A case study of *Apple v. Samsung* and how big data analytics might have changed the outcome

Chloé Salome Margulis and Brian J. Galli*

School of Computer Science, Innovation, and
Management Engineering,
Long Island University,
Brookville, New York, USA
Email: musikhan15@gmail.com
Email: brian.galli@liu.edu
*Corresponding author

Abstract: In today's rapidly advancing society, big data analytics is increasingly adopted to enhance business functions. Some industries, however, are more reluctant to adopt this new technology, namely the legal industry. The legal industry would benefit from big data analytics in the form of preventing lawsuits from occurring and predicting the outcome of cases. This study presents an analysis of the infamous and ongoing *Apple v. Samsung* case. It then suggests and concludes that a link can be made between litigation strategies and big data analytics, specifically by presenting a hypothetical application of big data analytics to *Apple v. Samsung*.

Keywords: Apple v. Samsung; patent litigation; big data; predictive analytics; litigation; smartphone technology.

Reference to this paper should be made as follows: Margulis, C.S. and Galli, B.J. (2018) 'A case study of *Apple v. Samsung* and how big data analytics might have changed the outcome', *Int. J. Qualitative Research in Services*, Vol. 3, No. 1, pp.21–35.

Biographical notes: Chloé Salome Margulis graduated from the Long Island University Post Honors College with major in Computer Science. She studied big data analytics and its application to patent litigation, and then wrote an honours thesis bridging the gap between the two fields. She is in her 1L of law school at the U of MN and has a concentration in intellectual property and tech law.

Brian J. Galli received his Doctoral degree in Engineering Management from the Old Dominion University. He also received his Bachelor of Science in Industrial Engineering from the Binghamton University and Master of Science in Engineering Management from the Missouri University of Science and Technology. He works as an Assistant Professor of Management Engineering at the Long Island University – Post. He also owns Apex Strategies, Ltd., a company that specialises in continuous improvement consulting and training. He has over nine years of experience in applying continuous improvement tools in many arenas.

1 Introduction

1.1 What is big data?

In today's rapidly changing society, businesses require fast and constant access to data. They will increasingly turn towards big data because of its strength in predictive analytics, data manipulation, and high return on investments. What makes big data unique from other data sources is that it encompasses a massive amount of information that appears in different data formats. Prior to the introduction of big data, businesses were unable to create trends and knowledge out of data that appeared in different formats and types. Now, with big data analytics, data point correlations are made visible when, at first glance, they may appear nonexistent, regardless of the data type and format [EMC, (2015), p.4].

Big data exhibits a high presence in Internet activity. Companies, such as Amazon, Facebook, and Twitter, benefit from collecting this information. They are able to make more advanced, accurate, and thorough predictions and trend analysis about individual customers. As a result, the companies can further organisational goals and maintain consumer satisfaction by targeting specific patterns or anomalies in the produced reports. Big data analytics aids companies to customise individual user pages with suggestions about what to buy, what websites to visit, and what social media accounts to follow. Because of big data analytics, the world wide web is transformed into an intimate interface catering to the specific needs and interests of each individual.

A 2011 McKinsey Global Report officially defines big data as "data whose scale, distribution, diversity, and/or timeliness require the use of new technical architectures and analytics to enable insights that unlock new sources of business value" [EMC, (2015), p.5]. It is crucial that big data optimisation systems are created and maintained so that big data can accomplish all its complex functions per the requests of its various users. Otherwise, minimal value will be gained. For businesses to keep up with the technologically advancing global market and influx of customer wants and needs, it is imperative that they update and expand their resources to encompass all aspects of big data analytic capabilities.

Big data is composed of three key characteristics, those being volume, variety, and velocity. Volume indicates that there is much data to access and use. This characteristic also indicates that some type of expansive storage will be required. Since big data implies the storing of thousands, if not millions, of data points at a time, it is necessary to have storage space that is not excessively expensive to expand. As of now, the more widely accepted storage service is through the cloud, which implies that available space is limitless. Utilising the cloud does come at a cost of security, however, because the data will only be as secure as the security system implemented by the cloud service. Organisations can then build personalised Cloud systems with enhanced security features. Facebook, for example, analyses and stores all user activity in data warehouses. These data warehouses are in the Cloud, and built specifically for Facebook. It is encryption and security features are key to company success, because all Facebook users release personal information to Facebook. Should Facebook experience security breaches and identity theft, its customers will lose faith, and customer satisfaction and profitability will plummet.

The second characteristic of variety implies that the available data will appear in numerous forms, including text, audio, video, and photos. Variety enhances the accuracy

of resulting predictions because more information is utilised to create trends and conclusions. Although a positive attribute, it also guarantees an increased level of complexity. Now, businesses must add an extra step in the data manipulation process. Before performing any analytics, they must convert the data into a single, readable type. Once the data is converted, it ready for use in predictive analytics.

One way to convert the data is via algorithms that assign values to different aspects of the inherited data. Then, the algorithm adds up the values in order to create a general profile for that piece of data. One such algorithm is called a semantic search. For example, in a consumer review of a product, the algorithm will seek out words that correlate with positive emotions and assign those words with +1. The algorithm will determine words that correlate with negative emotions and assign them values of –1. Finally, the algorithm will add up all the assigned values, and if the value is positive, then the user knows that the review was intentionally praising the product. On the other hand, if the resulting value is negative or close to 0, then the user knows that further analysis has to be completed in order to discern what features of the product ought to be fixed in order to improve consumer satisfaction.

The final characteristic of velocity exemplifies the rapid speed at which new data can be created. As a result, trends are constantly changing. It is prudent that a big data analytics system be up to date and able to automatically update with new data. This attribute presents a challenge, but can be overcome with support and effective system implementation that accommodates to a constantly changing environment.

The 3 V attributes make big data more challenging to manipulate and analyse than any other data. This does not deter businesses from pursuing a route towards utilisation of and reliance on big data analytics. As a result, expenditures into its implementation and maintenance are expected to steadily increase. Because of its distinctiveness and popularity, "it [was] predicted the market of big data technology and services will reach $16.9 billion in 2015, up from $3.2 billion in 2010" [Vanarse, (2014), p.4]. This prediction is affirmed by big data's 40% annual growth rate, which is seven times the growth rate of general information and communication technology [Vanarse, (2014), p.5]. Big data acquisition is a developing field that many industries can tap into in order to further organisational goals, profitability, and consumer satisfaction.

Recent studies demonstrate a similar trend in general big data growth. According to the International Data Corporation, in November 2015, "big data technology and services market [will grow] at a compound annual growth rate (CAGR) of 23.1% over the 2014–2019 forecast period with annual spending reaching $48.6 billion in 2019" (IDC, 2015). The three V's and these statistics are some reasons businesses are increasingly attracted to invest in big data. Velocity, variety, and volume are advantageous attributes, but they are hard to capitalise on unless properly addressed in an up-to-date big data system. With such a system, big data's characteristics can be exploited to the upmost advantage for its users.

1.2 What is Apple v. Samsung?

Since 2011, Apple and Samsung have been fighting over the originality of their patents. Apple was the first to bring legal charges by informing the public that Samsung "systematically copied Apple's innovative technology and products, features, and designs, and deluged [the] markets with infringing devices in an effort to usurp market

share from Apple" (Tibken, 2014). These bold claims were the beginning of what would become known as one of the most historic technology lawsuits of our time.

Apple and Samsung argued over five Apple patents and two Samsung ones. Some of the debated features include universal searching, background syncing, quick links, slide-to-unlock, and automatic word correction (Tibken, 2014). By today, such features are expected, if not required, in almost all smartphones. So how can one company claim all rights to such fundamental and necessary smartphone features?

One presumption is that Apple intentionally filed patents that were broad in scope in order to make patent overlap unavoidable. As a result, Apple would be able to leverage itself over competing companies in the industry. If competitors wish to include the features that Apple patented, they would need to engage in cross licensing agreements or face potential lawsuits down the road. Either way, Apple would benefit, and come out of both situations with profit. Creating and holding patents that are broad in scope go against one of the primary guidelines for patent acceptance, however that does not deter Apple from owning such patents.

In the case of *Apple v. Samsung*, since Apple filed allegations first, it had more leverage in the courtroom. Even if the patented improvements were obvious and fundamental to keep the new technology afloat in the market, a company that implemented the same ones would be held liable to Apple. Apple's legal aggressiveness demonstrates its primary objective to maintain an industry clout. If Apple monopolised the industry, however, it would reduce market competition. This leads to a lack of innovation along with slowed improvement and development processes. A competitive market, on the other hand, will force all companies to constantly release new products and improvements to uphold business and consumer satisfaction. Maintaining such competition is crucial to keep the industry and marketplace afloat and advancing.

The initial 2011 allegations stretched out for many years, and its outcome is still in question. In 2012, a decision was reached, thus dictating a win and loss for both parties. The Samsung Galaxy phones and the iPhone 4, 4s, and 5 were said to infringe on each other (Tibken, 2014). Samsung was found more at fault than Apple, owing Apple a settlement of upwards $1 billion. As of December 2015, Samsung brought to court a plea to receive $548 million in settlement back from Apple. It argues that "while Apple may hold the cash for now, if patents are deemed invalid or it were to win a case on appeal, it should be allowed to retrieve at least part of its payment" (Reisinger, 2015). In other words, Samsung believes that it can prove in court that Apple's patents are invalid. This could potentially reverse parts of the 2012 decision. Furthermore, recent research indicates that several of Samsung's patents in fact did not infringe on Apple's intellectual property. This can lead to re-opening the case, re-assessing damages incurred by both parties, and determining what about smartphone technology is fundamental and what is patentable.

In July 2015, Facebook, Google, and eBay joined forces with Samsung to appeal in court for the reimbursement from Apple. They argued, "if Apple's victory is allowed to stand, the possibility of subsequent patent lawsuits utilising the same intellectual property could negatively impact the development of useful modern technologies and have a devastating impact on companies" (Reisinger, 2015). Just as these allies suggest, continuous debate in court can hinder growth in the technology industry by limiting resources for innovation. It may be that the future of patent litigation and technological development depends on how *Apple v. Samsung* continues to unfold and shape the legal playing field.

As long as companies conduct basic improvements on their smartphones and patent them, infringement lawsuits will not cease to exist. Patent monetisers, such as Apple, are ruthless when it comes to suing others for supposed violation because everyone wants to maintain a competitive edge in the marketplace. *Apple v. Samsung* is the major precedent case that does not set a good example, but acts as a fruitful case study in this paper.

1.3 Research methodology

For this paper, research was conducted on existing studies and articles. Searches were made in scholarly databases and search engines in order to uncover articles and literature that pertains to *Apple v. Samsung* and/or big data analytics. No physical study was conducted to collect sample and testable data. Assumptions and suggestions for the application of big data analytics to the *Apple v. Samsung* case were derived from a comprehensive review of all collected literature.

1.4 Objective

The purpose of this paper is to present a case study of *Apple v. Samsung*. The case study is offered in a different light, in which the application and understanding of big data analytics is hypothetically applied. A comprehensive review of existing research on the case will be included in the proceeding sections. An analysis of the decision outcome in the USA court system and its applicability in the Japanese and Chinese courts is identified. Finally, throughout the paper, suggestions on how big data analytics might have changed the outcome of *Apple v. Samsung* are detailed.

1.5 Hypothesis

The primary purpose of this paper is to present a case study of *Apple v. Samsung* and highlight the effectiveness of big data analytics in predicting and preventing such costly litigation. The paper will demonstrate that there is an advantageous bridge between big data analytics and litigation, and that it should be further explored by parties in the industry in order to reap numerous benefits and same time, money, and innovations.

2 Literature review and discussion

2.1 The decision and its applicability in the USA, Japanese, and Chinese legal systems

Although *Apple v. Samsung* was brought to several international courts, the USA decision held the most weight. The US jury decided that Samsung infringed on more of Apple's patents than Apple infringed on Samsung's patents (Tibken, 2014). They reached the consensus that Samsung did not copy some features, such as slide-to-unlock, which appears on almost every smartphone. The jury's reasoning was that one company could not hold a monopoly over such a universal mechanism (Tibken, 2014). This decision confirms the requirement in the patent application process that a patent cannot protect something deemed common or necessary. A majority of smartphones, if not all, requires

a version of the slide-to-unlock feature for security purposes. Therefore, it cannot be defined as a unique characteristic that only one developer has the rights to own.

Apple v. Samsung was also settled in Japan, which has a specialised court called the IP High Court. The High Court's judges are more experienced in IP than former patent examiners who yield a more thorough understanding of technology limitations and jargon aid those presiding in the USA. Furthermore, these judges. Thus, "armed with complete knowledge of technical details and the patent application examination procedures, the system ensures that the IP judges clearly understand the technology and the law when deciding the scope of a patent claim and whether an infringement has occurred" [Bajwa, (2014), p.97]. As a result, there are more informed decision-making and less biased results. In comparison, in the USA legal system, the responsibility rests on the judge or a 'non-biased' jury is required to apply their limited understanding of technology and patents to preside over cases. The judges and/or jury do not receive adequate professional guidance as they do in Japan, making final decisions rather questionable. Additionally, if the USA judges and/or jury are more knowledgeable about the technology in question, it may be perceived as a sign of bias in the final decision.

Whereas the USA jury favoured Apple, the Japanese IP High Court favoured Samsung. Research can only speculate as to why Japan supported Samsung's allegations. Some suggest that the USA jury was predominantly composed of Apple users whereas the Japanese court was biased towards Samsung because it is an Asian company, headquartered in Korea (Bajwa, 2014). Regardless, both decisions demonstrate the difficulty that a legal system has in reaching an unbiased and educated decision. It is even more challenging to reach such a conclusion regarding technology, because everyone relies on it in different countries around the world; we all experience favouritism towards one product versus another. This leads to an inherent bias and obvious divide amongst technology users and future jurors that may preside over patent suits in the USA. Another issue that arises as a result of having to litigate technology is that the laws are different in each country that the technology is used in. As a result, it is more challenging to reach an internationally accepted settlement in technology lawsuits.

Intellectual property disputes are also settled differently in China. Although China did not settle directly on *Apple v. Samsung*, it did bring claims and charges against Apple for selling iPhones that compete with Chinese phones. Many thought China's Shenzhen Baili Marketing Services Co. would lose to Apple, however they were proven wrong. China comes out strong in patent litigation because of its recent reforms. Before 1985, patent law did not exist. Since then, China developed specialised courts with seasoned judges for intellectual property disputes. They follow the same change as Japan undertook. China also offers rewards for noteworthy inventions and ideas in order to stimulate idea development and innovation (Nicas and Chin, 2016). At first, there was an influx of patent filings, because people jumped at the notion of receiving cash rewards. There was also a proportional increase in overlapping ideas. As a result, the Chinese legal system tightened the application process, making it more difficult to patent ideas, and therefore significantly diminishing the rate of infringements. Even with these system changes, China demonstrates competitive innovation unlike any other country.

Although infringement damages in the Chinese legal system are astronomical, the Chinese courts constantly deal with patent suits, especially ones involving foreign parties. USA companies, for example, prefer settling IP suits in China because they are more likely to win overseas than in the USA. To put it into perspective, from 2006 to 2014, 81% of foreign plaintiffs won in Chinese courts. The reasons for this are that "foreign

firms...only sue in China if they are confident they can win" and it is extremely inexpensive to file compared to filing in US (Nicas and Chin, 2016). For example, for a claim of up to $25 million, a firm can spend less than $30,000 on a case in China. On the other hand, in the USA, it would cost anywhere between $650,000 and $5 million (Kerstetter, 2012). In China, as the number of patent infringement cases increased four times over the past three years, so did China's enforcement of the law. The USA system does not practice the same way, and that is another reason we see issues in our system.

A strong patent legal system, as seen in China, may not be feasible to implement in US, mostly because of its cost. To implement China's system here, we would have to reconstruct the entire US legal system, including passing new amendments and bills to install specific IP courts and judges, incentives, rules, and regulations. The cost to train judges to become professionals in IP would be astronomical. Therefore, to achieve similar success rates as witnessed in China, the USA patent system ought to adopt a more affordable route to change. This route is through building and implementing a universal big data analytics system.

Had the USA system already been built with the legal ramifications that China and Japan possess, Samsung might not have received the backend of the settlement in the USA. This brings into question just how effective a big data risk analysis could have been. With such a tool, Samsung could have studied the composition of the jury to foresee the direction the case would take. If the company saw beforehand that a majority of the jury consisted of Apple users, it might have approached the trial differently. The legal costs of pursuing such a case in court would have been more clearly defined, resulting in a pursuance of alternative settlement options. Or, Samsung might have taken advantage of big data risk analysis by pressing charges in a different country where its word may have more weight over Apple. Without a doubt, the costs of litigation could have been compromised if Samsung held access to a big data risk analytics system.

2.2 Thoughts on the outcome

In 2012, Apple received upwards of $1 billion in settlement from Samsung. As of December 2015, Samsung agreed to pay the remainder of $548 million. However, Samsung continues to fight for reimbursement of this amount, as previously discussed (Tibken, 2014). Despite the costs that Samsung owes to Apple, the company is otherwise relieved about the case turnout. Sales did not diminish – they only took a plunge after the Samsung Galaxy Note 7 caught on fire, was banned from aircrafts, and eventually recalled. If anything, following the suit, Samsung began producing bigger and better devices in an effort to make Apple step up its game.

Apple, on the other hand, is not as pleased with the decision, even though it won more than Samsung. Its representatives' adamant opinions were expressed in a press release shortly after the 2012 verdict: "Today's ruling reinforces what courts around the world have already found: that Samsung wilfully stole our ideas and copied our products. We are fighting to defend the hard work that goes into beloved products like the iPhone, which our employees devote their lives to designing and delivering for our customers" (Tibken, 2014). What Apple does not want to acknowledge is that Samsung developers put in the same effort and creativity to create successful and advanced devices.

Regardless of the company that they work for, creators do not receive the proper recognition and protection for their creations. With big data analytics, however, creators

can be ensured a more secure level of protection, whether a large company like Apple or they are independent inventors backs them. The ideal big data system would help creators identify possible overlaps, gaps, and/or similarities amongst existing patents. As a result, creators would not have to fight in court to defend their work, and lawsuits such as *Apple v. Samsung* would not be necessary to assert creative and design rights.

Apple and Samsung continue to pester each other with infringement suits so long as reform is not made to the patent industry. Unfortunately, this means that smartphone patent wars are unavoidable in the short term. We can implement a big data system to resolve the current issues. This would be a necessary precaution in order to prevent smartphone wars from overpowering the global market and hindering growth and innovation. Already, *Apple v. Samsung* shaped the landscape for patent creation and infringement. We do not want the landscape to keep changing, so it is necessary that we stabilise it with big data analytics.

2.3 Decision effects on current calls for change

A jury comprised of smartphone users and patent holders decided Apple v. Samsung. Researchers question if this knowledgeable jury did make an unbiased decision. Despite its educational background in technology, the jury struggled to understand the legal jargon and instructions. That is one reason that experts believe the jury reached a hasty and biased decision (Galli, 2017).

Member of the jury and patent holder Ilagan Hogan admitted that the jurors skipped over several key questions because they did not know enough about the topic nor understand the instructions. One question the jury ignored was whether Apple patents were considered prior art. As previously mentioned, prior art is a key component used to validate patents in the preliminary stages of the patent acceptance process. To gloss over such a key factor is a large mistake on part of the jury. If the jury had been knowledgeable enough in legal jargon and all the characteristics regarding patents, then it may have changed its verdict. It may have recognised that Apple in fact copied existing patents. In this regard, a jury educated in legal concepts and technology is required in order to reach a non-biased, thorough decision on patent related cases. Ideally, a big data system would fill in the gaps of knowledge for the jurors, thus easing the decision making process. A big data system would act similarly to the patent examiners who accompany and aid judges in the Japanese and Chinese courts. Major factors, such as prior art, will be fully accounted for and hasty decisions can be avoided.

In the *Apple v. Samsung* case, when the jury submitted its decision to the overseeing Judge Koh, it was immediately returned to them. Judge Koh found there were careless mistakes in monetary damages and other aspects of the suggested settlement. The errors are a direct result of their hastiness to reach a verdict and to skip important questions. The complicated jargon and instructions deterred the jurors from seeking additional sources to help them better understand the issue, and so they resorted to their own bias to decide on the matter. As previously mentioned, such prevalent cases ought to be dealt with by a technologically and legally educated jury, whose "brand loyalty may [not] cloud the jury's judgment" (Galli, 2017). This is a flaw that can be mended with a big data system that educates the jurors throughout the decision making process. If all the facts are laid out and legal jargon is simplified, the jury will not have a reason to resort to bias.

The big data system should serve an initial purpose of preventing these cases from being brought to court. However, if this fails and a jury must preside, the system would

be advantageous to the jurors. Just like Japan has IP professionals aiding judges, a big data system would be the jurors' aid. Complicated legal jargon and instructions can be translated into simpler terms. The questions and case details also can be simplified so that the jury would not feel pressured to skip important decision factors. A big data system would improve court efficiency, thus deterring mistakes from occurring in the decision making process.

Understanding how different countries presided over *Apple v. Samsung* can help us understand what changes should be made to the USA system. Japan's IP High Court ensures a level of increased efficiency so that the judges who are "armed with complete knowledge of technical details and the patent application examination procedures ... clearly understand the technology and the law when deciding the scope of a patent claim and whether an infringement has occurred". On the other hand, US judges do not receive adequate professional guidance, and often have to rely on their own interpretation of the technology in question. As previously mentioned, US judges and jurors can be professionally guided with the use of big data analytics.

Judges ought to have a blend of technical and legal expertise if they are expected to preside over patent suits. Over the long run, the easiest way to achieve this is with a big data system, which will take the place of a legal assistant. Big data can store mass amounts of information and then apply algorithms in order to create trends, highlight patterns, defragment complex information, and predict future trends.

There are several reasons physical changes to the USA court system are not feasible. For example, we can hire law clerks with a background in technology and patents. However, this would be a costly change and require excess time to find qualified candidates. Creating patent-specific courts is also out of the question. They would most likely be instantiated in the Northern District of California, where *Apple v. Samsung* took place, and the Eastern District of Texas, where the majority of patent cases are filed. As a result, these districts may yield heavier caseloads and run into resource depletion while other districts experience an imbalance between assigned cases and available resources. Finally, there would be an overall tipping of the scale in the judicial system from the poor resource distribution. Solving such an imbalance would require creating new courts that specialise in patent law and spreading them out across many districts. But even doing this would require altering district lines, making costly changes to jurisdictional statutes, and hiring and training of new judges. These are the reasons we cannot pursue physical change to the USA patent system.

The above alternatives are expensive and timely, offering no guarantee of efficacy. On the other hand, big data analytics that is available to lawyers, patent industry workers, and now judges and jurors can help avoid the discussed implications. There would be no need to change the physical judicial system, nor violate constitutional rights or require new amendments to be drafted. The only change would require building a big data system, off a platform of existing systems such as Lex Machina, and educating patent industry practitioners how to take advantage of this service.

2.4 How big data analytics might have changed the decision?

Access to a big data system has the potential to change the landscape of *Apple v. Samsung*. On the most fundamental level, when Apple first sued Samsung, Samsung could have used big data analytics to identify if Apple's allegations were in fact valid. If

they were, then Samsung would have reason to worry when brought to court. Even if Apple was truthful about Samsung's infringements, Samsung could have used the predictive analytics component of big data to identify what the outcome might be and what were the best preventative strategies to pursue in court. If, on the other hand, the analytics demonstrated no evidence in favour of Apple's claims, then the case could have avoided court, saving millions of dollars for both parties.

When in court, big data might have aided both parties with extensive research in case development. The early warning component would not have benefitted either party by this point, but the risk analysis would. The risk analysis would demonstrate trends, patterns, and anomalies based off relationships it could identify between all the stored structured and unstructured data. Finally, the insights gleaned from this type of analysis would better help the lawyers formulate the most concrete evidence-based responses to allegations made against them. If the two companies could present more clear and concise cases and evidence that did not require back and forth debate, perhaps the presiding jury would have had less difficulty breaking apart the evidence and to reach a conclusion. In addition, the case could have been settled in less time.

On the judicial level, big data analytics would have helped judges and jurors. Analytics can perform searches and provide information within seconds, saving the judicial system the four years that Apple and Samsung sat in court. When the parties were in court, the decision and evidence searches were conducted without the use of technology. Sifting through data about existing and historic patents to find clues about potential overlaps is a timely procedure. In an applied big data scenario, the judges would have access to limitless resources and would not have to pick their own brains to procure shallow assumptions. Furthermore, the jury would be educated on legal jargon, decision questions, and patent details. There would be no justification for the jury to skip important questions. The fear that the jury would make a biased decision would be diminished. If the jurors have the facts laid out before them in simpler terms, there is no reason for favouritism and loyalty towards one company to weigh into the decision process.

There are many ways in which big data analytics would benefit patent litigation like *Apple v. Samsung*. The patent industry currently uses big data to an extent. Some firms utilise big data to identify cases that will yield a high success rate. LexisNexis is known for its new service Lexis Advance MedMal Navigator, a platform that predicts the outcomes of medical malpractice suits. Studying past cases enables the software to identify potential witnesses and experts to consult during case development (Dysart, 2013). Lex Machina is also similar to LexisNexis, but is used primarily for intellectual property law. With more than 4 million documents pertaining to patents, lawyers have access to the most lucrative information in the market. Backed by Apple, Cisco, Intel, Microsoft, and Oracle, Lex Machina can map "every electronically available patent litigation event and outcome to bring openness and transparency to IP law" [Katz, (2013), p.32; Galli et al., 2016]. Lex Machina's limitation is its lack of applicability to many firms. It is a beneficial service, but only those who can afford paying for it have access to it. It also does not offer early warning and predictive analytics, both crucial in this context. Lex Machina acts as an online repository of intellectual property court cases and patent records.

Although existing, the current implementation of big data services in the legal industry is severely limited. A key aspect of intellectual property litigation is risk analysis. Because it is "a matter of critical importance for parties involved in IP litigation

to continually assess their respective risks during the entire progression of the case, starting from its filing time, or even prior to filing," proper risk analysis should be available to diminish the possibility of exacerbating costs [Surdeanu, et al., (2011), p.1]. As a result of not performing proper and thorough risk analysis, inventors may lose patent ownership and become indebted to patent monetisers. In some instances, damages paid by the defendant can amount to millions of dollars, as seen with *Apple v. Samsung*. It is usually hard to settle IP litigation because while the law says one thing, consumer use and favouritism might say another about the technology in question.

Risk analysis yields several key benefits. Its immediate effect is to help decide whether to file a case or not. Involved parties can predict which direction the case will take and what costs they will incur. Ideally, this would deter long term and expensive cases from being filed. The proper risk analysis would be composed of "a logistic regression classifier to capture historical features and a novel relational model using conditional random fields to jointly predict the outcomes of concurrent and related cases" [Surdeanu et al., (2011), p.5]. In other words, the logistic regression classifier seeks out anomalies and patterns by analysing historic, unstructured data and applying these findings to the current cases and data that might be related. This algorithm has the capability to produce 22% more accurate outcomes than when a lawyer conducts manual research [Surdeanu et al., (2011), p.6]. In so doing, the user can determine if the current case has the potential to be as expensive and lengthy in court as past cases. The logistic regression classifier will offer the user a prediction of the case costs and time. This is one reason big data is important; it can be used to access past information and create value from it. This is something lawyers have not been able to do in the past.

A risk analysis demonstrates if the costs of arguing in court outweigh the benefits of settling outside. Settling outside requires different mitigation techniques and an agreement where both parties compromise. Usually, one party will not suffer indefinitely at the hands of the other. A risk analysis that compares the two mitigation strategies for individual cases will ease burdens on the judicial system, judges, and juries, resulting in a lighter caseload and more time devoted to analyse and decide on serious lawsuits.

One challenge with a big data risk analysis is keeping it up-to-date. IP cases and patents are filed daily alongside the general trend of technological growth. An analysis platform derived from a traditional relational database would not solve this challenge like big data. Data warehouses are automatically scheduled to constantly update the stored information. Since big data performs functions automatically, involved parties would not have to expend valuable time searching for evidence. A user may only need to know keywords to perform big data fact-finding in case development.

The ideal risk analysis would consider the case's merit and prior factors. Case merit should yield any conflicts that may currently exist or arise in the future between the questionable patents. If there is no relation between the two, then the case has no merit and can be dismissed. A big data platform accelerates this process by reporting about similarities and differences through the implementation of an algorithm, such as a semantic search.

Prior factors will determine the likelihood for success in the courtroom. Factors under consideration include opposition success rates, potential bias of the presiding judges, and circumstances that may result in unprecedented outcomes [Surdeanu et al., (2011), p.2]. This research tool can be implemented through an algorithm, similar to those comprising LexisNexis. However, it is noted that the algorithms in LexisNexis are only advanced

enough to compare competitor costs and success rates. The ideal tool would build upon these capabilities, and further detect relationships amongst cases and judge bias.

Another issue industry professional's face is a language barrier. Patents may implement differing legal syntax, especially if they originate from different countries. Differences in global legal systems make arbitration challenging for universally accepted patents. Because technology is widespread and patented globally for the international market, infringements are often unavoidable. Resolving these cases is challenging because they need to comply with the differing legal systems, and yet they are often resolved in a country that is not their filing country. Patent holders will do this because the decision might be more favourable than in another legal system.

Apple v. Samsung was brought to different countries, each offering somewhat contradictory decisions. Ideally, a big data platform should adapt patents to every legal system, so this would not occur again. It would translate languages and legal procedures, finding similarities amongst the different legal systems. This can help avoid international litigation that does not reach the same decisions and deter companies from choosing some legal systems over others (Martin, 2013; Galli et al., 2016).

Something lawyers currently are not privy to is metadata. Digging into patent metadata can uncover trails to other patents and useful information for case development. The user may be overwhelmed with information, however big data brings together the structured and unstructured data, therefore saving time in research. This also avoids alienating the user with an abundance of information. With big data, lawyers can capitalise on metadata trails (Martin, 2013). For example, creating these links will expose the frequency that similar cases go to court. As a direct result, lawyers can make more informed decisions when preparing for court so that they are not there for a long time. Lawyers can study in more depth the patents' contentious parts and become aware of what the future may hold for the case.

As previously mentioned, LexisNexis is a popular service that firms use. One California based firm, Dummit, Buchholz and Trapp, uses it because it "allows [one] determining in 20 minutes – versus 20 days – if a case is worth taking on. This tool is a powerful, one-stop solution for attorneys" (Dysart, 2013; Galli et al., 2016). Malpractice attorneys are LexisNexis' biggest clients. The service stores domestic and international cases, which is crucial for the reasons previously mentioned. Chief Architect of the company, Ian Koenig, says, "We have the tools and analytics to look across [all] this data in a way that helps attorneys glean insights into potential case outcomes. There are things law firms do not know today simply because they' have never stored and correlated all this data" (Dysart, 2013). This is important because it opens many doors for predicting case outcomes and finding the most efficient way to resolve a case. In patent litigation, it will be incredibly beneficial since patents reach beyond domestic boundaries and question other patents around the world.

Alongside LexisNexis is Lex Machina, offered only to IP firms to reap similar benefits as malpractice attorneys do. Lex Machina searches the Internet for information that can be used for existing cases. Collected data is stored in a database, where basic algorithms are implemented to map out possible correlations. The service can predict the direction a case may take. Silicon Valley IP attorney Vicki Veenker expressed her faith in Lex Machina: "My prediction is that in the long term, others will join me and big data will bring more transparency to the IP marketplace" (Dysart, 2013). Her support for big data implies that it can be very helpful to the patent industry. This provides hope for a future in this field in which big data is widely accepted and used for all its capabilities.

All the features of a big data analytics system could be taken advantage of in the *Apple v. Samsung* case. An effective analysis and mining of historical data, past cases, and online articles and information about smartphone technology and the two companies would have been made possible. Informed results would be presented to the judge, jury, and the two parties in less time than it took to settle the case in court. Imagine the costs saved from such a service. And then imagine the effects that this risk analysis would have had on the future of smartphones and patenting technology in general.

Overall, the ideal big data system would have saved time, money, and efforts during *Apple v. Samsung*. In the long run, it would reform the current issues outlined above in the patent industry. Implementing a universal big data system would be less expensive than applying physical changes to the legal system, therefore making it a more reasonable option that is also available to all patent industry practitioners. Big data has the potential to decrease the amount of smartphone wars that fill up the courts. In so doing, a big data system will not diminish innovation or the competitive nature of companies in the global market; it will merely make companies and inventors more precautious and proactive in their decision-making and pursuit of new ideas.

3 Suggestions for future research and contributions

It is suggested that future research be conducted on the actual application of predictive analytics to the *Apple v. Samsung* case. A study can be conducted in which big data analytics is applied to a simulation of the case in order to discern how effective the analytics would be in predicting the case flow and outcome. Furthermore, research can be conducted into what features of a big data analytics system would be most beneficial to resolving issues seen in *Apple v. Samsung*. Future research can be used to identify where big data analytics may be applicable to other specialties in the legal industry, such as bankruptcy, criminal, and corporate law.

This research study was limited in scope of actual testing and implementation due to minimal funding. However, this study contributed valuable ideas and input into the existing body of knowledge. It suggests that a link can be made between litigation strategies and big data analytics, specifically by analysing and presenting a hypothetical application of big data analytics to *Apple v. Samsung*. This case study creates a new avenue for research to be conducted that will further the ideas presented here.

4 Conclusions

The application of big data analytics to *Apple v. Samsung* might have resolved some of the underlying issues inherent in the patent industry. Revisiting *Apple v. Samsung* and applying big data's features exemplifies the bridge we seek to build between patent litigation and big data. In this case, predictive analytics would have illuminated the direction the case might take in court. It would have prepared both parties for the astronomical fees they would incur during the four years of courtroom proceedings. Big data's risk management would have helped both parties allocate resources more efficiently and perhaps would have urged Samsung to bring the case to another country, such as China or Japan. In China, it would have been cheaper to pursue, and the legal

system might have demonstrated more favour towards Samsung versus Apple. This is because China's IP specific court has educated judges in intellectual property and patent law who take into account all the factors of a case. No question is left unresolved. The decision might have been different, resulting in a smaller settlement payout than in the USA. Based on some facts from the case, foreign legal systems may have deemed Apple more liable for infringement than Samsung.

If this big data feature were brought to Japan's patent system, the decision in Japan would not have changed. Big data would have no major impact on the *Apple v. Samsung* decision because Japan already has a specific court for IP law. This country focuses its resources on developing the IP court, which is led by educated judges and patent experts. Therefore, Japan would not even require a big data system, because the system is not as flawed and uneducated as US intellectual property judicial system. The Japanese court is precise in its decision-making and is unlikely to oversee important case details like US jury did for *Apple v. Samsung*. USA, however, does not have the same resource allocation in its judicial system, and therefore the best option to reform the flaws of the patent system is to develop a big data analytics platform.

Finally, big data's early warning and risk management functions might have helped Samsung avoid a lawsuit in the first case. Courtroom litigation would be avoidable and the two parties might have had the opportunity to alter their patents so that they did not overlap. With the application of big data's three major features, *Apple v. Samsung* could have been resolved from the get go. Such real-time application and proactive planning would save money and resources for all involved parties in the patent industry. As one can see, the application and implementation of a big data analytics system would have been tremendously beneficial in resolving the *Apple v. Samsung* case. It is only hoped for that in the short term, companies such as Apple and Samsung will develop an interest in big data analytics in order to protect their ideas from costly patent warfare and predict how existing cases will turn out, preferably to their advantage.

References

Bajwa, S. (2014) 'Apple v. Samsung: is it time to change our patent trial system?', *Global Business & Development Law Journal*, 4th ser., Vol. 27, No. 1, pp.1–31 [online] http://digitalcommons.mcgeorge.edu/cgi/viewcontent.cgi?article=1003&context=globe (accessed 10 July 2017).

Dysart, J. (2013) *How Lawyers are Mining the Information Mother Lode for Pricing, Practice Tips and Predictions*, 1 April [online] http://www.abajournal.com/magazine/article/the_dawn_of_big_data/ (accessed 28 March 2016).

EMC (2015) 'Data science & big data analytics: discovering, analyzing, visualizing and presenting data', *EMC Education Services*, pp.1–27, DOI: 10.1002/9781119183686.fmatter.

Galli, B. (2017) 'How to truly win in business with leadership – a case study report', *Middle Eastern Journal of Management*, Vol. 4, No. 3, pp.235–245.

Galli, B., Santos-Arteaga, F.J., Di Caprio, D. and Kennedy, D. (2016) 'Do ethical leaders exist? A unified theoretical framework to identify and evaluate them', *International Journal of Management & Decision Making* (*IJMDM*), Vol. 15, Nos. 3–4, pp.10–23.

International Data Corporation (IDC) (2015) *New IDC Forecast Sees Worldwide Big Data Technology and Services Market Growing*, 9 November [online] http://www.idc.com/getdoc.jsp?containerId=prUS40560115 (accessed 31 March 2016).

Katz, D.M. (2013) 'Quantitative legal prediction', *Emory Law Journal*, Vol. 62, No. 4, pp.1–58 [online] http://law.emory.edu/elj/_documents/volumes/62/4/contents/katz.pdf (accessed 10 July 2017).

Kerstetter, J. (2012) *How Much is that Patent Lawsuit Going to Cost You*?, 5 April [online] http://www.cnet.com/news/how-much-is-that-patent-lawsuit-going-to-cost-you/ (accessed 30 July 2016)

Martin, J. (2013) *Big Data for Patents: Deriving Novel Insights from Publicly Available Data*, 28 October [online] http://insights.wired.com/profiles/blogs/big-data-for-patents-deriving-novel-insights-from-publicly#axzz3pLzu0Xhb (accessed 28 March 2016).

Nicas, J. and Chin, J. (2016) *Stronger Chinese Patent Laws Also Help U.S. Companies* [online] http://www.wsj.com/articles/stronger-chinese-patent-laws-also-help-u-s-companies-1468994404 (accessed 30 July 2016).

Reisinger, D. (2015) *Samsung Will Pay Apple for Damages – But Wants The Cash Back*, 4 December [online] http://fortune.com/2015/12/04/samsung-pay-apple-patents/ (accessed 10 July 2017).

Surdeanu, M., Nallapati, R., Gregory, G., Walker, J. and Manning, C.D. (2011) 'Risk analysis for intellectual property litigation', *Proceedings of the 13th International Conference on Artificial Intelligence and Law*, pp.116–120 [online] http://0-dl.acm.org.liucat.lib.liu.edu/citation.cfm?id=2018375 (accessed 10 July 2017).

Tibken, S. (2014) *Apple v. Samsung Patent Trial Recap: How It All Turned Out*, 7 May [online] http://www.cnet.com/news/apple-v-samsung-patent-trial-recap-how-it-all-turned-out-faq/ (accessed 28 March 2016).

Vanarse, S. (2014) *Big Data Breathes Life Into Next-Gen Pharma and R&D*, WIPRO: Applying Thought [online] https://www.wipro.com/documents/big-data-breathes-life-into-next-gen-pharma-RD.pdf (accessed 15 July 2017).