

DATA-DRIVEN IS COOL: PERSPECTIVES, POSSIBILITIES, AND LIMITATIONS OF DATA-DRIVEN LAW APPLICATIONS

Lorenzo Antonini Itabaiana *
Letícia Camargos Ferraz **

Abstract: This paper analyzes the lessons of data-driven economics, investigating the prospects for applying the principles of Data Science on Law, or simply data-driven law. By understanding the possibilities that already exist, particularly in some Brazilian contexts, it addresses the limitations that permeate the ever-increasing interaction between Law and Data Science.

Keywords: law; Data Science; data-driven law.

INTRODUCTION

In a world increasingly driven by information and technology, the intersection between Data Science and Law emerges as a fertile ground of opportunity and innovation. The ability to extract valuable insights from massive data sets and apply them to the legal context is revolutionizing the way lawyers, judges and legal professionals' approach contemporary challenges. As we enter an era of legal complexity and the need for informed decisions, Data Science presents itself as a powerful ally, capable of revealing hidden patterns, predicting judicial outcomes, and providing a solid foundation for strategic decision-making.

In this article we will look at the evolution of the impact of data analysis on legal decision-making (data-driven law). From a deductive perspective, based on a literature review and applied to practical national and international experiences, we will try to answer the following question: how can law benefit from data science to be more data-driven? To this end, we will address four central topics: (i) how data-driven economics can contribute to law; (ii) what is data-driven law and what are its prospects for application; (iii) how Brazilian legislation approaches the data-driven law; (iv) what are the limitations and main challenges to be overcome for law to better absorb the principles of data science.

* Master of Laws candidate, Federal University of Minas Gerais, Brazil. Email: lorenzo.antonini@dcom.law / ORCID iD: <https://orcid.org/0009-0008-2966-9644>

** Bachelor of Laws from the Federal University of Minas Gerais. Corporate Consulting Lawyer, Brazil. Email: letqsf@gmail.com / ORCID iD: <https://orcid.org/0009-0007-3849-2185>

Initially, we will recap the application of techniques in areas other than law, such as finance and streaming services, with the aim of providing concrete examples of how technology has been used to make better decisions. These techniques and examples will be cited to introduce the potential of such applications in the field of law, demonstrating how other areas have benefited from this phenomenon, both qualitatively (improving services and products) and quantitatively (increasing profit margins).

Next, we will detail what data-driven law is, indicating practical applications of these concepts. We will cover everyday experiences that can be explored right now, such as those linked to the concept of Legal Ops, and the implementation of artificial intelligence in Brazilian Courts, moving on to more advanced possibilities with machine learning applications, up to propositions considered to be "futuristic", such as those that intend to use big data and machine learning to make law personalized¹.

It will be necessary to check whether Brazilian law supports these applications, or whether there is a legal ban on their development and use on a large scale. This analysis will allow the authors to identify possible limitations and challenges, to contribute to the debate surrounding the legal professional in the 21st century.

The intersection of this article with Data Science is aimed at reducing the anxiety generated by the lack of dialog between the "social sciences" and the "natural sciences"². Without pretending to exhaust the subject, this work seeks to be a bridge between the two worlds, helping to broaden the horizons of legal professionals in an increasingly technological world.

I. WHAT CAN DATA-DRIVEN ECONOMY TEACH US?

The application of mathematical techniques to business is not a phenomenon of the 21st century. However, the amount of personal data (big data), processed on a large scale by innovative tools such as machine learning, driven by high computer processing capacity, has opened up avenues for innovation in various sectors of society, uncovering hidden patterns previously unidentifiable to human reason. In this sense, there is a tendency for 21st century institutions to be data-driven in order to maximize their productivity³ and remain competitive.

Data Science can be defined as a set of fundamental principles that guide

¹ Omri Ben-Shahar; Ariel Porat. *Personalized Law: Different Rules for Different People*, 2021

² Charles Snow. *The Two Cultures*, 2013, p. 57.

³ Foster Provost; Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*, 2013, p. 10.

the extraction of knowledge from data⁴. This area of knowledge seeks to support decision-making by agents and institutions, making them more assertive by supporting subjective and unscrutinized factors such as intuition and common sense⁵. The advantages of this application could be seen through experiments carried out in the banking sector in the 1990s, which was considered an "early adopter"⁶ due to its early implementation of the principles of Data Science.

The Americans Richard Fairbanks and Nigel Morris realized that information technology had the potential to generate more sophisticated predictive models that could be applied to banking services, such as granting credit lines, analyzing payment rates, and even analyzing user churn⁷. To do this, however, it was necessary to obtain quality data in a minimally structured way. With this in mind, Fairbanks and Morris suggested offering various services at random to customers of the Signet bank in the state of Virginia in the United States. Although the initiative initially generated monetary loss for the bank, eventually the discovery of financial consumption patterns and decision making based on these patterns made the operation so profitable that the company was split up, creating Capital One, a major US bank.

Another recent example of the ingenious application of data in decision-making is Netflix. In 2011, the streaming service bought the original series *House of Cards*, which was set in the United Kingdom, for US\$100 million⁸. In 2013, the American version of the series was launched, with the participation of producer David Fincher and actor Kevin Spacey⁹. The choice of the two was not random but was based on an analysis of the consumption habits of Netflix customers. The evaluation, carried out by the then CCO (Chief Content Officer) Ted Sarandos, determined that customers who consumed the original *House of Cards* series also consumed films and series produced by David Fincher and starring Kevin Spacey, suggesting that they both take part in the American version. In 2013 the series was one of Netflix's most watched, and by mid-2021 the company's market value was valued at

⁴ Foster Provost; Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*, 2013, p. 2.

⁵ Ed Walters. *Data Analytics and the New Legal Services*. 2019, p. 257.

⁶ Foster Provost; Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*, 2013, p. 7.

⁷ Foster Provost; Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*, 2013, p. 10.

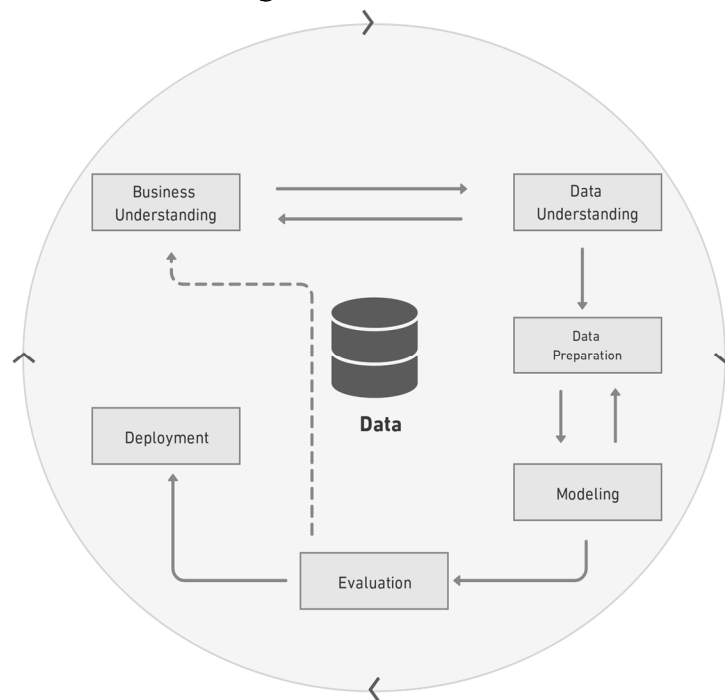
⁸ On said topic, cf. <https://variety.com/2011/tv/news/netflix-seals-house-of-cards-deal-1118034117/>. Accessed on: Jun 24th 2023.

⁹ For more information, check http://www.huffingtonpost.com/2013/01/24/netflix-show-house-of-cards_n_2545332.html. Accessed on: Jun 24th 2023.

approximately US\$267 billion¹⁰.

So far, we understood how valuable data is to an institution. However, getting to this point requires a great deal of effort from multidisciplinary teams who must be able to mine this data, a process known as data mining¹¹. There are various ways of carrying out such tasks. In this article, we will use the CRISP-DM (Cross Industry Standard Process for Data Mining)¹² method as a reference. According to this methodology, the data analysis cycle goes through six non-linear stages¹³, which can be carried out concurrently or not. Figure 1 illustrates the stages:

Figure 1 – CRISP DM



Source: Foster Provost; Tom Fawcett, p. 27.

¹⁰ <https://companiesmarketcap.com/netflix/marketcap/>. Accessed on Jun 25th 2023. This is not to say categorically that the analysis done for the House of Cards series in isolation was responsible for Netflix's success, but only that it is part of a broader data-driven decision strategy that has paid off for the company and its shareholders.

¹¹ Foster Provost; Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. 2013, p. 14.

¹² There are several possible models for carrying out data mining tasks, such as KDD (Knowledge Discovery in Database) and SEMMA (Sample, Explore, Modify, Model and Asses). For more information on this topic, see: Andres Fortino. *Data Mining and Predictive Analytics for Business Decisions: A Case Study Approach*, 2023

¹³ Foster Provost; Tom Fawcett. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. 2013, p. 27.

Although this paper is not intended to exhaust and detail the stages of the CRISP-DM cycle, it is important that they are known so that the principles can be applied by legal professionals. Here we highlight the business understanding phase, through which the objectives and requirements of the project are analyzed and then converted into a problem and a plan to achieve these objectives¹⁴. We also point to data understanding as an essential stage in the legal professional's knowledge. In this phase, data is scanned to understand what is available, what the limitations are, and what needs to be bought or produced. The other phases involve data preparation, modeling, evaluation, and deployment. But what does this have to do with law?

In general, decisions made by jurists are based on their knowledge of doctrine, legislation, and case law. They are therefore based on experience and subjective analysis, for example, checking the political profile of a judge to predict the chances of their client's claim being accepted or not. To check the productivity of his sector, a senior lawyer would use his subjective perception of the other lawyers, analyzing the quality of the documents produced and the assertiveness of the theses proposed. In none of the examples cited is there any support from data to define a procedural strategy or the allocation of human resources within a law firm.

This is where data-driven law seeks to act.

II. PERSPECTIVES AND POSSIBILITIES FOR DATA-DRIVEN LAW

Data-driven law can be defined as the application of a set of techniques from data science and statistics to guide or automate decision-making¹⁵ in the legal sphere, making it more productive¹⁶, precise and efficient for legal operators, reducing the impact of cognitive biases¹⁷ and heuristic cognitive¹⁸

¹⁴ Andres Fortino. *Data Mining and Predictive Analytics for Business Decisions: A Case Study Approach*. 2023, p. 41.

¹⁵ The term "decisions" is used in this paper with a broad concept, referring to the cognitive process of reaching a conclusion on how to act based on an analysis of the possibilities, data and regulations presented by the concrete situation. As an example, "decision" can refer to the adoption of procedures legally provided for by the Public Administration to promote certain public policies, or the deliberation on how human resources will be allocated in a public judicial body, or how the division of the different procedural classes will be carried out in a district court, or even the judgment of a lawsuit. Other more concrete examples will be cited throughout the text.

¹⁶ Ed Walters. *Data Analytics and the New Legal Services*, 252

¹⁷ In short, cognitive biases are misrepresentations of judgments due to the insertion of irrationality resulting from intuition or previous experiences. For further information on this subject, see: Nathália Medeiros. *Use of Artificial Intelligence in Judicial Decision-Making: an analysis from the perspective of the normative theory of participation*, 2019

¹⁸ In the epistemological field, "heuristics" refers to intuitive, unsubstantiated

processes in the legal field. It can be applied to a wide range of legal activities, so the following is a selection of the possibilities.

A. Jurimetrics

Perhaps the oldest application of data-driven principles to law is Jurimetrics. It has its origins in the dissatisfaction of some authors - such as Lee Loevinger - with the stagnation of legal reasoning to the detriment of progress in other areas. In various fields of knowledge, the adoption of scientific methods has made it possible for primitive knowledge resulting from speculation to be overtaken by scientific knowledge resulting from investigation¹⁹. The law, on the other hand, would not have achieved such significant progress.

Jurimetrics is based on the principle that legal problems will be solved by investigating the law through statistics²⁰ rather than introspection. In other words, it is necessary a broad scientific investigation of legal problems²¹, i.e. the adoption of scientific methods to identify and solve problems in law²². Jurimetrics does not attempt to tackle broad, general and age-old problems, such as those restricted to the philosophy of law, but specific, practical, real problems with testable solutions²³. These problems can be investigated, partially solved, tested and then investigated again by reinserting data as new situations arise: this is the incorporation of the scientific method into legal practice.

In this scenario, in addition to the broad proposal to change the legal epistemological paradigm through Jurimetrics, Loevinger also mapped out some concrete tools that could be devised to fully implement the principles of Jurimetrics in the reasons for decisions. These tools include: (i) quantitative analysis of judicial behavior; (ii) application of communication and information theories to legal language; (iii) use of mathematical logic in law; (iv) retrieval of legal data for electronic media; and (v) formation of legal predictability calculations²⁴.

judgments based on imprecise correlations made by humans, using simplification or mental automatism to make decisions. For further information on this subject, see: Nathália Medeiros. Use of Artificial Intelligence in Judicial Decision-Making: an analysis from the perspective of the normative theory of participation, 2019

¹⁹ Lee Loevinger. Jurimetrics: The Next Step Forward, 467

²⁰ Marcelo Guedes Nunes. *Jurimetria: Como a estatística pode reinventar o direito*, 112

²¹ Lee Loevinger. Jurimetrics: The Next Step Forward, 467

²² Lee Loevinger, Lee. *Jurimetrics: The Methodology of Legal Inquiry. Law and Contemporary Problems*, 08

²³ Lee Loevinger. Jurimetrics: The Next Step Forward, 467

²⁴ Lee Loevinger, Lee. *Jurimetrics: The Methodology of Legal Inquiry. Law and*

In the case of Brazil, several of the tools proposed by Loevinger have already been widely adopted. An example of item (i) is the annual statistical survey of the performance of the Judiciary by the Brazilian National Council of Justice (“*Conselho Nacional de Justiça – CNJ*”²⁵) and the diagnoses carried out by the Brazilian Association of Jurimetrics (“*Associação Brasileira de Jurimetria – ABJ*”²⁶) based on the data collected on a given matter.

As an application of item (ii), we can mention the theories of legal design and visual law, used in the private and public sectors, which aim to transmit information in a clearer and more objective way not only to legal operators, but also to the general population²⁷. As an application of item (iv), mention should be made of the case law databases that are present in practically all Brazilian courts, either via the electronic judicial process (“*Processo Judicial Eletrônico – Pje*”²⁸) or via digitization of physical cases and making them available online, which allow users to search for specific matters using filters. Finally, items (iii) and (v) have been developed especially by law firms and legal departments, mainly by creating predictive models to define the mathematical possibility of a particular appeal being granted, taking into account the matter, the legal grounds and the understandings of each court. The legal predictability made possible by Jurimetrics can also guide departments in their decision-making and, why not say it, in their own management²⁹.

B. Legal Departments and Legal Ops

Legal departments can already benefit from data-driven principles,

Contemporary Problems, 08

²⁵ Conselho Nacional de Justiça. *Justiça em Números*. 2022. Available at <<https://www.cnj.jus.br/wp-content/uploads/2022/09/justica-em-numeros-2022-1.pdf>>. Accessed on Jun 20, 2023.

²⁶ Associação Brasileira de Jurimetria. *Pesquisas*. Available at: <<http://abj.org.br/pesquisas/>>. Accessed on Jun 28, 2023

²⁷ The subject of legal design is beyond the scope of this article and will not be dealt with in detail. However, we recommend: Colette Brunschwig. *Visual Law and Legal Design: Questions and Tentative Answers*, 2021.

²⁸ A more in-depth study on this topic can be found in Rômulo Valentini. *Electronic Court Systems and Mechanical Trials: Developing Standards for a "Legal Turing Test"*. In: Leonardo Parentoni; Renato Cardoso. *Law, Technology and Innovation II: Insights from Artificial Intelligence and the Law*. Editora Expert: Belo Horizonte, 2021.

²⁹ A classic example of a predictive model is that used by the American company Lex Machina, which predicts the behavior of courts, judges and lawyers based on the analysis of historical data. To find out more, see: <<https://lexmachina.com/>>. Accessed on June 28th 2023.

especially those that apply the principles of Legal Ops³⁰. The 2022 Legal Department Operations Index, a survey conducted by the Thomson Reuters Institute and the Legal Value Network³¹, sought to understand the main trends and priorities of these agents by analyzing information provided by more than 107 legal departments.

In this context, it was reported that the adoption of technology to simplify workflow is a high priority for almost 70% of legal departments. However, when it comes to adopting Business Intelligence (BI) to inform decision-making, this figure drops to 41%³². In a similar vein, a survey by Deloitte found that 45%³³ of departments have some kind of BI analysis based on central technology. Another report, presented by the Corporate Legal Operations Consortium (CLOC) showed that, among the sample of legal departments, 56% adopt Jurimetrics, metrics and dashboards, while only 22% adopt Data Science, including the use of artificial intelligence³⁴.

The use of BI by legal departments has the potential to add value in the generation of Key Performance Indicators (KPIs), analysis of metrics and comparisons with other departments or firms (benchmarking)³⁵. To do this, the data identified as necessary in the Business Understanding phase must be available or accessible. To this end, it is highly recommended that platforms be used to collect the necessary information, such as activities carried out, time spent, story points³⁶, and other data that the manager deems necessary. There are various tools that can support departments in carrying out these

³⁰ CLOC (Corporate Legal Operations Consortium) defines Legal Ops as: "a set of business processes, activities and professionals that enable legal departments to serve their clients more effectively by applying business and technical practices to the delivery of legal services. Legal ops provides the strategic planning, financial management, project management and technological expertise that allow legal professionals to focus on providing legal advice." Available at <<https://cloc.org/what-is-legal-operations/>>. Accessed on June 25, 2023.

³¹ Thomson Reuters Institute and the Legal Value Network. *2022 Legal Department Operations Index. Seeking stability amid uncertainty*. 2022.

³² Thomson Reuters Institute and the Legal Value Network, 6.

³³ Deloitte. *2022 State of Legal Operations Survey. Taking steps to modernize Legal Operations*. 2022.

³⁴ CLOC. *2021 State of the Industry Survey. Results and Analysis*. 2021, p. 13. In the document, Data Science was treated as something distinct from Jurimetrics, metrics and dashboards, so the authors chose to maintain the original CLOC understanding.

³⁵ Thomson Reuters Institute and the Legal Value Network. *2021 Legal Department Operations Index. The risk of being left behind*, 17

³⁶ A story point is a number from the Fibonacci sequence assigned to a task to be carried out by a person (Layton et al, 80), and can take into account factors such as complexity and difficulty of execution. For more information on the Scrum method, we suggest the humorous book: Mark Layton, Steven Ostermiller, Dean Kynaston. *Scrum for dummies*, 2023.

tasks. Examples include Clickup, Monday, Trello and Clokify³⁷.

CLOC also points to BI as a core competence for legal departments³⁸. The purpose is for the organization to be managed based on data, not just intuition, discovering hidden trends, increasing efficiency and focusing the team on clear and measurable results. To achieve this result, legal departments need to determine the appropriate data for analysis, implementing metrics and dashboards that are easily understood in order to detect patterns. All of this, if done well, has great potential to reduce costs and improve team performance.

Nowadays, being a lawyer is much more than knowing laws, doctrines and jurisprudence. It involves acting strategically and connected to various areas of the business under analysis, such as operations, risks, talent and technology³⁹, which would require a greater ability to understand the data generated. Without such indicators, managers are doomed to rely on their intuition, which, although still extremely relevant, should not be taken as an isolated factor in decision-making. The benefits highlighted in this section, however, can also be enjoyed by courts and other public bodies.

*C. Statistics panel and Brazilian National Council of Justice (CNJ)
report: Justice in Numbers*

New technologies are essential tools for court management⁴⁰. The technologies provided by big data have enabled a revolution in the way the administration of justice is managed, which lacked interconnected data to map performance and identify problems⁴¹. In the case of the "Brazilian Justice in Numbers Report", data mining, combined with statistical models, made it possible to identify data and trends related to judicial management⁴².

Every year, the Brazilian National Council of Justice (CNJ) publishes an extensive report which statistically examines, among other topics, the effectiveness of judicial protection, the performance and productivity indicators of judicial management, the expenses incurred, conciliation rates and the length of time cases have been pending. Periodically, the Judiciary

³⁷ The platforms have been listed as examples and are not recommendations for use. The authors are not responsible for the use and effectiveness of these platforms, nor do they have any links with the companies that develop the applications.

³⁸ <https://cloc.org/what-is-legal-operations/>. Accessed on Jun 25, 2023.

³⁹ Deloitte. *2022 State of Legal Operations Survey. Taking steps to modernize Legal Operations*. 2022.

⁴⁰ Salomão, Luis Felipe. *Inteligência artificial: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário Brasileiro*, 12

⁴¹ Salomão, Luis Felipe. *Inteligência artificial: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário Brasileiro*, 12

⁴² Conselho Nacional de Justiça. *Justiça em Números*, 2022

Statistics Panel⁴³, created by Resolution n. 333/2020⁴⁴, is also updated with data related to procedural management. With these documents, it is possible to see where the Judiciary's current problems lie and map out strategies to better remedy them. On the other hand, it is impossible - or at least unlikely - to solve problems that are not even known about⁴⁵.

D. Justice 4.0 and Artificial Intelligence in the Brazilian Judiciary

CNJ Resolution n. 332/2020 made it possible to develop AI tools to help the judiciary make decisions. According to data provided by the CNJ, what is motivating the use of AI tools by the Courts is the improvement in the quality of services, increased productivity, the possibility of innovation and cost reduction⁴⁶.

The tools currently being developed focus on⁴⁷ (i) automated procedural classification based on the application of natural language processing techniques, such as argument mining⁴⁸; (ii) better organization and availability of qualified precedents handed down by higher courts, with a view to improving case law databases⁴⁹; (iii) the grouping of cases with similarities in content, semantics or case law⁵⁰; (iv) the automatic screening

⁴³ Available at: < <https://painel-estatistica.stg.cloud.cnj.jus.br/estatisticas.html>>. Accessed on: 28 jun. 2023

⁴⁴ Resolution n. 333/2020 determined the inclusion of the "Statistics" field on the websites of the bodies of the Judiciary with a view to gathering open data, Business Intelligence Panels and Statistical Reports relating to the core activity of the Judiciary. In: Conselho Nacional de Justiça *Justiça em Números, 2022*

⁴⁵ For example, in the Report covering data for 2021, it was possible to verify that, of the 77 (seventy-seven) million cases in progress, 53% are in the execution phase, which corresponds to a congestion rate of 90% on average. With this data in mind, it becomes feasible to study alternatives for improving judicial effectiveness in enforcement proceedings. In: Conselho Nacional de Justiça *Justiça em Números, 164-180*

⁴⁶ Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁴⁷ Salomão, Luis Felipe. *Inteligência artificial: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário Brasileiro, 27-76*

⁴⁸ This is the case with Miner Jurs, used by the TJTO (Tribunal of Justice of the State of Tocantins), which guarantees greater reliability in the classification of court cases when the initial petition is registered in relation to the subject, using natural language processing techniques to extract information. At the Federal Regional Court of the 4th Region (TRF-4), the system makes a classification suggestion when the case is filed. In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁴⁹ This is the case with INDIA - an artificial intelligence-based indexer of court documents, used by the Pará Court of Justice to enable the migration of proceedings from physical to digital media, and HORUS, used by the Federal District Court of Justice (TJDF). In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁵⁰ This is the case of Victor, used by the Brazilian Supreme Court, which identifies appeals that fall under one of the most recurrent themes of general repercussion and groups them by similarity. In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder*

of cases⁵¹; (v) voice transcription for texts with content⁵²; (vi) the semi-automatic generation of pleadings⁵³ and (vii) the use of language processing models to answer simple questions⁵⁴. Other ambitious models are being developed in an attempt to automate repetitive tasks, such as sending and attaching documents like summonses, subpoenas and certificates⁵⁵.

E. Argument mining

A bolder application of data-driven law is argument mining. The proposal is to apply natural processing language models to law, an artificial intelligence⁵⁶ technique whose aim is to automatically extract arguments from generic texts, allowing data to be structured for computational models⁵⁷. Based on the basic assumption that argumentation is made up of arguments, and that arguments are made up of premises and conclusions⁵⁸, it would be possible to apply this model to extract arguments from complex legal texts, such as pleadings. Detecting arguments in a text is similar to a binary classification task, as it aims to distribute sentences between "argumentative" and "non-argumentative", while also making it possible to classify the arguments identified⁵⁹. Figure 2 can help the reader understand the proposal:

Judiciário em 2022

⁵¹ This is the case of Bem te Vi, an initiative of the Superior Labor Court (TST) that screens cases. In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁵² This is the case of Grafo, created by the Rio Grande do Sul Court of Justice (TJRS). In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁵³ This is the case with JANUS, a tool designed by the Regional Electoral Court (TRE) of the State of Bahia, which is also in use in the States of Piauí, Pernambuco and Maranhão. JANUS identifies the candidates' information and the opinions of the Public Prosecutor's Office on whether they should be accepted or rejected. It then issues a standard monocratic decision in the PJE in cases of automatic approval of candidacies. In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁵⁴ This is the case of the chatbot for the public at the Superior Electoral Court (TSE). In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

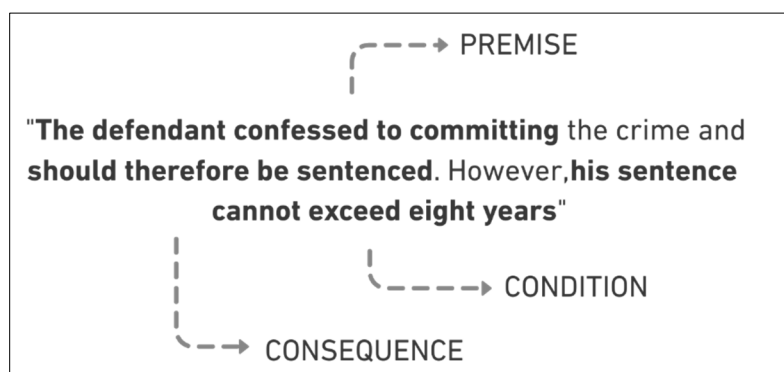
⁵⁵ Models are being developed by the Tribunal of Justiça of the State of Bahia. In: Conselho Nacional de Justiça. *Resultado Pesquisa IA no Poder Judiciário em 2022*

⁵⁶ In this paper we do not intend to go into the conceptualization of Artificial Intelligence (AI). For a more in-depth conceptualization of AI, see: Leonardo Parentoni. What should we reasonably expect from artificial intelligence? In: *Il Diritto Degli Affari*, 2022; and Alan Turing. Computing machinery and intelligence. In *Mind*, n. 49, p. 433-460

⁵⁷ Marco Lippi; Paolo Torroni. Argumentation Mining: State of the Art and Emerging Trends, 2

⁵⁸ Raquel Mochales; Marie Francine Moens. *Argumentation mining*, 5

⁵⁹ Raquel Mochales; Marie Francine Moens. *Argumentation mining*, 5

Figure 2 – *Argument mining*

Source: drawn up by the authors.

Thus, in order to detect textual elements in law, the algorithms would identify expressions common to legal texts, such as "for all the above reasons", "in light of principle X", "however", "in addition", among others. A study by Mochales and Moens obtained an accuracy of 60% in detecting argumentative structures in texts from the European Court of Human Rights⁶⁰, proving that it is possible to identify written arguments and structure them using artificial intelligence methods.

It is true that these applications need to be explored in greater depth, but in principle it is assumed that the technique could be used, for example, to help lawyers search for case law based on the processing of the initial petition. Another possible application would be to infer a certain position of a judge or a Chamber from the automatic processing of past judgments.

F. Exploring the limits of technology: personalized law

Law is made up of laws, most of which tend to be general and abstract. As Norberto Bobbio teaches us, "'The law is the same for everyone' is undoubtedly the generality of the rules, that is, the fact that the rules are not intended for this or that citizen, but for all citizens or for an abstract type of operator in social life"⁶¹. From this premise, we can draw the fact that, in general, laws are not aimed at a specific individual, and do not take the internal and external characteristics of the recipient into account when applying them.

The Professors Omri Ben-Shahar and Ariel Porat have proposed a thesis that challenges this premise. By analyzing large volumes of data (big data)

⁶⁰ Raquel Mochales; Marie Francine Moens. *Argumentation mining*, 18

⁶¹ Norberto Bobbio. *Teoria Geral do Direito*, 176

and using artificial intelligence models, the authors intend to make the law personalized. A law for every citizen. This is called Personalized Law. For these researchers, "personalized law is precision law that is characterized by two main features: individualization and information selected by machines"⁶². As an example provided by the authors, in order to verify whether a driver's license should be suspended for alcohol intoxication, in addition to external factors, intrinsic factors should also be analyzed, such as the subject's ability to absorb alcohol. Thus, for each subject there would be a specific rule, personalized according to their characteristics, which would replace the jurisdictional activity of applying the rules according to the facts (subsumption of the rule to the fact) by an artificial intelligence, responsible for processing the facts and identifying the individual attributes that would lead to a certain decision, which could be automated or not.

Among the greatest advantages identified is precision, considering that by defining a specific rule for each individual, the objective of the law would be achieved more efficiently. Like a tailor-made suit, a customizable law would provide more precise commands on the conduct to be adopted by each individual. Thus, if the law aims to prevent accidents due to alcohol intoxication, it would not matter whether or not someone ingested alcohol, but rather their ability to process it, and therefore to cause - or not - traffic accidents. This would allow different conducts for the same purpose, which, in principle, would be beneficial for individual rights and freedoms. However, further studies are needed to delve deeper into the risks involved in applying this theory, which is not the aim of this paper.

G. Limitations: Brazilian legislation and practical consequences

So far, it has been possible to understand that the data-driven mentality is a power to be exploited in the field of law. Possible applications have been detailed, as well as ongoing or consolidated experiences. However, we have not neglected to note that there are limitations and challenges to be overcome for an ethical and legitimate application of data-driven principles. Without presuming to exhaust them all, in this first topic we will analyze legislative clippings to determine whether there would be legal openness to data-driven law, as well as observing possible ethical impacts resulting from the use of a data-driven logic, such as discrimination, preservation of privacy, protection of personal data, among others. We will then investigate the impacts of data-driven law in other fields, such as ethics and information security.

⁶² Omri Ben-Shahar; Ariel Porat. *Personalized Law: Different Rules for Different People*, 19.

III. DOES THE BRAZILIAN LAW SUPPORT DATA-DRIVEN LAW?

Brazil progressively builds and enhances databases⁶³ for several different purposes⁶⁴, prioritizing, on doing so, transparency⁶⁵ and structured data dissemination⁶⁶. Furthermore, it is a guideline of the Public Administration to enable interoperability among databases⁶⁷ and public services provided by the government to promote the exchange of information and procedural efficiency. Hence, since these databases are being fulfilled with historical data and made publicly available to the population in most cases, is safe to assume that there is a significant space for data mining initiatives for decision-making procedures.

The Law of Introduction to the Rules of Brazilian Law (*Lei de Introdução às Normas do Direito Brasileiro* or “LINDB”) was amended in 2018 to oblige that decisions must not be based on abstract legal values without considering the practical consequences of the decision⁶⁸. Consequently, this legal meta-norm demonstrates a greater concern for the rationale behind decisions⁶⁹, curbing the use of abstract parameters in reasoning procedures.⁷⁰

Subsequently, Brazilian Decree No. 9.930/2020⁷¹, which regulated changes to the LINDB, clarified that decisions should exhibit consistency between norms and the facts that underpin the decision. Moreover, the decision's rationale, in addition to legal reasoning, jurisprudence, and doctrine, can also delve into technical notes, opinions, and other relevant information related to the case, providing the necessary support for justifying the judgment. The primary objective is to elevate the level of legal certainty

⁶³ As an example, we mention ‘Redesim’, which is a platform created in 2007 that enables data integration regarding legal entities for all public bodies (such as states, municipalities, State Board of Trade, Federal Revenue and the State Treasury Department) that adhere to the system. As a result, all legal entities are benefited to a single entry of registration data and documents to constitute or change the constitutive data of legal entities, so that the constitution or alteration is processed faster and automatically.

⁶⁴ Art. 24, VI e VII of the Brazilian General Internet Regulation (“*Marco Civil da Internet*”).

⁶⁵ As an example, see the database of the Brazilian Institute of Intellectual Property (INPI). Available at: <https://busca.inpi.gov.br/pePI/servlet/LoginController?action=login>. Accessed on: 28 jun. 2023.

⁶⁶ It is relevant to note that the Brazilian’s federalism contributes to the formation of a reliable and centralized database.

⁶⁷ Article. 24, III and IV of *Marco Civil da Internet*.

⁶⁸ Article 20 e 21 of LINDB (Decree no. 4.657).

⁶⁹ Brazilian Federal Constitution, Article. 93, IX.

⁷⁰ José Luiz de Moura Faleiros Júnior; Alexandre Walmott Borges; Alex Cabello Ayzama. *Reflexões sobre o Consequencialismo Jurídico na Lei da Segurança para Inovação Pública (Lei n.º 13.655/2018)*, 112.

⁷¹ Brazil. Decree No. 9.930/2020.

within the legal system.⁷²

Therefore, based on this preliminary study, it is safe to say that there is a general authorization or, even better, a directive from the Brazilian legal system to incorporate data in administrative, regulatory, or judicial decisions. This observation opens the door for the exploration of existing databases in Brazil to guide legal practitioners. Several legislations seem to adhere to this logic, defining to a certain extent the use of data for various purposes, as we analyze in the following topics.

A. Brazilian General Internet Regulation

The Brazilian General Internet Regulation (*also known as Marco Civil da Internet or “MCI”*)⁷³ established the following guidelines to be observed by the Public Administration: (i) the rationalization and technological interoperability of electronic government services to enable information exchange and procedural efficiency (Article 24, III); (ii) the publicity and dissemination of public data and information in an open and structured framework (Article 24, VI); e a (iii) the optimization of infrastructure and promotion of the establishment of data storage, management, and dissemination centers within the country, fostering technical quality, innovation, and the spread of internet applications (Article 24, VII). Furthermore, the MCI stipulated that internet applications by public entities should strive for compatibility with both human reading and automated processing of information (Article 25, III). Lastly, the MCI also prioritizes the protection of privacy and personal data (Article 3, II and III), emphasizing that entities will be held accountable according to their activities (Article 3, VI), in accordance with the law.

Hence, it is evident that the efforts of the Public Administration aim to encourage the creation of interoperable and structured databases, which are essential for processing such information using data-driven law techniques. In an incipient framework, the MCI has paved the way for the possibility of automations in decision-making by public entities, and it has partially regulated the protection of personal data – that is, until the enactment of the Brazilian General Data Protection Law, mentioned below.

⁷² José Luiz de Moura Faleiros Júnior; Alexandre Walmott Borges; Alex Cabello Ayzama. *Reflexões sobre o Consequencialismo Jurídico na Lei da Segurança para Inovação Pública (Lei n.º 13.655/2018)*. Revista Jurídica da Universidade do Sul de Santa Catarina, p. 114.

⁷³ Brazil. Law No. 12.965. Marco Civil da Internet. Available at: <https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm>. Accessed on: 27 jun. 2023.

B. Brazilian General Data Protection Regulation

The Brazilian General Data Protection Regulation (*Lei Geral de Proteção de Dados “LGPD”*) has significantly contributed to establishing a general legal framework for the processing of personal data. By setting forth foundations⁷⁴, principles⁷⁵, and outlining the legal possibilities that authorize data processing, the LGPD has provided greater legal certainty for large-scale data processing, including automated decision-making. An important advancement in this legislation is ensuring that the regime of personal data protection, a fundamental right⁷⁶, must be compatible with economic, technological, and innovative developments⁷⁷.

Another significant contribution was the limitation of data processing by the Public Administration and public databases. Because of LGPD, public sector initiatives applying data-driven principles to personal data must align with the public interest, aiming to execute legal competencies or fulfill legal responsibilities of public services.

In a broader sense, it's understood that any data-driven decision-making involving personal data must be embedded within a governance principle. This governance framework should establish organizational conditions, operating procedures, mechanisms for addressing complaints and petitions from data subjects, security standards, technical specifications, and specific obligations for all stakeholders involved in data processing⁷⁸. Therefore, from the perspective of the LGPD, it can be assured that there is a legal framework for large-scale data processing in Brazil.

C. Economic Freedom Act and Regulatory Impact Analysis

The Law of Economic Freedom (*Lei da Liberdade Econômica “LLE”*⁷⁹) introduced data analysis by the Legislative branch by stipulating the requirement for regulatory impact assessment before making changes to normative acts of general interest. As per the LLE, the regulatory impact report must contain information and data regarding the potential effects of the normative act to assess the reasonability of its economic impact.⁸⁰

Subsequently, the Regulatory Impact Analysis (AIR) was further

⁷⁴ Article. 2º, LGPD.

⁷⁵ Article. 6º, LGPD.

⁷⁶ Article. 5º, LXXIX of Brazilian Federal Constitution.

⁷⁷ Leonardo Parentoni. *Compartilhamento de Dados Pessoais e a Figura do Controlador (Personal Data Sharing and the Role of the Data Controller)*. 2021, p. 3.

⁷⁸ Article. 50, *caput*, LGPD.

⁷⁹ Art. 5º LLE.

⁸⁰Subsequently, the need for regulatory impact analysis was also ratified for the acts of regulatory agencies.

regulated through Decree No. 10.411/2020, which states that the regulatory impact report, equipped with the obtained results⁸¹, should support decision-making⁸². To account the reasonability of the economic impact, entities must adopt specific methodologies, such as cost-benefit analysis or risk analysis.⁸³ The Decree also mandates that agencies should employ efficient strategies for data collection and processing⁸⁴ to enable the development of quantitative analyses.

Hence, it becomes evident that decisions to modify regulations must be substantiated by data, not solely by abstract legal principles or formal justifications. In other words, the impact analysis appears to necessitate the use of data-driven law through an examination of the problem to be addressed and the potential material consequences of the envisioned normative act.

D. Digital Government Law

The Digital Government Law (*Lei do Governo Digital* “LGD”), applicable to the administration entities of all federative entities, encompasses principles and guidelines including, among others,, (i) integrated cooperation among entities involved in the provision and oversight of public services, with the secure sharing of personal data (Article 3, VII) (ii) system interoperability and open data (Article 3, XIV)⁸⁵; e (iii) utilization of data for policy formulation, scientific research, business generation programs, and social control (Article 3, XXIII). The law also mandates that entities responsible for digital public service provision must manage their public policies based on data and evidence through the application of data intelligence on digital platforms.⁸⁶

The LGD also enables the establishment of Innovation Laboratories, which, among other directives, support data-driven policies and evidence-based decision-making to inform decision-making and enhance public administration (Article 45, VIII). Furthermore, it stipulates the implementation of governance mechanisms, instances, and practices that

⁸¹ Article. 6º, Decree No. 10.411/2020.

⁸² Article 2º, I, Decree No. 10.411/2020.

⁸³ Article 7º, Decree No. 10.411/2020.

⁸⁴ Article 17, Decree No. 10.411/2020. Here, we make a simple criticism: the use of the words “collection” and “treatment” seems inappropriate, as the LGPD provides that collection is a species of the genus treatment (art. 5, X of the LGPD).

⁸⁵ The objective of interoperability is to improve the management of public policies, increase the reliability of records, enable the creation of unified means of identifying citizens for the provision of public services, facilitate the interoperability of data between government agencies and carry out the processing of database information (Digital Government Law, article 39).

⁸⁶ Article. 24, VII of LGD.

include instruments to promote evidence-based decision-making (Article 47, III).

Clearly, this legislative framework aims to facilitate the convergence of technology with the public sector to enable improved services and optimize government functions⁸⁷. As such, the LGD also accommodates data-driven law techniques, requiring that the public administration base decisions on verifiable evidence derived from data intelligence.

*E. CNJ Resolution No. 332/2020:
the Artificial Intelligence forerunner within law courts*

The most recent experiences in the development and application of artificial intelligence in Brazilian courts are supported by CNJ Resolution No. 332/2020. In the preambular clauses of the regulation, it is expressly acknowledged that AI applications can contribute to the agility and coherence of the decision-making process.

According to the Resolution, judicial decisions supported by artificial intelligence models must uphold principles such as equality, non-discrimination, plurality, and solidarity, and also assist in fair judgment and to minimize judgment errors resulting from biases⁸⁸. For its application, it is also necessary that the models could be audited and provide satisfactory explanations in decision-making processes. In fact, the explanation of the steps leading to the outcome is a predominant criterion for adopting AI model in a judicial framework.⁸⁹

Resolution No. 332/2020 was issued based on the European Charter of Ethics for the Use of Artificial Intelligence in Judicial Systems. The European document established principles including respect for fundamental rights⁹⁰, non-discrimination⁹¹, technological quality and safety⁹², transparency,

⁸⁷ José. Luiz de Moura Faleiros Júnior. *Inovação, desenvolvimento e a Lei do Governo Digital (Lei nº 14.129/2021): perspectivas quanto às estruturas de governo como plataforma*. Inova Jur, Belo Horizonte, v. 1, n. 1, p. B1-B29, jan./jun. 2022, p. 08.

⁸⁸ Article 7º, CNJ Resolution 332/2020.

⁸⁹ Article 19, CNJ Resolution 332/2020.

⁹⁰ “When artificial intelligence tools are used to resolve a dispute or as a tool to assist in judicial decision-making or to give guidance to the public, it is essential to ensure that they do not undermine the guarantees of the right of access to the judge and the right to a fair trial (equality of arms and respect for the adversarial process).” In: European Commission for the Efficiency of Justice (CEPEJ). *European Ethical Charter on the use of Artificial Intelligence in Judicial Systems and Their Environment*. Strasbourg, 2018, p. 03.

⁹¹ “Given the ability of these processing methods to reveal existing discrimination, through grouping or classifying data relating to individuals or groups of individuals, public and private stakeholders must ensure that the methods do not reproduce or aggravate such discrimination and that they do not lead to deterministic analyses or uses.” In: *Ibidem*, p. 04.

⁹² “Data based on judicial decisions that is entered into a software which implements

impartiality, equity⁹³, and user control⁹⁴.

As a result of Resolution No. 332/2020, 111 projects were conceived by 88 Brazilian courts and, of that number, 63 tools are already in use⁹⁵. Of these, 98 have text analysis capabilities, and 49 have data organization features. Furthermore, there are also initiatives aiming to optimize processes and automate workflows, image recognition, speech analysis, and facial recognition. The progress made by the CNJ is significant as artificial intelligence can also be applied to data mining, particularly during the "Modeling" phase.

F. Reflections regarding the impacts of data-driven law

The field of Law is progressively embracing a data-driven concept through the utilization, by legal departments, law firms, and the public sector, of data-driven law techniques. Brazilian legislation not only allows for the use of data for decision-making but also seems to encourage such application.

On regards of this topic, without any intention of fully exhausting such a promising and ongoing subject, the authors propose some reflections aiming to identify and discuss the technical and practical boundaries of data-driven law. Therefore, we understand that each of the points identified in this topic should be browsed into in specific works, so the reflections presented here are merely embryonic given the magnitude of the challenge that scholars still

a machine learning algorithm should come from certified sources and should not be modified until they have been used by the learning mechanism. The whole process must therefore be traceable to ensure that no modification has occurred to alter the content or meaning of the decision being processed". In: *Ibidem*, p. 05.

⁹³ "A balance must be struck between the intellectual property of certain processing methods and the need for transparency (access to the design process), impartiality (absence of bias), fairness and intellectual integrity (prioritising the interests of justice) when tools are used that may have legal consequences or may significantly affect people's lives. It should be made clear that these measures apply to the whole design and operating chain as the selection process and the quality and organisation of data directly influence the learning phase". In: *Ibidem*, p. 05.

⁹⁴ "User autonomy must be increased and not restricted through the use of artificial intelligence tools and services The user must be informed in clear and understandable language whether or not the solutions offered by the artificial intelligence tools are binding, of the different options available, and that s/he has the right to legal advice and the right to access a court. S/he must also be clearly informed of any prior processing of a case by artificial intelligence before or during a judicial process and have the right to object, so that his/her case can be heard directly by a court within the meaning of Article 6 of the ECHR." In: *Ibidem*, p. 06.

⁹⁵ <https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-BR&theme=IA_PJ&opt=ctxmenu,cursel&select=language,BR>. Accessed on 20 jun. 2023.

must face.

1. Code-driven law and legitimacy

It is necessary to reflect on whether data-driven Law might eventually evolve into code-driven Law⁹⁶, that is, if exclusively automated decisions will be prevalent in the future – and whether there is a common democratic interest in this regard. Within the realm of intellectual tasks of AI, theoretically, models could conduct the adjudication of judicial cases based on the analysis of facts, applicable norms, and jurisprudence regarding similar cases. But is this desirable? To answer this question, it's mandatory to investigate if there's the possibility of models to absorb the axiological principles of Justice. Furthermore, legitimacy issues will need to be overcome for judgments made by models rather than by humans. There are doctrinal perspectives advocating that legal reasoning requires a human judgment not only of subsuming facts under norms but also of deliberations, historical processes, community affiliation, interest in social progress, and hence judgments cannot be performed by non-humans in a democratic state⁹⁷. Indeed, currently, judicial activities, such as judgments and rules, are non-delegable in Brazil.

2. Surveillance, privacy and data protection

To archive a relevant data base to be used in data-driven techniques' application, personal data, whether sensitive or not, may be processed, enabling models to make judgments, correlations, or classifications through them.⁹⁸ Thus, it is necessary to debate the possibility (and consequences) of state and private surveillance in all aspects of a citizen's life, especially when considering Personalized Law. After all, to make Personalized Law possible, knowledge of all characteristics of the citizen is required, which could include categorizing them based on political beliefs, health conditions, racial origin, among other sensitive personal data⁹⁹. How can privacy and protection of

⁹⁶ To delve into this topic, see: Mireille Hildebrandt. *Code-driven Law: Freezing the Future and Scaling the Past*. In: Simon Deakin; Christopher Markou. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 75-79.

⁹⁷ Lyria Bennet Moses. *Not a Single Singularity*. In: Simon Deakin; Christopher Markou. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 67-85.

⁹⁸ Ana Frazão. *Julgamentos Algorítmicos: a necessidade de assegurarmos as preocupações éticas e o devido processo legal*. In: Cristiano Colombo; Wilson Engelmann; José Luiz de Moura Faleiros Júnior. *Tutela Jurídica do Corpo Eletrônico*. Novos Desafios ao Direito Digital. Editora Foco, 579.

⁹⁹ *Ibidem*, 581.

personal data be ensured in this scenario, especially to harmonize such treatment with the Brazilian LGPD? How can it be guaranteed that data will only be processed for lawful purposes and guided by social interests?

3. Immutability

The data mining process reveals correlations, not causalities¹⁰⁰, as it identifies reliable patterns, trends, and associations among variables that describe or anticipate processes or events and project them into the future. In another words, what derives from the data is determined by what is contained within the data.¹⁰¹ On the other hand, future events necessarily will impact the interpretation of a legal norm, as interpretation is dynamic, constructed through social evolution and constantly interchangeable.¹⁰² In this context, how can we ensure that data-driven decisions and automated decisions are aligned with social advancements and not merely reproduce old patterns? Could data-driven normativity freeze the future?¹⁰³

4. Algorithmic discrimination

Given that the database is fed with historical data, data-driven decisions or automated decisions might incorporate existing discriminatory biases within society. How can we ensure that these decisions do not further exacerbate existing social problems? It's crucial to question which data can be collected and what purposes can justify algorithmic judgment. There are valid reasons to argue that the more complex the judgment and the social impact, the greater care must be taken.¹⁰⁴

5. Computer language and conversion of human language

Computational language is governed by binary logic. Legal language, on the other hand, contains open and abstract concepts, as well as ambiguous,

¹⁰⁰ Emre Bayamlioglu; Ronald Leenes. *The 'rule of law' implications of data-driven decision-making: a techno-regulatory perspective*. In: Law, Innovation and Technology, 2018, v. 10, n. 2. p. 302.

¹⁰¹ In: *Ibidem*, 302.

¹⁰² Mireille Hildebrandt. *Code-driven Law: Freezing the Future and Scaling the Past*. In: Simon Deakin; Christopher Markou. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 75-79.

¹⁰³ *Ibidem*, 75.

¹⁰⁴ Ana Frazão. *Julgamentos Algorítmicos: a necessidade de assegurarmos as preocupações éticas e o devido processo legal*. In: Cristiano Colombo; Wilson Engelmann; José Luiz de Moura Faleiros Júnior. *Tutela Jurídica do Corpo Eletrônico. Novos Desafios ao Direito Digital*. Editora Foco, 584.

imprecise, and indeterminate terms¹⁰⁵. Is it possible to ensure an efficient translation of legal language into the structure of data and in the creation of AI models based on computational language?

6. Inaccuracy

Small alterations in the input commands may result in entirely different correlations in the output¹⁰⁶. Hence, the objective function and the goals sought by a data-driven law technique may not align with the key result if there's a small change. In other words, is mandatory to choose wisely the inputs, as the attributes being considered in a data-driven decision-making process might not reflect the aspects that should legally be deemed relevant for the rationale of decision-making. Beyond this problem, a machine learning model could learn to make incorrect correlations while processing a database. How can we ensure the accuracy of models in making relevant and legally sound correlations? How can we ensure that the legislator's intention, or even the constitutionality of decisions, is observed in the produced outcomes?

7. Black box models, opacity and auditing

In the promotion of public policies, it is desirable that predictive or automated decision models are explainable, particularly considering the principle of transparency that governs the Public Administration¹⁰⁷. Is it feasible to develop explainable models for decision-making? Can models be audited for explanation or review of decisions made by them?¹⁰⁸ And if so, can it be ensured that the explanations are comprehensible to jurists and/or the general population? How can such decisions be harmonized with principles of the Public Administration?

¹⁰⁵ Mireille Hildebrandt. *Code-driven Law: Freezing the Future and Scaling the Past*. In: *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 2020 P. 71-72.

¹⁰⁶ Also see: Jane R. Bambauer; Tal Zarsky, Jonathan Mayer. *When a Small Change Makes a Big Difference: Algorithmic Fairness among Similar Individuals*. University of California Davis Law Review. Davis: UC Davis Law School. v. 55, n. 04, p. 2337-2420, Apr. 2022.

¹⁰⁷ Brazilian Federal Constitution, Art. 37, *caput*.

¹⁰⁸ Jane R. Bambauer; Tal Zarsky, Jonathan Mayer. *When a Small Change Makes a Big Difference: Algorithmic Fairness among Similar Individuals*. University of California Davis Law Review. Davis: UC Davis Law School. v. 55, n. 04, p. 2337-2420, Apr. 2022.

8. Information security

As previously discussed, the performance of data-driven techniques presupposes the maintenance of a large database, which may contain sensitive information, of substantial volume, and on a broad scale. Consequently, any information security incidents could impact the lives of thousands of individuals, including the basic routine decisions made towards them. How can satisfactory cybersecurity be ensured in the maintenance of this database? How can it be technically guaranteed, when applicable, that personal data is properly anonymized or pseudonymized, so that the data subject cannot be identified by any means?

FINAL CONSIDERATIONS: CAN LAW BE DATA-DRIVEN?

In this article, we set out to explore the data-driven appliances that have been embraced by the Law, on grounds to examine which applications are feasible, which are in perspective and what are the limitations, impacts and risks of the data-driven logic. As a conclusion, it was understood that a decision-making process based on data-driven logic is, above possible, also desirable, when provided that on the process an affective and ethical approach in data governance¹⁰⁹ is implemented align with data protection, information security and fundamental rights.¹¹⁰ Caution is mandatory when data are being use as support for decision-making, considering the potential for misinterpretation of collected data, both by Courts and legal departments, to avoid generalizations. The inputted data and the resulting outputs should be constantly subject to evaluation in order to preserve the critical thinking essential to legal professionals. In this sense, beyond the points raised in this article, the challenges of data-driven law should be consistently mapped, discussed and overcome.

Without misconception of the main critiques and most red flags, both arguments for and against data-driven law converge at a common point: programmers, data scientists and jurists need to work together and define principles, possibilities and priorities. In the interface between Law and Data Science, Law scholars too have much to teach.¹¹¹ Ethical concerns must

¹⁰⁹ José Luiz de Moura Faleiros Junior. *Accountability e Devida Diligência como vetores da governança corporativa nos mercados ricos em dados*. Rio de Janeiro: Revista Semestral de Direito Empresarial n.º 26, 185.

¹¹⁰ *Ibidem*, 211.

¹¹¹ “Designers of machine learning models should be able to draw widely on the expertise of the relevant justice system professionals. Forming mixed project teams in short design cycles to produce functional models is one of the organizational methods making it possible to capitalize on this multidisciplinary approach. Existing ethical safeguards should be constantly shared by these project teams and enhanced using feedback.” In: European

permeate not only all stages of algorithmic systems but also their execution, which leads to a new relationship among those two involved stakeholders.¹¹² Assim, deve ser incentivada a interdisciplinaridade para que a arquitetura dos sistemas jurídicos *data-driven* observe as normas jurídicas desde a concepção inicial, mapeando-se estratégias para incorporar proteções legais na infraestrutura tecnológica¹¹³. Thus, interdisciplinary collaboration should be encouraged so that the architecture of data-driven legal systems observes legal norms from the very outset, strategizing to incorporate legal safeguards into the technological infrastructure.

The best way to stay on the "technological track" is to be at the next station¹¹⁴. In that sense, being contrary to the innovations brought by data-driven approaches will not consequently avoid technological progress, but it may prevent the Law from boarding the train. From an optimistic perspective, we believe that Data Science will enhance the jurist's freedom of action, as it provides more efficient methods for conducting legal procedures¹¹⁵. It's undeniable: automated systems often collaborate with humans in legal tasks, offering legal advice and making (or at least participating on) legal decisions¹¹⁶. Therefore, before engaging in blind opposition, it's essential for the jurist to grasp the intersection between Law and Data Science in order to better harness technological appliances that can build solutions aligned with the real and verifiable issues in our field of knowledge.

Commission for the Efficiency of Justice (CEPEJ). European Ethical Charter on the use of Artificial Intelligence in Judicial Systems and Their Environment. Strasbourg, 2018, p. 05.

¹¹² Frazão, Ana. Julgamentos Algorítmicos: a necessidade de assegurarmos as preocupações éticas e o devido processo legal. In: Cristiano Colombo; Wilson Engelmann; José Luiz de Moura Faleiros Júnior. Tutela Jurídica do Corpo Eletrônico. Novos Desafios ao Direito Digital. Editora Foco, 2022. p. 594.

¹¹³ Mireille Hildebrandt. *Code-driven Law: Freezing the Future and Scaling the Past*. In: Simon Deakin; Christopher Markou. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 82.

¹¹⁴ Antônio Anselmo Martino. *Logic, Informatics, Artificial Intelligence And Technology In: Law: History And Challenges*. In: Leonardo Parentoni; Renato César Cardoso. *Law, Technology and Innovation II: Insights of Artificial Intelligence and The Law*. Editora Expert: Belo Horizonte, 2021. p. 47.

¹¹⁵ Lee Loevinger. *Jurimetrics: The Next Step Forward*. Minnesota Law Review. Minneapolis: University of Minnesota Law School. v. 33, n. 05, p. 477, Apr. 1949.

¹¹⁶ Lyria Bennet Moses. *Not a Single Singularity*. Simon Deakin; Christopher Markou. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 2020. P. 208.

REFERENCES

- Bambauer, Jane R.; Zarsky, Tal; Mayer, Jonathan. When a Small Change Makes a Big Difference: Algorithmic Fairness among Similar Individuals. *Univ. of Cal. Davis L. Rev.* 55, no. 4 (Apr. 2022): 2337-2420.
- Bayamlioglu, Emre; Leenes, Ronald. The ‘Rule Of Law’ Implications Of Data-Driven Decision-Making: A Techno-Regulatory Perspective. *Law, Innov. & Tech.* 10, no. 2 (2018): 295-313. Available at: <https://doi.org/10.1080/17579961.2018.1527475>. Accessed on 27 Jun. 2023.
- Ben-Shahar, Omri; Porat, Ariel. *Personalized Law: Different Rules For Different People*. Oxford: Oxford University Press, 2021.
- Bobbio, Norberto. *Teoria Geral Do Direito*. São Paulo: Martins Fontes, 2010.
- Brasil. Conselho Nacional de Justiça. *Justiça em Números*. 2022. Available at <https://www.cnj.jus.br/wp-content/uploads/2022/09/justica-em-numeros-2022-1.pdf>. Accessed on 20 Jun. 2023.
- Brasil. Conselho Nacional de Justiça. *Resolução n.º 332, de 21 de agosto de 2020*. Available at <https://atos.cnj.jus.br/files/original191707202008255f4563b35f8e8.pdf>. Accessed on 20 Jun. 2023.
- Brasil. Conselho Nacional de Justiça. *Resolução n.º 333, de 21 de setembro de 2020*. Available at <https://atos.cnj.jus.br/atos/detalhar/3488>. Accessed on 20 Jun. 2023.
- Brasil. Conselho Nacional de Justiça. *Resultado Pesquisa IA No Poder Judiciário em 2022*. Available at https://paineisanalytics.cnj.jus.br/single/?appid=9e4f18ac-e253-4893-8ca1-b81d8af59ff6&sheet=b8267e5a-1f1f-41a7-90ff-d7a2f4ed34ea&lang=pt-br&theme=ia_pj&opt=ctxmenu,currsel&select=language,br. Accessed on 20 Jun. 2023.
- Brasil. *Decreto n.º 10.411, de 30 de junho de 2020*. Regulamenta a Análise de Impacto Regulatório. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2020/Decreto/D10411.htm. Accessed on 27 Jun. 2023.
- Brasil. *Decreto n.º 8.777, de 11 de maio de 2016*. Institui a Política de dados Abertos do Poder Executivo Federal. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Decreto/D8777.htm. Accessed on 28 Jun. 2023.
- Brasil. *Decreto n.º 9.830, de 10 de junho de 2019*. Regulamenta os Artigos 20 a 30 da Lei n.º 13.655/2018. Available at:

- https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Decreto/D9830.htm. Accessed on 27 Jun. 2023.
- Brasil. *Decreto-Lei n.º 4.657, de 4 de setembro de 1942*. Lei de Introdução às Normas do Direito Brasileiro. Available at: https://www.planalto.gov.br/ccivil_03/Decreto-Lei/Del4657.htm. Accessed on 25 Jun. 2023.
- Brasil. *Lei n.º 11.598, de 3 de dezembro de 2007*. Cria a Rede Nacional Para Simplificação do Registro e Legalização de Empresas e Negócios. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Lei/L11598.htm. Accessed on 28 Jun. 2023.
- Brasil. *Lei n.º 12.965, de 23 de abril de 2014*. Marco Civil da Internet. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2014/Lei/L12965.htm. Accessed on 27 Jun. 2023.
- Brasil. *Lei n.º 13.655, de 25 de abril de 2018*. Inclui na Lei de Introdução às Normas do Direito Brasileiro disposições sobre segurança jurídica e eficiência na criação e na aplicação do Direito Público. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13655.htm. Accessed on 25 Jun. 2023.
- Brasil. *Lei n.º 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais (LGPD). Available at: https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm. Accessed on 27 Jun. 2023.
- Brasil. *Lei n.º 13.848, de 25 de junho de 2019*. Dispõe sobre a Gestão, a Organização, o Processo Decisório e o Controle Social DAS Agências Reguladoras. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Lei/L13848.htm. Accessed on 27 Jun. 2023.
- Brasil. *Lei n.º 13.874, de 20 de setembro de 2019*. Institui a Declaração de Direitos de Liberdade Econômica. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Lei/L13874.htm. Accessed on 27 Jun. 2023.
- Brasil. *Lei n.º 14.129, de 29 de março de 2021*. Lei Do Governo Digital. Available at: https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2021/Lei/L14129.htm. Accessed on 27 Jun. 2023.
- Brunschwig, Colette. Visual Law and Legal Design: Questions and Tentative Answers. *Internationales Rechtsinformatik Symposium*, Editions Weblaw, 2021.
- Cavalcanti, Eugênia Giovanna Simões Inácio. *O dever de motivar os atos administrativos como princípio implícito na Constituição Federal*. Dissertação De Mestrado. Recife: UFPE, 2004.
- CLOC (Corporate Legal Operations Consortium). 2021 *State of The Industry Survey*. Results And Analysis. 2021.

- Coelho, Fábio Ulhoa. O uso da jurimetria. *Rev. de Direito Bancário e do Mercado de Capitais* 63 (2014): 193-199.
- Comissão Europeia para Eficácia da Justiça. *Carta Europeia de Ética sobre o uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente*. Strasbourg, 2018.
- Deakin, Simon; Markou, Christopher. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 2020.
- Deloitte. *2022 State of Legal Operations Survey*. Taking Steps to Modernize Legal Operations. 2022. Available at: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/advisory/us-rfa-legal-op-infographic-2023.pdf>. Accessed on 25 Jun. 2023.
- Faleiros Júnior, José Luiz De Moura. Accountability e devida diligência como vetores da governança corporativa nos mercados ricos em dados. Rio De Janeiro: *Revista Semestral De Direito Empresarial* No. 26, 2020. p. 183-211. Available at: <https://rsde.com.br/wp-content/uploads/2021/12/rsde-026-completa.pdf> Accessed on: 27 Jun. 2023.
- Faleiros Júnior, José Luiz de Moura. Inovação, Desenvolvimento E A Lei Do Governo Digital (Lei Nº 14.129/2021): Perspectivas Quanto Às Estruturas De Governo Como Plataforma. *Inova Jur*, Belo Horizonte, V. 1, N. 1, P. B1-B29, Jan./Jun. 2022.
- Faleiros Júnior, José Luiz De Moura; Borges, Alexandre Walmott; Ayzama, Alex Cabello. Reflexões sobre o consequencialismo jurídico na Lei da Segurança para Inovação Pública (Lei n.º 13.655/2018). *Revista Jurídica da Universidade do Sul de Santa Catarina*, p. 111-121. Available at: https://www.researchgate.net/publication/345413837_Reflexoes_Sobre_O_Consequencialismo_Juridico_Na_Lei_Da_Seguranca_Para_A_Inovacao_Publica_Lei_N_136552018. Accessed on: 27 Jun. 2023.
- Fortino, Andres. *Data Mining and Predictive Analytics for Business Decisions: A Case Study Approach*. Dulles: Mercury Learning & Information, 2023, p. 40.
- Frazão, Ana. Julgamentos Algorítmicos: A Necessidade De Asseguarmos As Preocupações Éticas E O Devido Processo Legal. In: Colombo, Cristiano; Engelmann, Wilson; Faleiros Júnior, José Luiz De Moura. *Tutela Jurídica do Corpo Eletrônico*. Novos Desafios Ao Direito Digital. Editora Foco, 2022. p. 579-588.
- Hildebrandt, Mireille. Code-Driven Law: Freezing the Future and Scaling the Past. In: Deakin, Simon; Markou, Christopher. *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*. Oxford: Hart Publishing, 2020. P. 67-85.

- Hod, Shlomi; Karni, Chagal-Feferkorn; Niva, Elkin-Koren; Avigdor Gal, Technion. Data Science Meets Law: Learning Responsible Ai Together. *Communications Of The ACM*. Vol. 65 No. 2, Pages 35-39. 2022. Available at: <http://dx.doi.org/10.2139/ssrn.4077216>. Accessed on: 20 Jun. 2023.
- Layton, Mark; Ostermiller, Steven; Kynaston, Dean. *Scrum For Dummies*. A Wiley Brand. Hoboken: John Wiley & Sons Inc., 2023.
- Lippi, Marco; Torroni, Paolo. Argumentation Mining: State of the Art and Emerging Trends. *ACM Trans. Internet Technol.* V, N, Article A. 2015. DOI: <http://dx.doi.org/10.1145/0000000.0000000>.
- Loevinger, Lee. Jurimetrics: The Methodology of Legal Inquiry. *Law And Contemporary Problems* No. 28, 1963.
- Loevinger, Lee. Jurimetrics: The Next Step Forward. *Minnesota Law Review*. Minneapolis: University Of Minnesota Law School. Vol. 33, No. 05, p. 455-493, Apr. 1949.
- Martino, Antônio Anselmo. Logic, Informatics, Artificial Intelligence and Technology In: Law: History and Challenges. In: Parentoni, Leonardo; Cardoso, Renato César. *Law, Technology and Innovation II: Insights on Artificial Intelligence and the Law*. Editora Expert: Belo Horizonte, 2021. p. 28-47.
- Medeiros, Nathália Robert Fett Viana. Uso da inteligência artificial no processo de tomada de decisões jurisdicionais: uma análise sob a perspectiva da teoria normativa da participação. 2019. Dissertation (Master of Laws) – Faculty of Law, Pontifical Catholic University of Minas Gerais, Belo Horizonte, 2019.
- Mochales, Raquel; Moens, Marie Francine. Argumentation Mining. *Artificial Intelligence And Law*, Berlim, V. 19, N. 1, P. 1–22, Mar. 2011. Available at: <https://doi.org/10.1007/s10506-010-9104-x>. Accessed on: 17 Jun. 2023.
- Monteiro, Renato Leite. *Desafios para a efetivação do direito à explicação na Lei Geral de Proteção de Dados do Brasil*. 2021. 383 f. Tese (Doutorado Em Direito Comercial) – Faculdade de Direito, Universidade de São Paulo, São Paulo, 2021.
- Moses, Lyria Bennet. Not A Single Singularity. In: Deakin, Simon; Markou, Christopher. *Is Law Computable? Critical Perspectives on Law And Artificial Intelligence*. Oxford: Hart Publishing, 2020. p. 205-223.
- Nunes, Marcelo Guedes. *Jurimetria: como a estatística pode reinventar o direito*. São Paulo: Revista Dos Tribunais, 2016.
- Parentoni, Leonardo. *Compartilhamento de dados pessoais e a figura do controlador* (Personal Data Sharing and the Role of the Data Controller). 2021.

- Parentoni, Leonardo. What Should We Reasonably Expect from Artificial Intelligence? In: *Il Diritto Degli Affari*, 2022.
- Provost, Foster; Fawcett, Tom. *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*. Sebastopol: O'Reilly, 2013.
- Salomão, Luis Felipe. *Inteligência Artificial: Tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário Brasileiro*. São Paulo: Fundação Getúlio Vargas, 2020.
- Snow, Charles. *The Two Cultures*. Cambridge: Cambridge University Press, 1959.
- Thomson Reuters Institute and The Legal Value Network. *2021 Legal Department Operations Index. The Risk of Being Left Behind*. 2021. Available at <https://legal.thomsonreuters.com/content/dam/ewp-m/documents/legal-tracker/en/pdf/reports/2021-ldo-report.pdf>. Accessed on 25 Jun. 2023.
- Thomson Reuters Institute and The Legal Value Network. *2022 Legal Department Operations Index. Seeking Stability Amid Uncertainty*. 2022. Available at https://legal.thomsonreuters.com/content/dam/ewp-m/documents/legal/en/pdf/reports/legal-dept-operations-index_2022.pdf. Accessed on: 25 Jun 2023.
- Turing, Alan M. Computing Machinery and Intelligence. *Mind*, N. 49, P. 433-460. Available at: <http://phil415.pbworks.com/f/TuringComputing.pdf>. Accessed on: 28 Jun. 2023.
- Valentini, Rômulo Soares. Electronic Lawsuit Systems and Machine-Made Judgments: Developing Standards for a “Legal Turing Test”. In: Parentoni, Leonardo; Cardoso, Renato César. *Law, Technology and Innovation II: Insights on Artificial Intelligence and The Law*. Editora Expert: Belo Horizonte, 2021. P. 328-346.
- Walters, Ed. *Data Analytics and The New Legal Services*. Boca Raton: CRC Press. 2019.

* * *

Lorenzo Antonini Itabaiana

Master of Laws candidate, Federal University of Minas Gerais. Specialist in Digital Law, Innovation Management, and Intellectual Property from PUC Minas. Holds a Bachelor of Laws degree from the Federal University of Minas Gerais, with a period of study at Universitetet i Oslo, Norway.

Email: lorenzo.antonini@dcom.law

ORCID iD: <https://orcid.org/0009-0008-2966-9644>

Leticia Camargos Ferraz

Bachelor of Laws from the Federal University of Minas Gerais. Corporate Consulting Lawyer.

Email: letqsf@gmail.com

ORCID iD: <https://orcid.org/0009-0007-3849-2185>