovery

One of the biggest areas that Data Science can impact the legal profession is through intelligent software providing electronic discovery solutions. Many organisations often store hundreds of thousands, if not millions, of documents in electronic formats. The main reason they do this is for regulatory purposes, but this can mean that for a law firm they will need to deploy many hours painstakingly sifting through all of that data. However, through semantic analysis and predictive solutions it is possible for a machine to handle the grunt work that goes into this stage of discovery. This can be a huge boost to a law firm as not only does it free up associates and paralegals for other tasks but it can also be done in a smaller time frame meaning you get more time to prepare your case.

With Data Science techniques, this type of automation can be taken a step further. Through the analysis of the correct types of data it is possible for a system to automatically present key information from related cases and citing instances of similar cases or rulings in the past. Packaged all together with natural language processing this type of solution can become an invaluable source of research to any law firm.

“Automatically present key information from related cases and cite instances of similar cases or rulings in the past.”

## Judge Analysis

“Highlight relevant arguments that have previously resonated well with that judge.”

We all know that judges are supposed to be impartial, and to the extent they can be most of them are. However, the law is all about interpretation and as not everyone is the same it is inevitable that one judge may interpret something different from another. Through the correct analysis of legal documents, it is possible to identify hidden patterns that would not have otherwise be visible.

This method of analytics can provide valuable insight into how to craft legal arguments in a way that might be better received by one judge over another. Predictive analytic techniques can be deployed here to aid with specific phrasing and can even highlight relevant arguments that have previously resonated well with that judge in similar cases previously.

Over time it is likely that as more law firms gain access to this type of awareness it will put pressure on judges to be more consistent in their rulings. This in turn will both strengthen this type of system but also mean that incorrect rulings will become less frequent.

## Opposition Scrutiny

When planning a strategy for a case it helps if you can understand the opposing lawyer and their company. Understanding how they think and what arguments they are likely to use can give a huge advantage. There is a greater number of open data sources that contain case logs and other legal documentation relating to cases that can be harvested and used in conjunction with Data Science to provide this level of insight into your opposition.

Previously a law firm would conduct opposition research manually, paying out huge amounts for the hours they spent pouring over previous case files looking for relevant information. Using a Data Science enabled solution it is possible to build a profile of your opposition that can be consulted throughout the case. As well as instant access, not just to their previous cases but to the relevant sections of those cases as well through the power of Natural Language Processing.

“Build a profile of your opposition that can be consulted throughout the case.”

## Client and Witness Insights

“Emphasis should be placed on the variety of data.”

The term Big Data is often confused as simply referring to the volume of data. In fact, just as much emphasis should be placed on the variety of that data as well. Every document, photograph, video and social media post is a source of data that can be analysed and turned into useful information. Accessing this data can be a useful way for a legal team to build up a clear picture of a client or a potential witness. This type of information can then be used to indicate any potential areas of concern or advantage that might be exploited during the case in your favour.

Harvesting such data is not as out of reach as it may sound. In reality a law firm could have its own software solution that acquires the relevant data and then automatically links the necessary information insights to the relevant cases.

## Case Strategy Development

Legal proceedings often generate a lot of documents and if analysed properly the statistical data that they provide can be extremely valuable. Data Science can enable law firms to easily examine large amounts of information to find underlying patterns that could not previously be identified by humans alone. Every new case adds mountains of documents that can take a large proportion of a lawyer's time to sift through for each case. This whole process can be streamlined through the use of machines to present lawyers with a condensed set of relevant facts and data relating to their current case. This in turn allows the lawyers more time to understand how to use the material effectively.

More than simply streamlining the process Data Science can provide actionable prescriptive analytics as well. With the right tools, it is possible to take the patterns and connections from thousands or more legal precedents and suggest important strategies that may have been missed through traditional case research. Advice on the way an argument should be crafted depending on the type of case, the presiding judge, the opposing law firm and many other factors can be invaluable, especially to a firm wishing to reduce its overheads without sacrificing quality.

“More than simply streamlining the process Data Science can provide actionable prescriptive analytics as well.”

## Conclusion

With the increase of data that is produced and with no sign of slowing down any time soon, coupled with advancements in machine hardware the legal profession is set for a major data driven overhaul in the coming years. Data Science offers opportunities to increase efficiency and reduce costs as well as allowing greater transparency across the judicial system. The only obstacle to these advantages is the legal profession itself with change not always quick to be adopted.

There are those that fear the loss of jobs due to automation, however for at least the short to mid-term the likelihood of such systems replacing lawyers is far from a reality. However, that does not mean that jobs in the legal profession are not about to change. Previously skills surrounding basic legal research were a large part of a lawyer or paralegal's job. However, with new technology being applied in the legal profession there will be an increased need for legal experts with data literacy skills as well.

We are not currently at this point yet however and with law firms being slow to adopt this technology it is likely that those who do so first will reap the greatest rewards, being able to offer better pricing without compromising service. It is also likely that those who make the shift to data driven services first will be better placed to take advantage of new technologies in the Data Science space as they are developed.

Such combinations between lawyers and machine learning, data and analytics provide a powerful mix that can enable legal firms to better and more efficiently serve their clients while also finding ways to separate themselves from competitors and do a better job of retaining clients and bringing in new ones.