Doctrine.fr,
artificial intelligence at the service of the law

Antoine Dusséaux,
co-founder of Doctrine.fr,
&
Hugo Ruggieri,
legal advisor, Doctrine.fr

Abstract:
The law seems to be among the sectors of the economy least affected by the digital revolution. Given that court decisions and texts of law are studded with references and links to other decisions and texts, the law seems to lend itself to the invention of a Google-like search engine. Artificial intelligence can thus make the law more accessible to litigants and more practicable for professionals. Doctrine.fr, a search engine for the law, places artificial intelligence at the service of people for the purpose not of providing predictive justice but, instead, of assisting justice. Its development depends on the general availability of raw data from judiciary sources (in particular court decisions). These data are necessary to perfect the tools for extending internationally the sway of the French economy and legal model.

A law google?

Twenty years ago, two researchers from Stanford discovered that information on the Internet could be made more accessible. Although the Web had just started developing, it was already possible to imagine its disruptive effects.[1] At the time, the dominant search engines were AltaVista and Yahoo!, which mainly analyzed key words and ranked as first the web page with the most occurrences. At the start, this basic procedure was good enough, but its limits soon appeared. A key word could be hidden in a web page so that the page be automatically ranked better among the findings. The major success of Larry Page and Sergey Brin, Google's two cofounders, was to put in order the mass of data available on the Internet.[2] Google's algorithm determined a website's reputation with the help of a slue of criteria, a principal one being the links toward the site from other sites but along with the reputation of these third-party sites. Google thus worked out a reputation-based grid for displaying not the findings with the most occurrences of the keywords in a search but the findings that the public recognizes as pertinent to the keywords.

The coming of Google disrupted how the Internet operates, since the vast and ever vaster quantity of information on the Web was now being sorted and organized. The promise of opening information to the public implied offering everyone the same access to this world library. This opening of information held promise for developing new services thanks to the knowledge thus gleaned. Twenty years have gone by, and new firms are offering many a service in fulfillment of this promise. Amazon, eBay, Wikipedia, Spotify, Netflix: in each case, a mass of information has been analyzed. Our approach to knowledge, culture and consumption has been changed.

---

[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.

[2] Cf. "How PageRank works", *The Economist*, 16 September 2004.

There is, however, a sector of society still shrouded in mystery because the access to knowledge there is restricted, even though it is one of the institutions most important to society's smooth operation. In France, the realm of justice has, till now, benefitted very little from the advances made in digital technology. Jurisprudence is a matter for a handful of publishers, centenarians that also publish the annotated codes of law used by lawyers and judges. Few court decisions are uploaded on line and made available to clients. Out of the 2,677,253 civil court decisions in France in 2016, only 13,360 were posted on Légifrance (the official website for publishing legal information): approximately 0.5%![3] These publishers also provide the search engines for French law, which, more like AltaVista than Google, mainly carry out searches using keywords.

As a consequence, the realm of justice is accessible only to insiders. This impairs its operation and runs counter to the values of transparency and openness.[4] In contrast, developing a search engine adapted to current needs would enable not only litigants to better understand the institution of justice (which safeguards the conditions for living in society) but also legal practitioners to concentrate on more gratifying aspects of their job. Why has this sector remained so compartmentalized? In other words: where is a law google?

Artificial Intelligence, a new technology for analyzing legal jargon

A major factor in the emergence of such a search engine for the legal field is the need to understand laws and court decisions in order to cogently index them.

The law provides a subject matter suited to the application of a search engine, as on the Internet, since texts of law and court decisions also rely on a tight grid of links and references. Unlike the Internet however, the points in the grid are not clickable links easily visible in a website's code (via the use of the html tags: <a>). Instead they are the citations, precedents, references incorporated in texts of laws and court decisions. Even when the links are straightforward (*e.g.*, the citation of another article in an act of law or of another court decision), they are normally written in what has been called "natural language" in opposition to programming languages (computer code). Very often, the most important references are far from evident for outsiders; or they can but be inferred.[5]

Furthermore, it is not enough just to detect references to other legal texts or court decisions. These references must be understood along with their contexts. Decisions might cite another text but in order to declare it inapplicable, or refer to precedents in case law but to underscore differences with the case at hand. In these examples, it would be misleading to link the referenced texts as if they were positively connected; but this is exactly what happens when a classical search engine analyzes a query by using keywords. Developing an engine that searches the law and comes up with pertinent findings depends not only on identifying links among legal texts and court decisions but also on understanding the context.

---

[3] Statistics from: CADIET L. (2017) "L'open data des décisions de justice", a report at the request of the Ministry of Justice on court decisions as open public data. Available via http://www.justice.gouv.fr/publication/open_data_rapport.pdf.

[4] The right to public hearings and sentencing is a key principle of justice, stipulated in Article 6 of the European Convention for the Protection of Human Rights and Fundamental Freedoms.

[5] For instance, a reference to "Chronopost case law" makes sense, very much so, to private attorneys, but an algorithm will have difficulty identifying it as a keyword in decision n° 93-18.632 of 22 October 1996 by the commercial division of the Court of Cassation.

Processing natural language has made strides thanks to recent developments in artificial intelligence (AI). The latter handles mathematical problems not as a series of conditions (as computers usually do) but with methods that try to imitate human reasoning. In particular, machine learning and deep learning methods are being used to advance as never before by enabling algorithms to learn and improve themselves through testing.

The progress made in AI algorithms in the past few years now allows us to process legalese and the grid of links and references resulting from an attentive reading of legal texts and court decisions. It is now possible to create a literal "law goggle", a smart legal search engine, and use AI to improve the relation between justice, litigants and legal professionals.

So many possibilities at the service of the law

This progress opens a vast scope of possibilities. One is to advance democracy, since the social contract hinges on justice, on the relinquishment of taking justice into one's own hands but under condition that a neutral, independent and transparent institution renders justice impartially and publically. Confidence in the judicial system will be strengthened by opening data on court decisions and legal texts and by using a search engine to make this data available to the public.

An AI-based search engine brings several other possibilities for concrete uses by litigants and legal professionals. As the CNIL has stated in its report on the ethical issues related to AI, "*Algorithmic software programs can have a strategic utility by optimizing the identification of the statistically most probable solutions for a given dispute.*"[6] This means that legal professionals will save time, and that lawyers can "*devote themselves to more gratifying tasks related to cases and involving human contact*". As for litigants, they will be better equipped to evaluate the chances of a lawsuit by comparing their suits to similar cases.

Making the sources of the law more easily available is a factor for improving justice in every way since it facilitates: the work of judges, who will be better able to verify arguments at a hearing; the work of lawyers, who will have better tools for improving pleadings; the work of academics and litigants, who will have access to more pertinent search findings and thus better understand the issues. AI can be used to customize legal tools by adapting them to the practices of each person (*e.g.*, via a system of personalized alerts on new developments related to the specific points of law identified through a person's searches).

Such tools are complementary to human beings (for the purpose of helping them with their jobs or in understanding the legal realm) and not in opposition to human beings (for the purpose of replacing them with predictive algorithms). A modern search engine holds the promise of assisting justice. The key is human beings, since their work will be made easier and more efficient thanks to AI tools that are complementary to it.

Thanks to AI, search engines based on natural language processing, and self-learning recommendation systems, Doctrine.fr tries to see to it that this complementarity exists between human beings and AI and that access to the law is made easier for everyone.

---

[6] CNIL [Commission Nationale de l'Informatique et des Libertés] (2017) *Comment permettre à l'homme de garder la main? Les enjeux éthiques des algorithmes et de l'intelligence artificielle*, December, 80p. Available via: https://www.cnil.fr/sites/default/files/atoms/.../cnil_rapport_garder_la_main_web.pdf.

Open legal data, artificial intelligence and competitiveness

AI tools cannot exist without access to data for two main reasons. First of all, these tools need data as exhaustive as possible in order to not mislead users. It is, therefore, essential that all data be available and accessible so that the services rendered not become a disservice to users, who receive incomplete information. Secondly, improving AI models implies having access to a database as big as possible in order to compare the model with a very large number of cases so as to hone it as a function of the diversity of actual situations.

The promises of an AI-assisted justice can be kept only if court decisions are published, like texts of law, *i.e.*, systematically without modification of their contents. This opening of data, a trend that has been called "open data", now figures in the law.[7] The concrete realization of open court data is essential for France to be competitive internationally and for French search engines in the law to develop.

The open data movement started across the Atlantic, where it is still growing and stimulating startups in the legal field. Smart search engines are being created in the United States. For instance, Judicata has recently raised millions of dollars to develop its search engine and promote an American model for these search engines.[8]

In response to this situation, pursuing an open data policy for legal data is good not just for improving the access to justice but also for the nation's economic, legal and political competitiveness. As Loïc Cadiet stated in the report on court decisions as open data, "*The prospects for improving the law's predictability through open data are, under certain conditions, a factor for improving the competitiveness of all firms and the attractiveness of France.*"[9] There is no doubt but that American legaltechs will soon be setting up in France.[10] France must create its own success stories in this field by developing its own startups. The alternative is simple: accept the domination of the United States in this sector, even though the French government has declared AI a priority for investments.[11]

Apart from the question of the of the country's economic influence, open data in the law field will, according to the Cadiet report, help France "*preserve the independence of its legal system and the attractiveness of its law*". Improving access to the law is likely to spur competition between legal systems. Whereas France has a civil law tradition, the United Kingdom and United States, owing to their common law traditions, pay more attention to precedents in court decisions. Search engines specialized in the law are probably going to take on quickly there, thus facilitating access to their law for not only litigants but also interested third parties and foreigners. Developing search engines in the law is a vector for diffusing legal concepts and models, a field historically pioneered by France. The conditions for applying open data in the legal field will have

---

[7] Articles 20 and 21 of 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. These articles also called for what would become the Cadiet report, see note 3.

[8] Il is noteworthy that one investor was no other than Peter Thiel, who cofounded PayPal and was recently a member of Donald Trump's transition team. *Cf.* ETHERINGTON D. (2012) "Judicata raises $2M from Peter Thiel, Keith Rabois and others to give lawyers better research and analytics tools", *TechCrunch*, available at: https://techcrunch.com/2012/12/11/judicata-raises-2m-from-peter-thiel-keith-rabois-and-others-to-give-lawyers-better-research-and-analytics-tools/.

[9] See notes 3 and 7 on the Cadiet report.

[10] Evidence of this is the recent arrival in Paris of Rocket Lawyer, a Californian firm that targets lawyers. https://www.carrieres-juridiques.com/actualites-et-conseils-emploi-juridique/lancement-de-rocketlawyer-france-la-plateforme-qui-rend-le-droit-accessible-a-tous/1488

[11] On France's intention to become a leader in AI see: https://www.aiforhumanity.fr/.

a direct impact on the attractiveness of our law. On the subject of making legal professionals anonymous in open data, the Cadiet report has stated, "*In a context of competition, especially between systems of civil law and of common law […] maintaining the names of legal professionals is evidence of seriousness and reliability, and, therefore, of the credibility and attractiveness both of French law and its state jurisdictions.*" It is up to France to maintain the prestige of its legal model internationally by backing efforts for providing access to it.

In summary, AI can help develop tools for the justice system that will broaden access to the law and facilitate the work of lawyers. Where is the law google? In France, it — along with the search engine Doctrine.fr and legaltechs in data analytics — is waiting for court decisions to become open data.