The ethics of predictive justice

Louis Larret-Chahine,
co-founder and general manger of Predictice

Abstract:
The intrusion of digital technology in the law field, via the emergence of legaltechs and predictive justice, calls for updating the rules of ethics necessary for justice to operate as it should. The design and control of algorithms, the anonymization of data, the management of biases in databases, sanctions against pirates… clear rules have to be drafted and obeyed lest litigants' trust in the justice system be eroded.

To act ethically is to comply with a set of moral principles devised to protect an ideal perceived to be universal. The ideal of justice, an underpinning of a democratic society's social pact, enters in this category. However the concept of justice is not fully in this world and should not be — as legal professionals well know since they handle this concept day after day. They do not mix the ideal of justice up with the legal or political quarrels that agitate the public service of justice in France. It is, therefore, indispensable to separate these two imperatives. On the one side is the ideal of justice, nearly divine but, in reality, untouchable by technology, except in certain works of literature that are still (quite fortunately) merely dystopian fantasies. On the other side are the objectives of a public service: efficient, intelligible and pragmatic for citizens and necessary for upholding the social pact. On this side, which has ample room for perfection, a fertile hybridization will take place between artificial intelligence (AI) and the law.[1]

Listen to the outrage: the irruption of digital technology in the realm of the law will erode the transcendence of a court hearing or even of justice itself. But if justice has so awesomely withstood the death of God, it will also withstand the digitization of hearings or the reduction of "contingency" in routine cases of litigation, most of which are already receiving standardized treatment. Listen to other voices: our era is witnessing an alliance between jurists and engineers with, as one (unspeakable) goal, the automation of justice. Nothing is more false. Since the very start, ethical issues lie at the center of this debate, evidence of this being the many initiatives, joint efforts and research programs on this topic. The goal has always been to improve the performance of legal professionals and the transparency of justice as a public service while upholding the still indispensable ethical standards.

In this context, predictive justice has, at lightning speed, become a wedge issue in legal circles. Apologists or naysayers… everyone has something to say about how legal professionals will appropriate assisted decision-making tools. Setting aside all preconceived opinions, one point receives unanimous support: the revolution of predictive justice — a technological challenge to the practices of justice — cannot be accomplished without respect for ethical and deontological imperatives, some of which must be reconsidered in the light of the possibilities brought by this technology.

[1] This article, including quotations from French sources, has been translated from French by Noal Mellott (Omaha Beach, France). The translation into English has, with the editor's approval, completed a few bibliographical references.

This overview of the ethical problems raised by the development and deployment of predictive justice examines how these problems have been handled and solved up till now. The three principal meeting points between ethics and predictive justice will be discussed chronologically: the design of digital tools, the processing of data, and the impact on legal standards and judgments.

Ethics in the very design of predictive justice

Intelligibility, design flaws: the lack of understanding results in the black box syndrome, especially in the realm of the law, which has long been technology-proof. Concerns about this (those that are legitimate) lead to proposing solutions such as open source code or the creation of a regulatory authority.

Automatic processing of the language of the law: Find the errors!

The lack of intelligibility of digital techniques makes us wary, like certain inventions (*e.g.*, Louis Lumière's moving pictures) that were, at the start, considered to be magic. In such cases, effort must be put into education.

Carrying out statistical operations on a large volume of court decisions calls for machine learning or other automatic learning processes. Machine learning, a field of study in AI, is sometimes confused with AI itself in the articles and press releases made for the general public. It refers to the design, analysis, development and implementation of the methods whereby a computer program evolves through systematic processes. Natural language processing is a branch of machine learning with the goal of analyzing, understanding and even producing human language. These techniques were first rolled out in certain branches of industry, such as aeronautics, without raising questions of an ethical sort.

Long impervious to the latest technological trends, the realm of the law has reacted less favorably to this technology. Legal professionals, initially not much concerned by (nor trained in) technology, abstained, for a long time, from making any efforts to adapt. As a consequence, some of them consider this technology to be a "black box" with incomprehensible or unforeseeable operations. They lack the skills for taking apart software and putting it back together, for controlling and auditing programs. This problem of intelligibility usually arouses apprehension about flaws in the software's design.

Apprehensions about errors in predictive justice software have to be overcome by advocating transparency. In fact, teams are permanently testing solutions, which are usually built with "bricks" that other firms have made public (as happened with Syntaxnet, which Google opened).In the case of natural language processing, 100% reliability is hard to imagine, even theoretically. Language evolves fast (as do court decisions and legal concepts); and algorithms, even perfectly built, might always make mistakes when resolving ambiguous syntactical constructions. In others words, given that what is being analyzed evolves and sometimes has several meanings, algorithms do not always perfectly understand language or the meanings hidden in it (as happens with the very complex double negations used in court decisions).

Solutions for dealing with this lack of intelligibility and the risk of errors are limited: open source algorithms or an organization of an authority control or certify the software.

Transparency and regulation: Alternative or cumulative solutions?

The transparency of algorithms is a thorny question. Designing an algorithm entails bringing together a talented team whose members will work together for several months, or even years. This is costly, but a company that owns its software might thus gain a competitive advantage. The GAFA (Google, Apple, Facebook, Amazon) sometimes "open" their algorithms, mainly to stimulate an ecosystem of innovation and create "soft power" around their brands. But they always keep control over the data in their possession, thus maintaining a competitive advantage that cannot be offset.

So, should legaltechs make their code "open source"? Some authors[2] answer "yes" because this solution is a prerequisite for building up confidence. But in a world where data are open (as is the case of court decisions in France), software code cannot be fully open lest its developer lose the competitive advantages gained by building it. Forcing firms to make their algorithms available to everyone would block innovations.

An alternative exists: an independent, national regulatory authority could be given access to all source code from a project that uses the open data on court decisions. It would vouchsafe the integrity of the algorithms while protecting them. There is a precedent for this sort of regulation: the algorithms for high-frequency trading under EU regulations and directives (MiFID and MiFIR).[3] To avoid setting up yet another independent administrative authority in France, the assignment could be handed over to the National Commission on Informatics and Liberty (CNIL: Commission Nationale de l'Informatique et des Libertés). But a few questions would have to be settled. How often should the code be transmitted: once to be certified for implementation? continually? on demand? Does each update have to be recertified? Is the certification to be obtained *ex ante* or *a posteriori*? The crux of regulating this technology is that it keeps on evolving and that exercising oversight might be hard to do without restraining innovation.[4] Furthermore, the regulatory process ties up resources.

While waiting for an institutional form of regulation, spontaneous initiatives have sprung up. The associations ADIJ and Open Law have drafted a charter of ethics,[5] which has rallied most legaltechs. This not very restrictive charter makes a first step and is to be encouraged. In 2017, Predicte set up a scientific and ethics committee on predictive justice, whose members (nine legal professionals, law professors and practitioners) reflect on the ethics underlying the company's activities and thus have access to relevant in-house documentation.[6]

As in all data-processing industries, value is generated by using well-designed algorithms on data (in the present case, on court decisions). However collecting and processing data also raises questions of ethics

---

[2] For example: BENESTY M. (2017), "L'*open data* et l'*open source*, des soutiens nécessaires à une justice prédictive fiable?", *Journal of Open Access to Law*, 5(1), 11p. Available at https://docplayer.fr/65053702-L-open-data-et-l-open-source-des-soutiens-necessaires-a-une-justice-predictive-fiable.html.

[3] Cabinet Norton Rose Fulbright (2014), *MiFID/MiFIR Series*, April, available at https://www.regulationtomorrow.com/eu/trading-venues-and-market-infrastructure/.

[4] KAY J. (2017) "How do you regulate a self-improving algorithm?", *The Atlantic*, 25 October.

[5] https://www.charteethique.legal/charte-ethique

[6] https://predictice.com/. The Clinique de l École de Droit de Sciences Po Paris monitors this committee's work.

Ethics in collecting and processing court decisions

As customers are being asked to exchange their personal data to obtain for-free services, personal data have become a precious good, sensitive to ethics — even more so when the data concern litigants, in particular physical persons. The issue of personal data has cropped up in the plans for implementing articles 20 and 21 of the Digital Republic Act, which state that the decisions and rulings by the various jurisdictions in France are to be "*made available to the public for free while upholding the privacy of the concerned*".[7] This provision encounters two major ethical questions, the one about the anonymization of personal data and the other about biased data sets.

Anonymization and reidentification

Since an enabling decree has not yet been issued for the Digital Republic Act at the time of the writing of this article, it is hard to imagine how anonymization will be applied to open data on court decisions. Will persons' names simply be replaced with "Mr. X" and "Ms. Y" without any other modification? This solution is undeniably the most advantageous in terms of cost, feasibility and interest. The same could be done, without much difficulty and without deviating from the spirit of the law, to telephone numbers, addresses and others identifying pieces of information.

However anonymization could reach farther, to the context. This would come at a much higher price since, given the state of technology, this solution would entail manually rereading court decisions to verify the results afterwards. This would drastically reduce the law's utility and not necessarily prevent reidentifying private persons.[8] Let us note that, from the viewpoints of commerce and justice, reidentification is of very little interest. Besides, Article 226-19 of the Penal Code provides for a sanction of "*five years of imprisonment and a fine of €300,000 [… for] the fact, except for the cases foreseen under the law, of placing or storing in computerized memory data of a personal sort about offenses, convictions or detentions*". As we see, a dissuasive sanction already exists.

Managing biases in the data

The algorithms used to statistically analyze court decisions require a massive quantity of data; and biases can be lurking in the mass. These software programs, like the algorithms used for big data, are restricted to using the data fed into them. This input, even if limited to court decisions, is not exempt from biases, sometimes unconscious. Compas provides the best known example of biases in the training data used for automatic learning. This software program, developed by Northpoint for the purpose of predicting recidivism rates, reinforces racial prejudice by evaluating Afro-American offenders more severely.[9]

Several solutions can be imagined to rectify biases of this sort.

The first one is to be rejected from the start, namely modify the data set so that it matches what we want. This approach is dangerous, since it would undermine confidence in assisted decision-making tools. Furthermore, the analyses made using the modified data would lack interest.

---

[7] Act n°2016-1321 of 7 October 2016 for a "digital republic" available at https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000033202746&categorieLien=id.

[8] According to Tristan Allard, associate professor in information science at Rennes University, who spoke about the anonymization of court decisions during a conference on predictive justice held on 6 April 2018 in Laval. A publication is forthcoming.

[9] According to Laurel Eckouse, an academic. *Cf.* https://www.lebigdata.fr/justice-big-data-prejudice-race-1302

A second possibility is to point out the biases so that corrections are gradually made by intervening in the data set as it is formed and thus influencing the statistics to be drawn from it later on. A system of alerts could be set up. When, on a given subject, a judge pronounces sentences that are 90% more or less harsh than the sentences pronounced by all other judges, a public discussion could take place to understand the grounds for the decisions made. Biases would thus gradually vanish thanks to the collective intelligence of the profession and of the court officers responsible for detecting and reflecting on variations in court decisions throughout the country.

Another way to reduce risks is to refuse to work with algorithms in certain fields of law. Just as driverless vehicles learn to never cross a solid white line, algorithms can be programmed to not operate in certain branches of the law. Predicte's ethics committee on predictive justice has made this decision about penal.

Finally, an imperative condition must be pointed out: all algorithms must operate on the same database. This means being clear about the origin of the data and the method of collecting them; but it also implies harsh sanctions for uncontrolled data sets, which strongly skew results and deceive users. If data have been gathered from a given firm or region, the analyses made using them will no longer be based on a representative sample or on an exhaustive data set of court decisions, but on a database that we do not know, neither how it has been formed nor whether it is complete.

Ethics in the operation of justice

Apprehensions about the impact of predictive justice on justice itself underlie many of the debates on this new technology. Without letting our imagination wander so far as robot judges (completely unrealistic because the process for making a court decision is complex and multifactorial), we should inquire into the effects of predictive justice on access (or recourse) to the law and on the standardization of court decisions.

Dejudicialization, restricted access to the law?

Since the tools for statistically analyzing court decisions provide a glimpse of what can be gained from introducing an action in court, they are an amazingly effective weapon for use during negotiations. We can thus predict bright prospects for disputes being settled out of court; but we can also worry about the scale of this coming "dejudicialization".

The key point is access to the justice system. Predictive justice will not be appreciated unless it improves this access and the operation of justice. Antoine Garapon and Jean Lassègue have said as much: "*What would justice be if it did not address the sense of justice — what is human in human beings?*"[10] But does the recourse to justice necessarily involve appearing before a court that will examine the merits of the case, and then appealing that decision until finally obtaining (or not) what is due while wasting years and losing a colossal sum of money? After all, the reason that litigants sometimes have the feeling that there is no "*sense of justice*" is due to the justice system itself.

In a context where magistrates are overloaded with cases, where waiting periods are stretching out, and where litigants barely have confidence in justice, dejudicialization is to be taken under consideration — more seriously than by simply restricting the procedures for entering an appeal or by complicating the forms for petitioning a labor court. Alternative methods of dispute settlement have an obvious interest for litigants. These methods — more rapid, less expensive, less traumatic — come up with satisfactory settlements for a wide variety of cases. Dejudicialization does not mean obstructing the recourse to justice. On the contrary, it means

[10] GARAPON A. & LASSÈGE J. (2018) *Justice digitale* (Paris: Presses Universitaires de France), p. 134.

better managing the flow of cases so that settlements are found, whether in or out of the courtroom.

A performative risk: A new system for producing legal standards?

Performativity refers to the phenomenon of self-fulfilling prophecies whereby simply having access to information has an impact on reality. There are several examples, especially in finance, where predicting an event strongly affects the chances of the event being realized. However this risk seems less probable in the case of predictive justice, mainly because algorithms cannot analyze events that depend on irrational factors or lie outside the court's decision (the temperature in the room where the hearings are held, the judge's fatigue, the fact that a particular plea has been very well made — anything having to do with human sensitivity during the decision-making process). A study is under way on these performative effects at Paris-Dauphine University.

Let us not overestimate the influence of software on judges (under condition that they use the software). Let us not assume that magistrates modify their decisions out of fear of what others will think. This assumption overlooks the judge's intellectual capacity and courage. Besides, judges have to motivate their decisions on legal grounds, and this obligation is itself a form of protection.

Nonetheless, one effect that predictive justice will surely have is to upset legal standards. The deployment of the solutions proposed by predictive justice in the everyday activities of legal professionals is modifying the very concept of case law. Thanks to the combination of predictive justice and open data, the access to case law by these professionals will no longer be limited to a few cases of litigation that were brought before a high court. Lawyers will be able to study in granular detail the millions of decisions made by lower courts every year. This staggering possibility will lead to a new — what has been called "isometric" — judicial system: the normative importance of high court rulings will not disappear but will gradually be diluted insofar as all decisions, from lower to higher courts, can be entered into statistics so as to give a well-honed view of how judges interpret the law.

Open data and predictive justice might, ironically, make justice more transparent and, too, more human and more democratic (or at least isometric) by endowing the nine thousand magistrates in France with a normative voice.