Tracking the New Demand for Justice in the Big Data Ecosystem

Carmen Vargas Pérez[1*]
Juan Luis Peñaloza Figueroa[2]
[1]Department of Applied Economics, Public Economics and Political Economy, Complutense University of Madrid, Spain
[2]Department of Statistics and Operations Research II, Complutense University of Madrid, Spain
*Email: cvargas@der.ucm.es

DOI: 10.2478/ejme-2023-0017

Abstract

Many studies have focused on the possibilities that organizations have to mine and analyze, through computational analytics, the huge amount of structured and unstructured data that is now available, to determine correlations and thus reveal patterns, trends, and associations to predict human behaviour; and to transform this information into knowledge for companies and governments. That is, just from the point of view of the suppliers of good and services. In this paper we contribute to the Law and Economics literature by showing that the logic of Big Data, the access to the cloud, and the use of Artificial Intelligence are drastically changing the ordinary citizen's way of making decisions in the field of justice; and that this new paradigm in the Demand for Justice will mean improvements in terms of both equity and efficiency, and ultimately an improvement in social welfare.

Keywords: Litigation, Justice, Legal, Courts, Law and Economics, Big Data, Ecosystem, Demand.

JEL Classification: A12 C81 D11 D43 D63 D82 H49 H50 K40 K41 K49

Introduction

Big Data logic and solutions are increasingly penetrating all economic and productive industries. It provides a wide range of benefits, including innovation, communication, and freedom of expression. And as the Big Data ecosystem deepens its incorporation into everyday life, economic and social activity, it generates value and provides actionable intelligence for governments, businesses, and individuals.

Both the definition and the scope of Big Data have evolved exponentially in a short period of time. Big Data and the use of Artificial Intelligence allow us to convert that enormous amount of structured and unstructured data into knowledge, and it can be of great help in making decisions in any field. For example, many applications have been developed in the business environment, ranging from increasing and valuing the performance and productivity of businesses -enabling them to leverage data of different types and optimize their warehousing- to improving the security of companies in all industries.

As Siegel (2016) points out, predictive analytics may help find ways to increase consumer buying, limit bank loan defaulting, anticipate employees quitting, or predict which people will drop out of school, cancel a subscription or get divorced before they are even aware of it themselves. And of course, these applications include how insurers use life expectancy estimates to make policy approval and pricing decisions; how the state of Maryland (US) uses predictive models to detect inmates more at risk to be perpetrators or victims of murder; or how software like PredPol predicts progressions that could lead to waves of crime.

Today, Artificial Intelligence tools are used in a wide range of areas such as national security, disaster and resource management, climate change, conflict prevention, disease eradication, economic development, good governance, management of cities and many government economic policy instruments. On an individual level, an overwhelming majority of citizens participate in the Big Data on a day-to-day basis. To begin with, having his life so digitalized, the ordinary citizen is continuously transmitting a huge amount of data about where he is and how long he stays in each place, what device he is using, what he is doing with it, what he is searching for, what kind of physical or virtual products he is interested in, etc. And at the same time, Big Data allows him to receive personalized suggestions on purchases, know the time of the next bus, automatically include his next trip in his agenda, receive (official or other drivers´) alerts about traffic jams on the road, know the problems a patient has had in a hospital he is going to, or see photos in real time, made by tourists with their mobile phones and posted on tourist services websites; in addition to all the information shared on social networks. Almost without realizing it, humanity produces and sends millions of bytes to the Internet every second[1].

Various interesting applications already exist in the field of justice. For example, a large number of Cloud-based legal practice software applications have appeared, enabling lawyers to improve their performance by reducing search time for legislation and legal precedents, including tools for time tracking, billing and case management, among others. In addition, using the tools of Collective Intelligence, some researchers have conducted a crowdsourcing study that predicts Supreme

[1] According to Mikal Khoso from Northwestern University, 2.5 exabytes are produced every day, which is equivalent to 530.000.000 songs, 5 million laptops or 90 years of HD video. http://www.northeastern.edu/levelblog/2016/05/13/how-much-data-produced-every-day/.

Court decisions. And other studies look at how Big Data techniques are impacting environmental justice activism or how predictive policing software can contain and generate biases as pernicious feedback loops are created.

Surprisingly, little attention has been paid to the effect of Big Data, the access to the Cloud and the use of Artificial Intelligence on the demand for judicial and legal services. The aim of this article is to show that the logic of Big Data and these technologies are drastically changing the way the average citizen makes decisions in the field of justice; and that this new paradigm in the Demand for Justice will mean improvements in terms of both equity and efficiency, and ultimately an improvement in social welfare.

In the next Section we briefly discuss some examples of the breakthroughs that the use of Big Data has had in various spheres of economy and society, showing the challenges that it still presents; and in the third Section we describe the use of Big Data and Artificial Intelligence (AI) in the field of Justice in general. In the fourth Section, we analyze the effect that the Big Data ecosystem has on litigation. Finally, the last section presents some conclusions and final ideas on the new risks that its use can bring to the generation of new statistics in this field and to decision making in general.

The Big Data Breakthroughs

Although the first reference to the term Big Data at an academic level appears in Cox and Ellsworth (1997)[1], the widespread use of Big Data as we know it today dates back only from 2012 (a lifetime ago in the Internet time).

In this short space of time, the definition of Big Data has evolved significantly. The initial concept incorporated only a flat definition, which only associated it with the existence of large volumes of structured and unstructured data that could not be processed with traditional tools[2]. Later on, the concept was expanded to include not only the data itself, but also the set of technologies that allowed data mining and its conversion into useful information, and into tools for making business decisions using Artificial Intelligence (AI).

In the business world, during the last 30 years, we have talked about Business Intelligence (BI) to make reference to the set of strategies and tools that a company has at its disposal to be able to analyze the data of its organization. And about Data Mining, which uses mathematical analysis to derive patterns and trends in data, making it possible to detect actionable information from large datasets (Puyol Montero, 2014). We now use a broader concept, which incorporates a change in the

[1] *"Visualization provides an interesting challenge for computer systems: data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk. We call this the problem of Big Data. When data sets do not fit in main memory (in core), or when they do not fit even on local disk, the most common solution is to acquire more resources."* Cox and Ellsworth (1997).

[2] The most widely used initial definition includes the three V's of Volume, Variety and Velocity by analyst Doug Lane of Gardner Consulting. Subsequently, many definitions have emerged incorporating other V's such as Value, Veracity, Visualization, Viability, Variability, Validity, Vulnerability or Volatility. See for example PowerData (2014).

way in which a diversity of areas of society such as trade in goods and services of all kinds, industry, technology, healthcare, education, arts, humanities and social sciences, among many others, will develop in the coming years, thus becoming a paradigm shift in the use of information. As Hilbert (2016) writes, "*the key feature of the paradigmatic change is that analytic treatment of data is systematically placed at the forefront of intelligent decision-making. The process can be seen as the natural next step in the evolution from the' Information Age' and 'Information Societies'... to 'Knowledge Societies*' *"*.

During these years, a good part of the effort related to Big Data has been concentrated on the development of software that allows the management of this enormous amount of data, not only according to its volume but also according to the many and diverse sources from which it could come; and on the possibility of not hosting these data in a centralized way - which at first was thought to reduce costs-, but distributed in many nodes. This eventually led to the introduction of Block Chain technologies that, having their origin in the management of BitCoins, added advantages in terms of information security (transactions)[1].

Recognition of Big Data's scope has also grown significantly in this short period of time. The first practical advantages of Big Data were evident in its ability to predict customer behaviour, both on the type of products to be consumed -which allows the organization of the production structure- and on the quantity, place and dates of consumption -to speed up the distribution of goods in those places and moments where they will be demanded. This has made it clear that a substantial part of the company's revenue flow depends on the data provided by customers. In the digital age the most valuable resource is data, so since the dawn of this new era -the Big Data one- it became clear that the first requirement for selling in the Cloud was to convince potential customers to give up their data.

At this point it is also worth mentioning that, in parallel with this development, and particularly in recent years, various sources of concern have arisen from governments, citizens and experts in these technologies, both for the privacy of personal information and its ecological impact, and for the inability of old computer security systems to counteract new methods of cyberattack[2]. According to Kate Crawford, principal researcher at Microsoft, it is dangerous to think that Big Data, including that large-scale data, can be made anonymous are inherently objective and include tacit or explicit consent or an opt-out function; research by computer scientists continue to show that even anonymized data can be reidentified and attributed to specific individuals[3].

[1] The most popular strategy for analyzing these vast amounts of information is the Hadoop open source platform, inspired by Google GFS (Google File System) and the MapReduce programming model, which provides a parallel and distributed data processing system. See for example: https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/.
[2] See for example O'Neil (2016).
[3] See Neff (2013).

Much of Big Data's potential lies in discovering, with the help of the fastest equipment and new analytical techniques available today, the hidden correlations that may exist between the hundreds or thousands of existing variables in these enormous volumes of data. This allows patterns, trends, and associations to be revealed, with the aim of predicting human behaviour. One of the best known pioneering examples is that of Amazon company which, in order to increase book sales, used to have a team of editorial advisors. This team, after studying the past purchases of each individual client, offered recommendations for the purchase of other publications. This sales incentive system was not only expensive -as it was very intensive in the use of qualified personnel who dedicated individually to each former purchaser- but its effectiveness was not very high as it generally ended up in the mere recommendation of titles similar to those already purchased by each client. When Amazon began to use its own Big Data, compiled from the data of hundreds of thousands of its customers, not only with respect to their past purchases but also to countless personal characteristics, purchase recommendations began to become much more profitable. First, it allowed a wide variety of editorial suggestions, which became effective purchases at a higher rate[1]. In addition, being a result of digital algorithms, the service could be offered at a very low cost, in a massive way, and in a very short period of time.

Big Data's ability to predict future behaviour has been one of the main arguments for promoting the massive use of technologies to harness its potential. In some cases, this fervent defence of its use has come to the extreme, saying that these technologies will replace the statistical-econometric analyses that study causality[2]. The underlying argument would be that it is no longer necessary to know Why things happen. To give an example from the health sector, if Big Data can predict when the next peak of purchase of anti-flu drugs will occur, it would seem irrelevant to the laboratory to know how the virus entered the country or how it spread among the population[3]. In our view, however, it is quite clear that the sole predictive capability of Big Data does not solve the health problem. Although the possibility of predicting which people will develop cancer in the future has enormous value for early diagnosis, it does not solve the underlying problem. In order to obtain a cure or vaccine for this disease, it is necessary to study its mechanism; in other words, in order to finally discover how to reverse the process from its origin, it is necessary to analyze how the disease is born and how it spreads in the organism. As Dunham (2015) writes, "*problems arise when*

[1] See Mayer-Schönberger and Cukier (2013).
[2] In 2008 the editor-in-chief of Wired magazine had made the bold statement that Big Data has the potential to render the scientific method obsolete (Anderson, 2008).
[3] In 2009, the results of an already famous algorithm that allowed Google to detect influenza epidemics using search engine query data and CDC records were published in Nature. This made it possible to predict the next flu outbreak well before the health authorities. This could potentially lead to better health authorities preparedness and better distribution of medicines by laboratories, increasing their profitability. Although the results could not be replicated effectively in subsequent years, these results gave a great boost to the dissemination of the benefits of the Big Data exploitation (Ginsberg et al., 2009).

decision makers rely solely on data and omit vital knowledge about the subtleties of citizens and processes".

These limitations of Big Data for problem-solving are extrapolated to almost any other field, including of course the field of justice. While it is true that the possibility of predicting a sudden peak of demand for judicial services in a given geographical area could be very useful in the early allocation of support staff, it is clear that this short-term measure will not solve at all the structural problems of justice delay, as the increased supply of court services will most likely lead to a subsequent increase in litigation, largely offsetting the benefits of the reinforcements allocated. This feedback effect between supply and demand -which explains why continuous increases in the resources devoted to justice do not usually reduce delay-, is well known in the field of law and economics[1] and only working on it will help to mitigate it. Thus, the mere prediction of future litigation levels based on the new correlations generated by Big Data, although very useful for the study of new causal relationships, is not sufficient for the solution of existing problems. As Google's chief economist Hal Varian says, *"It's not the ingredients that are critical, it's the recipe... the algorithms, the analysis, the business decisions that come from looking at that data and analyzing and understanding it".*

Despite these shortcomings, the potential of the predictions offered by Big Data logic, along with the use of AI, is huge. And the scope of these technologies can be endless. For example, it is now publicly known that, using AI technologies, the US National Security Agency analyzes foreign communications, including screening locations, words and transactions in Internet searches, obtained from a range of sources including US-based communication providers such as Microsoft, Yahoo, Google and Facebook (Cumbley and Church, 2013). The justification given is that the analysis of that unstructured data, including facial recognition technology, has helped prevent many terrorist attacks. Also, with the use of AI, banks and insurance companies are improving their credit scoring systems, and tax offices increase the detection of benefit fraud or tax evasion cases. In addition, these technologies make it possible to predict some macroeconomic variables ahead of official statistics, allowing for earlier economic policy actions[2]. And the use of the Big Data, the Cloud and the AI applications in the field of justice is no exception. We are devoted to this subject in the following Section.

The Current Use of Big Data and Artificial Intelligence in the Field of Justice

Forty years ago Prendergast (1977) offered the first insights into how data processing could benefit legal practitioners. Of course, the World Wide Web did not exist at that time and there was no mass gathering of information. However, this author showed that it was already possible to visualize the potential that systematization of

[1] See Priest (1989), Dakolias (1999), Vargas Perez and Peñaloza Figueroa (2017) or the book by Pastor Prieto (2016).
[2] Hilbert (2016) presents a fairly detailed overview of the application of Big Data for Development.

information could have[1] and also the reluctance of some law professionals to rely on the results obtained.

For years many professionals have not considered the data as a source of value but only as a source of more work[2]. And although there were already some advocates, in the legal area mistrust of the potential of digital tools emerged long before the current concern about the problems that the Cloud and the AI could pose for the security of personal data[3].

What benefits has the current use of Big Data brought in the legal and judicial sphere? The IT tools associated with the Big Data ecosystem offer cost, time and money savings and increased profitability for law firms. As Prendergast (1977) forecasted, "*with an understanding of data processing, the attorney will be able to provide better service to his clients at reduced cost and handle complex cases involving masses of information. Using the technology available, he can effectively and efficiently marshal the masses of data at his disposal and correlate this data to his legal knowledge of the case; and this combination will be highly effective*".

In recent years, a large number of cloud-based legal practice software applications have appeared enabling lawyers to improve their performance by reducing time spent searching for legislation and legal precedents, including tools for tracking time, billing and case management, among others[4]. According to Nelson and Simek (2013), analytics software can speed up management tasks, such as distributing cases, projecting revenues, projecting case budgets and -most important of all- predicting outcomes and make a fee estimate based on previous matters involving similar factors. With all this, it is also possible that law firm management budget forecasts may be more reliable[5]. Moreover, in recent times, as a result of the increasing power of computers, large law firms start to incorporate Big Data in their private activity and their internal operations, gathering information from thousands of their own prior clients, which allows them to estimate their chances of success in each specific case, depending on which court the case has been assigned to and other specific factors[6].

In addition, although its use has not spread too widely, there are several interesting essays that apply Big Data and Collective Intelligence in the field of justice. For example, as mentioned above, a group of researchers has conducted the largest crowdsourcing study to predict Supreme Court decisions[7]. Players of Fantasy SCOTUS

[1] *"Once the attorney gains an understanding of data processing, the outstanding capabilities of this technology can be available to him, with imagination again being the only real constraint"* (Prendergast, 1977).
[2] For a discussion of this point in the medical field, see Neff (2013).
[3] Regarding the difficulties in disseminating the creation and use of digital databases in the field of Spanish justice, see for example Pastor Prieto and Vargas Perez (2001) and Consejo General del Poder Judicial (1997).
[4] See for example http://www.capterra.com/legal-case-management-software/ .
[5] On this topic see Vargas and Peñaloza (2017).
[6] Attention should be paid to this type of software, as it handles private data about millions of citizens and their cases, and is being used by (law) firms who, despite knowing the legislation on data protection, are not used or prepared to protect that information, so it could easily be captured via the internet for unintended uses.
[7] Emerging Technology from the arXiv (2017).

(an online fantasy league) do not need any special qualification, but as Katz et al. (2017) have shown, crowds can make surprisingly successful decisions, sometimes even better than the smartest among them. In fact, this type of results led these authors to construct a model designed to predict the behaviour of these courts in a generalized, out-of-sample context, achieving outstanding results (70.2% accuracy at the case outcome level).

In addition, as in the commercial field, AI's tools begin to help to advertise legal services more effectively. And at the same time, they increase the quality of services offered to clients by allowing lawyers to have easier access to more information on precedents and legislation (Vargas Perez and Peñaloza Figueroa 2017).

Of course, although its potential is enormous, the logic of Big Data and the AI cannot do without the expert judgment of lawyers and legal analysts. It is necessary to make sense of the data in order to transform that information into knowledge and be able to make efficient decisions. And that should be the goal of any Artificial Intelligence software applied to the field of justice. As Ford (2016) states, after observing patterns and trends of behaviour we need theoretical frameworks to interpret that behaviour: "*Once the goals are defined, the attorney, with the assistance of data processing personnel, can work backwards from these goals and define, step by step and in minute detail, exactly how the computer is to use the information to reach those goals. This process results in the systems design*". That is, we need what Wang (2013) calls Thick Data: "*Big Data delivers numbers; Thick Data delivers stories. Big data relies on machine learning; Thick Data relies on human learning*".

Finally, it should be remembered that the quality of justice received by the average citizen does not depend solely on the services received from their legal representatives. The functioning of justice depends on many other factors, being delay particularly worrying. This is the reason of why the study of the supply and especially the demand for court services has occupied a large part of the economic analysis of law in the last forty years[1]. The demand for justice in the new Big Data ecosystem is what we will discuss in the next Section.

The Potential of the Big-Data Ecosystem in Litigation

The way citizens approach the world of justice is changing. The lack of information, as a barrier to access to justice, is disappearing thanks to the thousands or millions of users who exchange experiences on the Internet. Thanks to the Internet of Things (IoT) and other devices, we all contribute to Big Data on a regular basis and can make use of the existing information on the Internet and the advantages of Artificial Intelligence. And there is not a big distance between asking other internet users about the characteristics and properties of an item for sale, and asking them about hiring

[1] See for example Voigt (2016), Dimitrova-Grajzl et al. (2012), Rosales-López (2014), Mora-Sanguinetti (2016), Pastor Prieto (2016) and Pastor Prieto and Vargas Perez (2001). And seminal papers by Landes (1971), Posner (1973) and Shavell (1982).

one lawyer or another, or about the convenience of filing a lawsuit or rather reaching an out-of-court settlement.

In the last decade, the Internet has replaced the telephone -and even part of face-to-face direct communication- as the basic, bidirectional and general-purpose means of communication. Cloud computing allows an on-demand self-service of information and most individuals now have instant access to vast amounts of information, which provides a wide range of benefits.

Unlike the legal business environment where it may be necessary an investment in software and, above all, a change in the mindset of lawyers to make intensive use of the logic of Big Data in their daily work, in the case of ordinary citizens this obstacle no longer exists. A large part of the population, at least in developed countries, has a smartphone or another device for accessing the Internet and they are increasingly used to turning to it for information, opinions, and experience of other citizens on various subjects.

In the last 40 years, several papers have examined the possible determinants of the demand for court services[1]. Decisions on whether or not to go to trial depend on, among others, the estimates of the costs of going to trial, the estimates of the chances of winning, the estimates of the waiting times to obtain a court judgement, the believe that obtaining a judgment means that the defendant will automatically pay what the sentence states or the belief that no tax must be paid on the amounts received after the trial (which includes the winning party's costs paid by the losing party). That is, the demand for court services crucially depends on the information parties have (and the estimations they make) on different variables that are often unknown to a greater or lesser extent and/or factors that are not under their control, even after filing the lawsuit and during the legal proceedings (Vargas Perez and Peñaloza Figueroa, 2017).

In the first place, with respect to the estimation of the Chances of Winning the Case, Voigt (2016) concludes that if both parties have a "realistic" evaluation of their chances of winning (but also of losing) the case, then we should not observe any civil trials as a pre-trial bargain is expected to be systematically cheaper than taking the case to court. On the other hand, Pastor (1993) states that agreements occur when the perception of results coincide, or when the parties are very pessimistic[2]. As demonstrated in Vargas Perez and Peñaloza Figueroa (2017) such negotiation does not depend on whether their estimates are realistic or not, but on the relationship (and difference) between such estimates. To a great extent, in the absence of more information, those estimates depend on the information transmitted by lawyers to their clients and in general on the amount of information that the parties have. But now, in the Big Data ecosystem, these estimates will be increasingly similar -as this same information will be available to both parties- which is likely to increase the

[1] See for example the survey by Voigt (2016).
[2] See also Voigt (2016).

number of cases solved by out-of-court settlement and will lead to a reduction in the demand for court services.

In the second place, when estimating the Litigation and the Agreement Costs, clients usually have much less information than lawyers about the stages of the proceedings, the time that can elapse until a final judgment is reached, the private costs of the proceedings or the possibility of an out-of-court settlement. With such lack of information, the estimations may be influenced by the strategic behaviour of lawyers[1], as their incentives to go to trial or to bargain may be different from those of their clients[2]. However, the increasing use of the logic of Big Data not only allows for more and better information obtained from "real users", but it also changes the way the client deals with the legal proceedings. When the ordinary citizen makes use of the logic of Big Data, he can get first-hand information about court proceedings and other possible ways of resolving the conflict, and also on the costs incurred by hundreds of users in the past; everything in real time, detailing the type of matter, the amount at issue, the place or any other variable that may alter the outcome; variables that may influence his decision to go to court, and the strategies to follow.

Also, the Big Data ecosystem not only changes the "estimates" of the costs of the proceedings but the costs themselves, especially those relating to lawyers' fees, for a couple of reasons. First, as the market becomes more competitive, the fees per case tend to fall; and secondly, because in the Big Data environment prices move towards contingency fees.

In addition, as shown in Vargas Perez and Peñaloza Figueroa (2017), this lack of information makes consumers of court services face a number of Asymmetric Information issues, including Adverse Selection and Moral Hazard/Principal-Agent Problems and has important effects on the possibility of reaching out-of-court agreements. Nevertheless, with the greater access to the Big Data ecosystem, the problem of adverse selection will be reduced as potential clients will be able to obtain very extensive information from previous clients about the type and quality of the services offered by different law firms, which will allow them to better choose the services they need to contract. And although the individual citizen will only have few additional tools to carry out the monitoring of his own lawyer´s performance, by making use of the logic of Big Data he will be able to get information from other clients' experience. This will be perceived by law firms as a greater control over their work and will change the lawyers´ performance, their strategies in court and the overall effectiveness of their work, thus reducing the moral hazard issues.

Moreover, the fact that citizens can study in depth the conditions they will face in a possible lawsuit, from their own home or office, encourages clients to be more daring[3] in hiring new services or professionals with whom they have not previously dealt,

[1] For some evidence on the effect of lawyers on litigation see Mora-Sanguinetti and Garoupa (2015).
[2] See for example Wistrich and Rachlinski (2013).
[3] See Johnson et al. (2002) and Universidad Barcelona, AQR Lab and Cambra de Comerc de Barcelona (2018).

allowing new law firms and lawyers to open up a place in the market. This further enhances competition, allowing customers to afford lower prices and access services closer to their preferences, thereby increasing efficiency in the professional services market.

Besides, the new exchange of information between users of justice will allow the number of disputes initiated by overly optimistic expectations to be reduced, so that some of the inefficient disputes will disappear (thus reducing the -inefficient- collapse of the courts). In turn, this will allow these courts to resolve cases of other citizens or companies that do require a judicial solution, in a shorter period of time.

In addition to the benefits mentioned above, this new access to the Big Data logic will encourage the ordinary citizen to have more incentives to know his rights and options for action in society. The breakthrough of the Big Data ecosystem into the environment of the potential user of justice will allow him to better understand the content and meaning of the laws in general, and the law applicable to his particular case, while improving his perception of having made the right decision.

Finally, with access to Big Data and the use of Artificial Intelligence tools, the potential user of the legal and judicial system begins to make his decisions with a broader knowledge of the existing alternatives and the results that each had in cases similar to his. This can be interpreted as a better-informed decision-making process -which increases the efficiency of the actions taken-, a greater and better access to justice in its broadest sense and also an incentive for users to become more involved in the process and to take more responsibility for the results.

Therefore, it is clear that although Big Data and the use of AI are generating enormous improvements in efficiency -with the consequent increase in profits- for suppliers of goods and services, considering Big Data as a useful logic just for companies and governments means underestimating the great potential that changes in human interactions with new technology have for the well-being of the average citizen. In short, this new Big Data ecosystem promotes greater information symmetry, more access to services that meet consumer needs, freer entry of new professionals and lower costs faced, increased competition, better services of lawyers to their clients and better decision making by and for citizens. Overall, it is a clear gain in terms of Efficiency and Equity, and ultimately an increase in Social Welfare.

Final ideas

The access to the Internet, the Cloud and the Big Data ecosystem, and the use of Artificial Intelligence are significantly changing the demand for court and legal services. It is not just a matter of having more data to make decisions. As we have seen, it is a new strategy that allows complementing -in important ways- the information usually offered by the lawyer, who usually gave a specific recommendation on the option to choose. This set of experiences, alien but useful, allows the potential user of courts to participate more fully in deciding how to resolve

the initial conflict, reduces uncertainty and information asymmetry, allows him to make better-informed decisions and encourages him to play a more important role in the decision to litigate -or not- and in the steps to be taken at each stage of the process. At the same time, and through a variety of mechanisms, the market for professional legal services is becoming more competitive, improving the services offered to clients, reducing costs for clients and allowing greater access to the market for younger law firms. Thus, the new Big Data Ecosystem is increasing the quality of court and legal services, improving the access to justice and increasing the efficiency in the judicial sphere, which is definitely leading to an increase in social welfare.

Of course, we still need to be more aware that these technological changes, which will soon penetrate the judicial sphere, can be a double-edged sword. In addition to the dangers already known coming from their ecological impact and the invasion of privacy -blurred in the increasingly confusing distinction between private and nonprivate data-[1] and the risky temptation to mistake "correlation" for "causality" discussed earlier in this paper, the immersion in the Big Data Ecosystem introduces new challenges.

Although the results of an algorithm, such as those that emerge from Artificial Intelligence applied to Big Data, have the appearance of being totally rational and objective, we must bear in mind that its application may introduce new biases. For example, in crime prevention, even small crimes threaten to skew the analysis, because once this data flows into a predictive model, more police are drawn in those areas, and they are more likely to arrest more people. This may lead to the conclusion that crime has increased in this more closely monitored area, whereas in reality, it is the application of the algorithm that is introducing a bias into the new statistics obtained on crime rate, in addition to promoting discriminatory criteria[2] [3]. As Neff (2013) claims, "*Data are not neutral... Put simply, data is only data in the eye of the stakeholder*". This kind of biases can be extrapolated to any field, from the detection of new needs for judicial services to the location of potential clients of professional services or the detection of underperforming employees or departments in a law firm. If the Big Data algorithms are applied blindly, regardless of this feedback loop, and without expert surveillance by specialists, this will lead to the generation of distorted statistical data that will lead to the application of biased corrective or preventive measures and to wrong decisions in general. Not only will the potential benefits of Big Data be largely offset, but new problems will be unknowingly introduced, reducing the well-being of citizens and the corporate profits.

References

[1] Anderson, C. (2008). The End of Theory: The data deluge makes the scientific method obsolete. *Wired* 16 (July).

[1] See for example Mayer-Schönberger and Cukier (2013).

[2] A 2014 White House report from the office of President Obama underlined that Big Data leads to 'vexing issues (Big Data technologies can cause societal harms beyond damages to privacy, such as discrimination against individuals and groups)', Hilbert (2016).

[3] For a broad discussion on this issue see for example O'Neil (2016).

[2] Consejo General del Poder Judicial (1997). El Libro Blanco de la Justicia. Ed. Comares.
[3] Cox, M. and Ellsworth, D. (1997). Managing Big Data for Scientific Visualization. ACM SIGGRAPH '97 Course #4, Exploring Gigabyte Datasets in Real-Time: Algorithms, Data Management, and Time-Critical Design, ACM SIGGRAPH '97, August.
[4] Cumbley, Richard and Church, Peter (2013). Is "Big Data" Creepy? In: *Computer Law & Security Review* 29 (5). DOI: 10.1016/j.clsr.2013.07.007.
[5] Dakolias, M. (1999). Court Performance around the World. A Comparative Perspective. In: *World Bank Technical Paper* (430).
[6] Dimitrova-Grajzl, Valentina; Grajzl, Peter; Sustersic, Janez; and Zajc, Katarina (2012). Court Output, Judicial Staffing, and the Demand for Court Services. Evidence from Slovenian Courts of First Instance. In: *International Review of Law and Economics* 32 (1). DOI: 10.1016/j.irle.2011.12.006.
[7] Dunham, Ian M. (2015). Big Data. A Revolution That Will Transform How We Live, Work, and Think. In: *The AAG Review of Books* 3 (1). DOI: 10.1080/2325548X.2015.985533.
[8] Emerging Technology from the arXiv (2017). Wisdom of the Crowd Accurately Predicts Supreme Court Decisions. Available at: https://www.technologyreview.com/s/609852/wisdom-of-the-crowd-accurately-predicts-supreme-court-decisions/.
[9] Ford, Heather (2016). The Person in the (Big) Data: A Selection of Innovative Methods, Strategies and Perspectives for Social Research in the Age of (Big) Data. In: *Working Papers of the Communities & Culture Network* (8). Available at: http://eprints.whiterose.ac.uk/114826/1/person_in_the_big_data_report.pdf .
[10] Ginsberg, Jeremy; Mohebbi, Matthew H.; Patel, Rajan S.; Brammer, Lynnette; Smolinski, Mark S.; and Brilliant, Larry (2009). Detecting Influenza Epidemics Using Search Engine Query Data. In: *Nature* 457 (7232).. DOI: 10.1038/nature07634.
[11] Hilbert, Martin (2016). Big Data for Development: A Review of Promises and Challenges. In: *Development Policy Review* 34 (1).
[12] Johnson, S.; McMillan, J.; and Woodroof, C. (2002). Courts and Relational Contracts. In: *Journal of Law, Economics and Organisation* 18.
[13] Katz, Daniel Martin; Bommarito, Michael J.; and Blackman, Josh (2017). A General Approach for Predicting the Behaviour of the Supreme Court of the United States. In: *PloS one* 12 (4), e0174698. DOI: 10.1371/journal.pone.0174698.
[14] Landes, William M. (1971). An Economic Analysis of the Courts. In: *The Journal of Law and Economics* 14 (1). DOI: 10.1086/466704.
[15] Mayer-Schönberger, Viktor; and Cukier, Kenneth (2013). Big data. A Revolution that will Transform How We Live, Work, and Think. Boston: Houghton Mifflin Harcourt.

[16] Mora-Sanguinetti, Juan S.; and Garoupa, Nuno (2015). Do Lawyers Induce Litigation? Evidence from Spain, 2001–2010. In: *International Review of Law and Economics* 44. DOI: 10.1016/j.irle.2015.06.003.
[17] Mora-Sanguinetti, Juan S. (2016). Evidencia Reciente sobre los Efectos Económicos del Funcionamiento de la Justicia en España. In: *Boletin Económico del Banco de España* 34.
[18] Neff, Gina (2013): Why Big Data Won't Cure Us. In: *Big data* 1 (3). Available at: http://online.liebertpub.com/doi/full/10.1089/big.2013.0029.
[19] Nelson, Sharon D.; and Simek, John W. (2013). Big Data_ Big Pain or Big Gain for Lawyers. In: *Law Practice Magazine* 39-4.
[20] O'Neil, Cathy (2016). Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy. First edition. New York: Crown.
[21] Pastor, Santos (1993). ¡Ah de la Justicia!, Política Judicial y Economía: Civitas.
[22] Pastor Prieto, S. (2016). Análisis Económico de la Justicia y Reforma Judicial.: Tirant lo Blanch.
[23] Pastor Prieto, Santos; and Vargas Perez, Carmen (2001). El Coste de la Justicia, Datos y un Poco de Análisis. In: *Cuadernos de Derecho Judicial* (15).
[24] Peñaloza Figueroa, Juan Luis and Vargas Perez, Carmen (2004). ¿Qué Costes Económicos y Sociales Comporta la Ejecución de Sentencias Judiciales? *Cuadernos de Estudios Empresariales* 1(4).
[25] Posner, R. A. (1973). An Economic Approach to Legal Procedure and Judicial Administration. In: *Journal of Legal Studies 2*.
[26] PowerData (2014). 7 Definiciones de Big Data. https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/381767/7-definiciones-de-big-data.
[27] Prendergast, James D. (1977). The Use of Data Processing in Litigation. In: *Jurimetrics Journal* 17 (3).
[28] Priest, G. (1989). Private Litigants and the Court Congestion Problem. Boston University Law Review, 69. In: *Boston University Law Review* (69).
[29] Puyol Montero, Javier (2014). Una Aproximación al Big Data. In: *Revista de Derecho UNED* 14.
[30] Rosales Lopez, V. (2008). Economics of Court Performance. An Empirical Analysis. *European Journal of Law and Economics*, 25(3).
[31] Rosales-López, V. (2014). Demanda y Oferta de Justicia Civil en España. In: Panorama de las Administraciones Públicas: OECD Publishing.
[32] Rosales Lopez, V; and Jimenez-Rubio, D. (2016). Empirical Analysis of Civil Litigation Determinants: the Case of Spain. *European Journal of Law and Economics*. *1*(18).
[33] Shavell, Steven (1982). The Social Versus the Private Incentive to Bring Suit in a Costly Legal System. In: *Journal of Legal Studies* 9. Available at: http://www.law.harvard.edu/faculty/shavell/pdf/11_J_legal_stud_333.pdf
[34] Siegel, Eric (2016): Predictive analytics. The Power to Predict who will Click, Buy, Lie, or Die. Available at:

https://www.synapse.koreamed.org/Synapse/Data/PDFData/1088HIR/hir-19-63.pdf.

[35] Universidad Barcelona, AQR Lab and Cambra de Comerc de Barcelona (2018). Impacto Económico del Sistema de Ejecución de Sentences Judiciales y Propuestas de Mejora. Available at: https://www.google.es/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwiitZKC5_rYAhVGsxQKHXv9BYQQFggoMAA&url=https%3A%2F%2Fwww.icpb.es%2Ffotos2%2FEJECUCIONES_SENTENCIAS_ENERO_2018.pdf&usg=AOvVaw2kYIy1D4m0ZetuvF7toTOb.

[36] Vargas Perez, Carmen; and Peñaloza Figueroa, Juan Luis (2017). Big Data and the Demand for Court and Legal Services. In: *European Journal of Interdisciplinary Studies* 9 (1).

[37] Vargas Pérez, Carmen; and Peñaloza Figueroa, Juan Luis (2005). The Duration of Civil Cases. A Survival Analysis. *Perspectivas. Revista de Análisis en Economía y Comercio Exterior*. *1*(1), México. http://eprints.ucm.es/42178/).

[38] Voigt, Stefan (2016). Determinants of Judicial Efficiency. A survey. In: *Eur J Law Econ* 42 (2). DOI: 10.1007/s10657-016-9531-6.

[39] Wang, Tricia (2013). Big Data Needs Thick Data. In: *Ethnography Matters.* Available at: http://ethnographymatters.net/blog/2013/05/13/big-data-needs-thick-data/.

[40] Wistrich, Andrew J.; Rachlinski, Jeffrey (2013). How Lawyers' Intuitions Prolong Litigation. In: *Cornell Law Faculty Publications* (602).