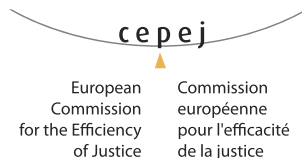# EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ)

European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment

Adopted at the 31st plenary meeting of the CEPEJ (Strasbourg, 3-4 December 2018)

c e p e j

European Commission for the Efficiency of Justice

Commission européenne pour l'efficacité de la justice

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

EUROPEAN COMMISSION FOR THE EFFICIENCY OF JUSTICE (CEPEJ)

European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment

Adopted at the 31st plenary meeting
of the CEPEJ (Strasbourg, 3-4 December 2018)

Council of Europe

Contents

Introduction

Acknowledging the increasing importance of artificial intelligence[1] (AI) in our modern societies, and the expected benefits when it will be fully used at the service of the efficiency and quality of justice, the CEPEJ formally adopts the 5 fundamental principles entitled "European Ethical Charter on the use of AI in the judicial systems and their environment".

The Charter is intended for public and private stakeholders responsible for the design and deployment of artificial intelligence tools and services that involve the processing of judicial decisions and data (machine learning or any other methods deriving from data science).

It also concerns public decision-makers in charge of the legislative or regulatory framework, of the development, audit or use of such tools and services.

The use of such tools and services in judicial systems seeks to improve the efficiency and quality of justice, and should be encouraged. It must, however, be carried out with responsibly, with due regard for the fundamental rights of individuals as set forth in the European Convention on Human Rights and the Convention on the Protection of Personal Data, and in compliance with other fundamental principles set out below, which should guide the framing of public justice policies in this field.

Judicial decision processing by artificial intelligence, according to their developers, is likely, in civil, commercial and administrative matters, to help improve the predictability of the application of the law and consistency of court decisions, subject to compliance with the principles set out below. In criminal matters, their use must be considered with the greatest reservations in order to prevent discrimination based on sensitive data, in conformity with the guarantees of a fair trial.

Whether designed with the aim of assisting in the provision of legal advice, helping in drafting or in the decision-making process, or advising the user, it is essential that processing is carried out with transparency, impartiality and equity, certified by an external and independent expert assessment.

_____

1. For the definition of artificial intelligence, see the Glossary appended.

Application of the Charter

The principles of the Charter should be subject to regular application, monitoring and evaluation by public and private actors, with a view to continuous improvement of practices.

In this respect, it is desirable that a regular review of the implementation of the principles of the Charter be made by these actors, explaining, where appropriate, the reasons for non-implementation or partial implementation, accompanied by an action plan to introduce the necessary measures.

The independent authorities mentioned in the Charter could be responsible to periodically assess the level of endorsement of the Charter's principles by all actors, and to propose improvements to adapt it to changing technologies and uses of such technologies.

The five principles of the Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment

1 **Principle of respect for fundamental rights:** ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights.

2 **Principle of non-discrimination:** specifically prevent the development or intensification of any discrimination between individuals or groups of individuals.

3 **Principle of quality and security**: with regard to the processing of judicial decisions and data, use certified sources and intangible data with models elaborated in a multi-disciplinary manner, in a secure technological environment.

4 **Principle of transparency, impartiality and fairness**: make data processing methods accessible and understandable, authorise external audits.

5 **Principle "under user control"**: preclude a prescriptive approach and ensure that users are informed actors and in control of the choices made.

<table>
<tr><td>**1**</td><td>**Principle of respect for fundamental rights:** ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights</td></tr>
</table>

▬▬ The processing of judicial decisions and data must serve clear purposes, in full compliance with the fundamental rights guaranteed by the European Convention on Human Rights (ECHR) and the Convention on the Protection of Personal Data (Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, ETS No. 108 as amended by the CETS amending protocol No. 223).

▬▬ 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).

▬▬ They should also be used with due respect for the principles of the rule of law and judges' independence in their decision-making process.

▬▬ Preference should therefore be given to *ethical-by-design*[2] or *human-rights-by-design* approaches. This means that right from the design and learning phases, rules prohibiting direct or indirect violations of the fundamental values protected by the conventions are fully integrated.

---

2. The ethical choice is made upstream by the program designers and is therefore not left to the user.

2

Principle of non-discrimination: specifically prevent the development or intensification of any discrimination between individuals or groups of individuals

━━━ 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.

━━━ Particular care must be taken in both the development and deployment phases, especially when the processing is directly or indirectly based on "sensitive" data. This could include alleged racial or ethnic origin, socio-economic background, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data, health-related data or data concerning sexual life or sexual orientation. When such discrimination has been identified, consideration must be given to corrective measures to limit or, if possible, neutralise these risks and as well as to awareness-raising among stakeholders.

━━━ However, the use of machine learning and multidisciplinary scientific analyses to combat such discrimination should be encouraged.

3

Principle of quality and security: with regard to the processing of judicial decisions and data, use certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment

━━ Designers of machine learning models should be able to draw widely on the expertise of the relevant justice system professionals (judges, prosecutors, lawyers, etc.) and researchers/lecturers in the fields of law and social sciences (for example, economists, sociologists and philosophers).

━━ Forming mixed project teams in short design cycles to produce functional models is one of the organisational methods making it possible to capitalise on this multidisciplinary approach.

━━ Existing ethical safeguards should be constantly shared by these project teams and enhanced using feedback.

━━ Data based on judicial decisions that is entered into a software which implements a machine learning algorithm should come from certified sources and should not be modified until they have actually 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.

━━ The models and algorithms created must also be able to be stored and executed in secure environments, so as to ensure system integrity and intangibility.

<table>
<tr><td>**4**</td><td>**Principle of transparency, impartiality and fairness:** make data processing methods accessible and understandable, authorise external audits</td></tr>
</table>

■ A balance must be struck[3] between the intellectual property of certain processing methods and the need for transparency (access to the design process), impartiality (absence of bias)[4], 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.

■ The first option is complete technical transparency (for example, open source code and documentation), which is sometimes restricted by the protection of trade secrets. The system could also be explained in clear and familiar language (to describe how results are produced) by communicating, for example, the nature of the services offered, the tools that have been developed, performance and the risks of error. Independent authorities or experts could be tasked with certifying and auditing processing methods or providing advice beforehand. Public authorities could grant certification, to be regularly reviewed.

----

3. Of interest in this connection is the suggestion made on page 38 of the Council of Europe's MSI-NET study on "Algorithms and Human Rights": "The provision of entire algorithms or the underlying software code to the public is an unlikely solution in this context, as private companies regard their algorithm as key proprietary software that is protected. However, there may be a possibility of demanding that key subsets of information about the algorithms be provided to the public, for example which variables are in use, which goals the algorithms are being optimised for, the training data and average values and standard deviations of the results produced, or the amount and type of data being processed by the algorithm." Or even the suggestions appearing on page 117 of the "AI for humanity" report drafted by Mr Cédric Villani, a member of the French National Assembly as part of a mission assigned to him by the Prime Minister of the French Republic: "The auditors may be satisfied with simply checking the fairness and equity of a programme (doing only what is required of them), by submitting a variety of false input data, for example, or by creating a large quantity of system user profiles according to precise guidelines." In addition, there are also the statements in the report by the House of Lords, "AI in the UK: ready, willing and able?", paragraphs 92, 96-99.

4. In this connection, it is interesting to note the solutions generally considered for ensuring the neutrality of algorithms in the above-mentioned report by the House of Lords (paragraphs 114, 115, 116, 119, 120): more diverse datasets, more diversity and multidisciplinary approaches, more auditing of aspects such as data processing and the manner in which the machine is constructed.

<table>
<tr><td>**5**</td><td>**Principle "under user control":** preclude a prescriptive approach and ensure that users are informed actors and in control of their choices</td></tr>
</table>

▬ User autonomy must be increased and not restricted through the use of artificial intelligence tools and services.

▬ Professionals in the justice system should, at any moment, be able to review judicial decisions and the data used to produce a result and continue not to be necessarily bound by it in the light of the specific features of that particular case.

▬ 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.

▬ Generally speaking, when any artificial intelligence-based information system is implemented there should be computer literacy programmes for users and debates involving professionals from the justice system.

Appendix I

In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data

prepared by Mr Xavier Ronsin, First President of the Court of Appeal of Rennes, scientific expert (France),
and
Mr Vasileios Lampos, principal research fellow at the Computer Science department of University College London (UCL), scientific expert (United Kingdom),

and with the contribution of Ms Agnès Maîtrepierre, judge, member of the Consultative Committee of the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data of the Council of Europe (France)

The following experts also contributed to fintune the Study:

Mr Francesco Contini, Senior Researcher at the Research Institute on Judicial Systems – National Research Council (IRSIG-CNR), Bologna (Italy)

Mr Francesco De Santis, Professor of Human Rights Procedures, University of Naples (Italy)

Mr Jean Lassègue, philosopher and epistemologist, research fellow at the Centre National de Recherche Scientifique (CNRS) and associate researcher at the Institut des Hautes Etudes sur la Justice (IHEJ) (France)

Ms Dory Reiling, Honorary Senior Judge, Independent Expert on Information Technology and Judicial Reform (Netherlands)

Mr Aleš Završnik, Chief Researcher at the Institute of Criminology, Associate Professor at the Faculty of Law, University of Ljubljana (Slovenia) and EURIAS Researcher 2017-18 at Collegium Helveticum in Zürich (Switzerland)

INTRODUCTION

1.   The wave of digital transformation in our societies still has an uneven effect on the judicial systems of Council of Europe member States. Many European countries seem to have already developed an extremely advanced approach to using practical applications (in terms of both technology and legal support), while for others, this is still just an emerging issue and the focus is solely on effective IT management.

2.   Among the technologies at work in this great digital transformation, artificial intelligence (AI) appears to be both the most spectacular and the most striking. In the United States, "robot lawyers" are already at work and seem to converse in natural language with humans. Legal tech start-ups specialising in the design of new legal services offer new applications to legal professions, mainly lawyers, legal services and insurers, allowing in-depth access to judicial information and case law. These private companies even aim to predict judges' decisions with "predictive justice" tools, although we will see that this may not be the best description for them[5].

3.   An initial examination of this phenomenon, however, prompts us to differentiate between this commercial discourse and the reality of the use and deployment of these technologies. For the time being judges in the Council of Europe member States do not seem to be making any practical and daily use of predictive software. Local tests[6] and academic work[7] have been carried out to explore the potential of these applications, but they have not yet been applied on a wide scale. The initiative for the development of these tools comes largely from the private sector, whose clientele so far has been made up mostly of insurance companies, lawyers and legal services wanting to reduce legal uncertainty and the unpredictability of judicial decisions. Nevertheless, public decision-makers are beginning to be increasingly solicited by a private sector wishing to see these tools – which are sometimes "beta" versions, i.e. they will evolve over time – integrated into public policies.

---

5.   See in particular the frame opening Chapter 9.
6.   For example, the Douai and Rennes Courts of Appeal in France conducted a three-month trial in 2017 with a software programme labelled "predictive" by a panel of judges.
7.   Work on a sample of 584 decisions of the European Court of Human Rights: Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preoţiuc-Pietro, Vasileios Lampos, "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", published on 24 October 2016, [Online], https://peerj.com/articles/cs-93/

4.   In the line of the thought process initiated in its "Guidelines on how to drive change towards Cyberjustice",[8] the CEPEJ proposes to provide public decision-makers and justice professionals with keys for a better understanding of the "predictive justice" phenomenon.

5.   The first task will be to clarify the questions relating to the intrinsic nature of these mass case-law data processing systems, along with their technical and theoretical limitations. These aspects have not often been mentioned in the debate on this subject in the judicial sphere but they are very well known and discussed by specialists in these technologies (mathematicians, statisticians and computer scientists) and deserve some attention.

6.   Secondly, this document will analyse the benefits and risks of these tools. While their supporters highlight their assets in terms of transparency, predictability and standardisation of case-law, their critics point to the limitations and the reasoning bias of the software currently on the market. The inherent risks in these technologies may even transcend the act of judging and affect essential functioning elements of the rule of law and judicial systems, to which the Council of Europe is particularly attached.

7.   These include principles such as the primacy of law. The effect of these tools may be not only to provide incentives but to be almost prescriptive, creating a new form of normativity, which could supplement the law by regulating the sovereign discretion of the judge, and potentially leading, in the long term, to a standardisation of judicial decisions based no longer on case-by-case reasoning by the courts, but on a pure statistical calculation linked to the average compensation previously awarded by other courts.

8.   There is also a need to consider whether these solutions are compatible with the individual rights enshrined in the European Convention on Human Rights (ECHR). These include the right to a fair trial (particularly the right to a natural judge established by law, the right to an independent and impartial tribunal and equality of arms in judicial proceedings) and, where insufficient care has been taken to protect data communicated in open data, the right to respect for private and family life.

9.   While taking these issues into consideration, the document highlights the great potential of AI to help legal professionals in their work. There is no doubt that some AI applications which are still under development or testing, such as those designed to improve legal research, could be very useful in rendering processing of the judicial workload faster and more efficient. The

---

8. See in particular paragraph 51 of the document CEPEJ(2016)13, Guidelines on how to drive change towards Cyberjustice.

document highlights these positive examples and generally advocates the use of AI by legal professionals according to their needs, provided that due regard is shown for the individual rights guaranteed by the ECHR and Council of Europe standards, particularly in criminal matters. Far from being a simple instrument for improving the efficiency of judicial systems, AI should strengthen the guarantees of the rule of law, together with the quality of public justice.

10.   Lastly, the document suggests means of monitoring this phenomenon in the form of an ethical charter, emphasising the need for a cautious approach to the integration of these tools into public policies. It is essential that any public debate involves all the stakeholders, whether legal professionals, legal tech companies or scientists, to enable them to convey the full scope and possible impact of the introduction of artificial intelligence applications in judicial systems and devise the ethical framework in which they must operate. Subsequently, this debate could go beyond a pure "business" framework, involving citizens themselves and hence contributing to some extent to general computer literacy, as has been achieved in Canada.[9]

1. State of the use of artificial intelligence algorithms in the judicial systems of Council of Europe member States

In 2018, the use of artificial intelligence algorithms in European judicial systems remains primarily a private-sector commercial initiative aimed at insurance companies, legal departments, lawyers and individuals.

11.   The use of AI in the judicial field appears to be quite popular in the United States, which has invested in these tools in a fairly uncomplicated way, both in civil and criminal matters.[10]

12.   Pinpointing instances of AI algorithm initiatives in the judicial systems of Council of Europe member States is a more difficult task, as most of the initiatives come from the private sector and are not often integrated into public policies.

13.   The question of the use of AI in judicial systems was dealt with in a specific online survey, launched in April 2018 for representatives of the

---

9. Declaration of Montreal, available at https://www.declarationmontreal-iaresponsable. com/demarche, accessed 16 July 2018.
10. See COMPAS algorithms or tools such as RAVEL LAW or ROSS chatbot

CEPEJ member States and civil society. The response level was relatively low and did not allow clear trends to be identified. Some private operators did not seem very receptive to this survey and the members of the CEPEJ, who belong for the most part to ministries of justice or higher councils of justice, were able to quote only the tools currently used by the public sphere.

14.   As a result, the inventory below is only partial and is based solely on research conducted by experts and the secretariat using publicly available literature.[11]

15.   Classifications can be made according to the service offered. The involvement of AI can vary greatly according to the applications. For illustrative purposes, the main categories are as follows:

► Advanced case-law search engines

► Online dispute resolution

► Assistance in drafting deeds

► Analysis (predictive, scales)

► Categorisation of contracts according to different criteria and detection of divergent or incompatible contractual clauses

► "Chatbots" to inform litigants or support them in their legal proceedings

16.   **Latvia** stated that it was exploring the possibilities of machine learning for the administration of justice. The main purpose would be to process court statistics to draw up provisional estimates of human and financial resources to be allocated.

17.   Other activities carried out by legal tech companies have not been included in this classification because they involve little or no artificial intelligence processing: some sites offer access to legal information, "cloud" solutions, electronic signatures, etc.

18.   A non-exhaustive list of legal services making use of artificial intelligence in their operations is set out below:

―――

11. See summary bibliography in Appendix IV – substantial contributions from Benoît Charpentier as well as Giuseppe Contissa and Giovanni Sartori (https://media.wix.com/ugd/c21db1_14b-04c49ba7f46bf9a5d88581cbda172.pdf) and Emmanuel Barthe (http://www.precisement.org/blog/Intelligence-artificielle-en-droit-derriere-la-hype-la-realite.html#nb14) (French only)

Software	State	Type
Doctrine.fr	France	Search engine
Prédictice	France	Analysis (except criminal cases)
Case Law Analytics	France	Analysis (except criminal cases)
JurisData Analytics (LexisNexis)	France	Search engine, Analysis (except criminal cases)
Luminance	United Kingdom	Analysis
Watson/Ross (IBM)	USA	Analysis
HART	United Kingdom	Analysis (criminal, risk of reoffending)
Lex Machina (LexisNexis)	USA	Analysis

2. Overview of open data policies relating to judicial decisions in the judicial systems of Council of Europe member States

The availability of data is an essential condition for the development of AI, enabling it to perform certain tasks previously carried out by humans in a non-automated manner. The more data available, the more AI is able to refine models improving their predictive ability. An open data approach to judicial decisions is therefore a prerequisite for the work of legal tech companies specialising in search engines or trend analysis ("predictive justice").

Processing of these data raises a number of issues, such as changes in the formation of case-law and protection of personal data (including the names of professionals).

19.   Computer-raised data are said to be the "oil" of the 21st century as their use and cross-referencing are producing a whole new wealth. Even though some stakeholders and authors dispute this argument, the global successes of the digital industry over recent decades have confirmed the enormous growth potential of this field of activity.

20.   The quantification of human activities, now on a global scale, could not fail to touch on the data produced by the public sector. This is what has prompted the movement to open up public data, based on much older imperatives which are the founding principles of our constitutional states.

21.   The major change in recent years has been brought about by the emergence of downloadable public data (open data), notably in the context of the "Partnership for Open Government" (OGP). The OGP is a non-governmental organisation bringing together nearly 70 member States (including many of the Council of Europe member States) with representatives of civil society and digital giants. The aim of this openness is to improve the transparency of public activities, encourage citizens in the development and assessment of public policies and guarantee the integrity of public service and those who perform it by processing considerable amounts of information, organised into databases (big data).

2.1. Definition of open data on judicial decisions

22.   First of all, let us redefine the notion of open data before dealing with the question of the impact of allowing open data on judicial activity. Firstly, there is often confusion between access to information and access to data (more precisely, access to information in the form of database)[12].

23.   A certain amount of public information, requiring wide publicity, is already disseminated using information technology. In France, the government site Légifrance.fr is the main online source of certified public information, comprising not only legislative and regulatory texts but also case-law and information on appointments to public posts. This unitary information, although available on the Internet, differs completely from direct access to data organised and included in a database that can be downloaded and processed by a computer.

24.   Open data therefore only involves the dissemination of "raw" data in structured computer databases. These data, aggregated in whole or in part with other structured sources, constitute what we call big data. The Council of Europe Convention 108 Consultative Committee defines big data as "the growing technological ability to collect, process and extract new and predictive knowledge from great volume, velocity, and variety of data. In terms of data protection, the main issues do not only concern the volume, velocity, and variety of processed data, but also the analysis of the data using software to extract new and predictive knowledge for decision-making

_____
12. The data are meaningless letters and numbers. Information is data included in a context. It is the context that gives meaning to the data. We can guess that 2005 is a year, but without context, we do not know. But in "in 2005, we completed 3 projects" the context gives meaning to the number. Therefore, 'open data' is not data in the sense of the definition, but information. Similarly, large data are also large amounts of information, not data.

purposes regarding individuals or groups. For the purposes of these Guidelines, the definition of Big Data therefore encompasses both Big Data and Big Data analytics".[13]

25.　**As this definition shows, open data should not be confused with their means of processing.** Some of the discourse on this issue actually relates to processing carried out by various advanced methods which are generally defined as data science. Predictive justice using artificial intelligence, advanced search engines applying extremely precise criteria and legal robots are all algorithmic applications which are fed with data but have nothing to do with the policy of open data itself.

26.　However, this policy must be examined in the light of the possibilities it offers for further processing, whatever its nature. If certain data are filtered upstream, taking account for example, of the need for confidentiality and respect for privacy, subsequent risks of misuse appear to be reduced.

2.2. State of development of open data on judicial decisions in Council of Europe member States and consequences for the development of case law

27.　What is the situation of the Council of Europe member States as regards open data on judicial decisions? The 2016-2018 CEPEJ evaluation cycle focused for the first time on the question of court decisions being provided in open data, for which some AI processing is used. The issue of data anonymisation or pseudonomysation within the European data protection framework[14] provided by the General Data Protection Regulation (GDPR, EU Regulation 2016/679) and Council of Europe Convention No. 108 was the subject of a specific question designed to identify the measures implemented by member States and observers in this particularly sensitive area.

---

13. T-PD(2017)1, Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data.
14. General Data Protection Regulations (DGPS, EU Regulation 2016/679, and Council of Europe Convention No 108 for the Protection of Individuals with regard to Automatic Processing of Personal Data).

Figure 1. Table showing CoE member States having introduced open data on judicial decisions in 2016 (Q62-4 of the evaluation scheme)
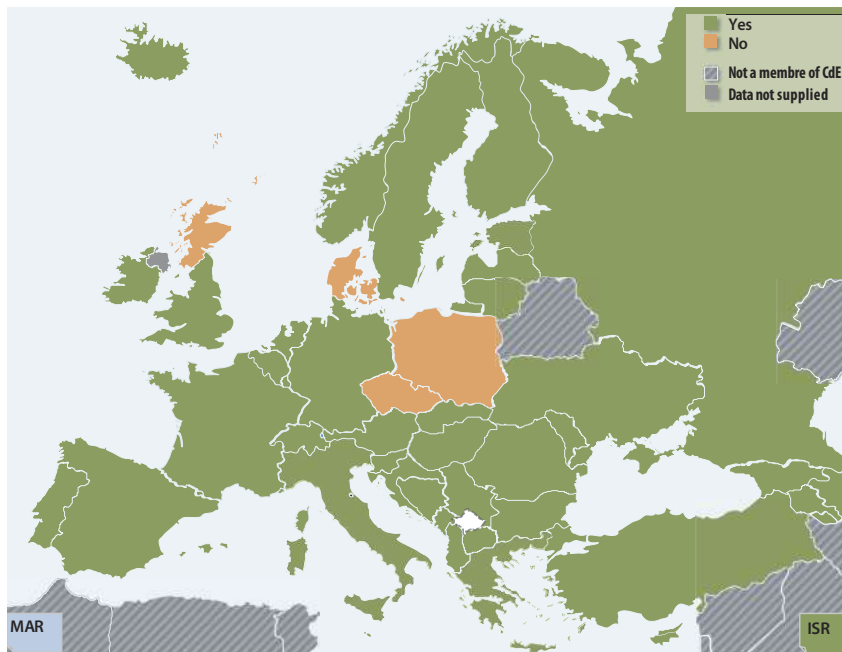
28.   Of all the States and observers surveyed, only 5 declared that they had not implemented an open data policy for judicial decisions in 2016. While this response rate should be put into perspective, since some answers confused public access to decisions with open data (**Armenia**, **Belgium**, **Bosnia and Herzegovina**, **Luxembourg**, **Russian Federation**, **Spain**, **Turkey**), it reveals, on the one hand, a desire for transparency on the part of European judicial institutions and, on the other hand, a desire on the part of many countries to make court decisions public and thus make it possible to deal with them later using AI tools. This also requires efforts by the institutions concerned, since a number of technical measures must be put in place to this end. In **France** specifically, some administrative case law is already available for download on the site *data.gouv.fr* (see below).

29.   With regard to the protection of personal data, 23 countries declared that they are pseudonymising[15] at least some types of disputes (e.g. personal status, family status) by erasing data making the parties or witnesses identifiable (names, addresses, telephone numbers, identity numbers, bank account numbers, tax numbers, health status, etc.). This work appears to be the responsibility of judicial personnel (e.g. **Israel**, **Republic of Moldova**) or public officials (e.g. **Bosnia and Herzegovina**, **Spain**). **Bosnia and Herzegovina** and **Hungary** on the other hand stated to publish the names of professionals.

30.   However, there is a real difficulty in measuring the impact of open data on the efficiency and quality of justice. As indicated above, the initiative to re-use these data is essentially private, targeting a professional clientele (lawyers, legal departments), and an exclusively intergovernmental activity is probably not the best means of fully identifying such positive results.

31.   The situation in France is representative of the questions raised by this approach and reveals a number of the issues at stake. First of all, it is important to underline that France enacted legislation in 2016 imposing a compulsory framework for the open data dissemination of decisions on its courts.

32.   Articles 20 and 21 of the Law for a Digital Republic[16] broke with the previous logic[17] of selecting which decisions from judicial and administrative courts and tribunals were to be disseminated if they were "of particular interest". Under the new French law, however, the opposite principle that everything is publishable has been set, except in specific cases identified by law (for judicial decisions) and with due regard for the privacy of the persons concerned. Provision is made, however, for judicial and administrative decisions to be published only after an analysis has been made of the risk of re-identification of the persons involved.

------

15. As defined by the T-PD in its "Guidelines on the protection of individuals with regard to the processing of personal data in a world of big data" T-PD(2017)1, pseudonymisation refers to the processing of personal data "in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

16. This law was adopted in order to bring French law into line with Directive 2013/37/EU of the European Parliament and of the European Council of 26 June 2013, which in turn amended the Council Directive of 17 November 2003 on the re-use of public sector information (the "PSI Directive").

17. Article R433-3 of the Code of judicial organisation

33.   There are numerous advantages to wide dissemination, promoted in particular by the Court of Cassation at two conferences in October 2016 and February 2018. They include greater awareness of judicial activity and case law trends, the increased quality of a justice system that knows it is being observed and the creation of a completely new factual reference base.

34.   This theoretical desire to "regulate" case-law through digital leverage raises however a number of general questions that are relevant to all countries considering a similar approach. Firstly, it should be placed in the context of some of the principles set out by the European Court of Human Rights in cases of differences in domestic case-law. The Court clearly emphasises the need to balance legal certainty, which makes decisions more predictable, against vitality in judicial interpretation.[18]

35.   This desire is also tempered by several aspects, of technical nature in first place:

a. **Collection of all judicial decisions eligible for publication is not necessarily well co-ordinated between all levels of courts:** some of the business applications in European courts have not been designed for this, particularly as regards first instance decisions, and some countries will have to set up new processes for collating judgments if they want collection to be exhaustive;

b. Despite some promising developments, **a fully effective automated post-identification mechanism that can prevent any risk of identification or re-identification has not yet been devised**;

And equally substantive aspects:

c. **Thought should be given to the transformation of the very logic of the production of case-law.** What is the value of the "standard" resulting from the number of decisions given on a specific matter? Does this "standard" add to the law? If so, is this a new source of law?

d. **Should judges be asked to provide additional explanations for decisions to deviate from this norm?**[19] This question is not insignificant and does not prompt the same reply in all judicial systems. In French civil law,[20] "judges settle disputes in accordance with the rules of law applicable thereto". In the French context, if they were asked to justify their decisions by providing all the reasons for which they deviated

---

18. *Greek Catholic parish Lupeni and Others v. Romania* [GC]. No. 76943/11, 29/11/2016, § 116.
19. Eloi Buat-Menard and Paolo Giambiasi, «La mémoire numérique des décisions judiciaires», Dalloz Reports, 2017, p. 1483. (French only)
20. Article 12 of the French Code of Civil Procedure

from the supposed majority case-law trend on how to resolve the dispute (while complying with the relevant rules of law), would this not be tantamount to removing them from office? Judges would not only settle disputes in accordance with the rules of law but also in accordance with case law trends derived from statistics compiled by a digital tool (which could also be biased or developed without external control by a private operator, see sections 6.1 and 9 below). The question that arises goes far beyond a specific national feature, relating more broadly to the place of AI tools in judges' decision-making processes. Note that this analysis is limited to a processing of case-law by algorithms and in no way refers to the role of case-law in the sources of law or the authority of case-law precedents, which, moreover, are well established principles at European level.[21]

e. **Moreover, would it not be the case that if norms were established according to the majority trend, judicial decisions would be rendered uniform, and no longer be ordered according to the hierarchy of the courts from which they emanate, disregarding the significance of the decisions of supreme courts, which are the guarantors of the uniform interpretation of law in many European States?** What would the relationship be between norms and case-law? Since they derive from the majority, would they become a criterion for these courts when determining their own case-law, which they would in turn have to justify when they deviate from the majority view?

f. **Finally, isn't there a risk of court decisions being written down according to a reproductive logic?** While court decisions are likely to evolve according to the evolution of a normative framework (national, European or international), the case law available for reference (for example, from supreme courts and European courts) or the socio-economic context, would not the norm resulting from the majority become a standard to which judges would be encouraged to refer without question, with an induced effect of the excessive standardisation of judicial decisions?

36. **Doubts might also be raised about the consequences for users, who are supposed to benefit directly from the transparency of activities:** will they really benefit from the publication of all judicial decisions on the

---

21. The ECJ has stated, on the conditions governing the liability of a Member State for the content of a decision of a supreme national court, that "an infringement of Community law will be sufficiently serious" and must give rise to compensation for the damage "where the decision concerned was made in manifest breach of the case-law of the Court in the matter" (ECJ, Case C-224/01, Koebler, §56).

Internet or are they not more likely to find themselves overwhelmed by the mass of decisions, without necessarily deriving qualitative benefits from them, for lack of legal knowledge and the critical apparatus to analyse them?

37.   Paradoxically enough, one may wonder of open data be in position of delivering meaningful information since it will certainly make it possible to download a considerable set of raw data with a simple click on a link, but the meaning will remain totally obscure for the majority of citizens.

38.   The first requirement for open data to be transparent and informative is for third parties to be able to use them for analysis and development. An economic model in which public case law data, which are the subject of work by judicial authorities in order to make them technically "readable" by the AI and compliant with the legal requirements for the protection of personal data, would be processed free of charge by the private sector and then sold by to the courts, professionals and citizens, therefore deserves to be questioned and even criticised.

2.3. Protection of personal data in open data policies for judicial decisions

2.3.1. The names of parties and witnesses

39.   In order to strike a fair balance in the digital age between the need to make judicial decisions public and respect for the fundamental rights of parties or witnesses, their names and addresses must not appear in published decisions, particularly in view of the risk of misappropriation and re-use of such personal information and the particular sensitivity of the data likely to be contained in the decisions.[22]Automated processes can be used to systematically to conceal such information.

40.   Other identifying information may also be obscured (for example, telephone numbers, e-mail addresses, dates of birth, children's given names, rare given names, nicknames and place names). In terms of personal data protection principles, this concealment amounts to a simple pseudonymisation of the data, not complete anonymisation. The volume and variety of information contained in court decisions, combined with the growing ease

-----

22. ECHR, *Z. v. Finland* No. 22009/93, §§95 et seq. and the European Commission's Green Paper on public sector information in the information society (COM(1998)585) ("If special precautions are not taken, case-law databases, which are legal documentation instruments, can become information files on individuals if these databases are consulted to obtain a list of court judgments on a specific individual rather than to find out about case-law").

of cross-referencing with other databases, makes it impossible, in practice, to guarantee that the person concerned cannot be re-identified. In the absence of such a guarantee, these data cannot be qualified as anonymous and are therefore subject to personal data protection rules.

41.   Some especially sensitive items of personal data warrant particular attention, as provided for in Article 6 of Convention 108. This applies to data revealing ethnic or racial origin, political opinions, trade union membership, religious or other beliefs, physical or mental health or sex life, which are considered intimate details.

42.   Court decisions may contain other, very varied, types of personal data that fall into this category of sensitive data. Courts dealing with criminal matters are particularly likely to process sensitive data such as those on criminal proceedings and convictions. All this sensitive data therefore deserves special vigilance. Their mass dissemination would present serious risks of discrimination, profiling[23] and violation of human dignity.

2.3.2. The names of professionals, including judges

43.   Obviously, knowing how a judgment will be arrived at is an essential element for lawyers in predicting the outcome of a case, and they believe that knowing one's judge is sometimes almost as important as knowing the law. They have long tried to make comparisons between panels of judges, more or less empirically, so as to give better advice to clients dealing with a particular judge or panel of judges.

44.   This method was sufficient when a lawyer was only speaking before a limited number of courts, but the gradual loosening of local restrictions on the bar in many countries and the freedom to move and work within the European Union make it reasonable for any national or even European lawyer to want to know the case-law of each national or European jurisdiction in which he is likely to plead in full detail.

45.   We cannot therefore exclude the possibility that, in the future, highly useful, and hence very expensive, machine learning applications will be much more effective than the experience and "good sense" of litigation lawyers working through cases in the traditional way. The use of such applications could further accentuate the distortion of competition and inequality

------

23. Profiling is defined in section 4 of the GDPR. It is processing an individual's personal data in order to analyse and predict his/her behavior or his/her situation, such as determining his/her performance at work, financial situation, health, preferences, lifestyle habits, etc.

of arms between law firms that have or have not used such "predictive" case-law analysis software.

46. There is a real risk that, for the sake of such a competitive advantage, the principle of a fair trial established by law will be undermined. The possibility of judge profiling through cross-referencing of public and private data could allow private companies and their lawyers to engage in even more forum shopping practices. This tactic has already been observed for a long time in the United States and in France for press offences and violations of privacy in the press, where plaintiffs have already been known to choose the court which appears to award the highest amounts of damages and interest.

47. In addition, much of the discourse on this subject confuses open data with the need to publish a certain amount of public information. As a result it is sometimes argued that the names of professionals should appear in open data for the sake of publicity and transparency.

48. However, the provision of computerised case-law data is a totally separate issue from the principles of publication of original or certified copies of decisions. The objective of open data is to allow automated processing of case-law under a low-cost operating licence. As stated previously, this provision is made in the form of a comprehensive computer database, which is opaque and not directly understandable to citizens.

49. This provision obviously does not meet the need to publicise the names of professionals who have contributed to a specific decision. It should be noted that, in law, mentioning the names of judges in judicial decisions is a common obligation for member States, in connection with the principle of public trial set out in Article 6(1) of the European Convention on Human Rights, in order to ensure the objective impartiality of judges (who must be identifiable, and lawfully appointed and assigned to the duties they perform) and compliance with procedural rules (e.g. publicity and collegiality).[24]

50. The answer to the question of the legitimacy or not of publishing the names of professionals[25] in open data therefore has nothing to do with the obligation to publish the names of professionals in decisions. Rather, it

_____

24. ECHR, *Vernes v. France*, No. 30183/06 as regards the identification of the judges who rendered the decision, *Pretto v. Italy*, No. 7984/77 for the publication of the judgment, *Kontalexis v. Greece*, No. 59000/08, § 38, *DMD GROUP, a.s, v. Slovakia*, No. 19334/03, § 66, *Miracle Europe KFT v. Hungary*, No. 57774/13, § 58 for the right to a court established by law in conjunction with the right to an impartial judge.

25. These questions may also apply to Rechtspfleger who make judicial decisions and to clerks, assistants to the judge mentioned in the composition of the formation of the court (albeit to a lesser extent).

seems that the challenge lies in reconciling often conflicting requirements: making public activities transparent by allowing citizens to know and evaluate their judges, on one hand, while protecting the privacy of professionals (whose functions should not limit their fundamental guarantees in this field), on the other hand. There are rigorous challenges where it comes to guaranteeing the impartiality of judges and even judicial institutions as a whole, which open data policies are actually designed to meet[26]. What practical measures can be taken to protect them from potential attempts at destabilisation which cross-reference judges' personal data included in databases with other sources (social networks, commercial sites) to try to identify hypothetical political, religious and other biases?

51.   These questions do not arise in the same form everywhere in Europe and depend on the specific features of the judicial system concerned (and on the nature of the judiciary's career management body), the collegial nature or not of the judgment and the level of court concerned. In Switzerland, for example, where judges are elected, publication is a guarantee of transparency and social responsibility of judges vis-à-vis citizens and political groups. This information is already available in online search engines (which are not strictly speaking open data).[27]

52.   Nor do these questions arise in the same form depending on the level of jurisdiction. The value of characterising the case-law of lower court judges may not be the same as for judges of supreme courts or international courts. For example, the European Court of Human Rights authorises searches for judgments by the names of the judges members of the decision panel, but does not allow the calculation of statistics relating to a particular judge.[28] On the other hand, in countries where the judicial bodies are unfamiliar with the practice of dissenting opinions (existing within this international court), it may seem unfair to assign a judge personal responsibility for a decision which he voted against during deliberation in a collegial court.

53.   These debates were well defined by a study mission conducted in France by Professor Loïc Cadiet. The conclusions of the mission remain limited since they do not recommend prohibiting publication but reserving it for certain types of litigation and ruling it out for others (for example, for

---

26. See *ECHR Previti v. Italy,* No. 45291/06, §§ 249 et seq., which recalls the principles of objective impartiality of the judge.

27. See the example of the Swiss Federal Court, whose case-law can be downloaded: https://www.bger.ch/fr/index/juridiction/jurisdiction-inherit-template/jurisdiction-recht.htm; or, for the cantons: http://ge.ch/justice/dans-la-jurisprudence (Canton of Geneva for example).

28. The decisions of the European Court of Human Rights are collegial. Publications include any dissenting opinions.

specialised criminal matters). The possibility of publishing only the names of supreme court judges was proposed, although it was conceded that this might result in a "one-way trip".

54.   As it stands, a simple precautionary principle could be applied in order to assess the interest, by type of litigation and degree of jurisdiction, in publishing the names of professionals in a downloadable database. Likewise, we cannot rule out the possibility of judicial institutions themselves or authorised third parties exploiting this information outside the open-data context to find out about case-law.

55.   Publication of this sort would in any event still need to be examined in light of applicable international European rules on the protection of personal data, such as those of Convention No. 108 and the European Union's GDPR and Directive 680/2016.

3. Operating characteristics of artificial intelligence (machine learning) applied to judicial decisions

Natural language processing and machine learning are the two techniques at the heart of processing of judicial decisions using artificial intelligence.

In most occasions, the objective of these systems is not to reproduce legal reasoning but to identify the correlations between the different parameters of a decision (for example, in a divorce claim the length of marriage, the income of spouses, the existence of adultery, the amount of the benefit pronounced, etc.) andn, through the use of machine learning, to infer one or more models. Such models would then be used to "predict" or "foresee" a future judicial decision.

The purpose of this chapter is to clarify the intrinsic nature of software described as "predictive" – sometimes "beta" versions, i.e. in a development phase – both in terms of its potential and its limitations. It is intended to provide a simple explanation of machine-learning algorithms, which are at the heart of the automated analysis of case-law.

3.1. The theoretical functionalities of "predictive justice" software

56.   By way of an introduction, we should briefly review the features promised by "predictive" software. They propose to establish the probabilities of the success (or failure) of a case before a court. These probabilities are established through the statistical modelling of previous decisions using methods from two broad computer science domains: natural language processing

and machine learning. These modelling approaches are often referred to as AI; in reality, these are "weak" AI (see the glossary, page

57.   It should be made clear straight away that the term AI is debated by experts as it leads to many ambiguities. The term AI has now entered our everyday language to describe a diverse range of sciences and technologies that allow computers to beat the best champions in the game of Go,[29] to drive a car, to converse with humans, etc. Researchers prefer to identify the different applictions through the exact technologies underlying them, including machine learning, and sometimes refer to all of these highly specialised AI resources as "weak" (or "moderate") AIs. These are distinguished from an ultimate – still totally theoretical – goal of creating a "strong" AI, i.e. a self-learning machine capable of automatically comprehending the world in general, in all its complexity.

58.   In relation specifically to justice, predictive justice systems are designed for use by legal departments, insurers (both for their internal needs and for their policyholders) as well as lawyers for them to anticipate the outcome of litigation. Theoretically, they could also assist judges in their decision-making.

59.   They provide a graphic representation of the probability of success for each outcome of a dispute based on criteria entered by the user (specific to each type of dispute). These systems claim to be capable to calculate the likely amount of compensation distributed by the courts.

3.2. The practical functioning of artificial intelligence: statistical machines constructing models based on the past

60.   A distinction must be made from the outset between what is a "prediction" and what is a "forecast". Prediction is the act of announcing what will happen (*prae*, before – *dictare*, say) in advance of future events (by supernatural inspiration, by clairvoyance or premonition). Forecasting, on the other hand, is the result of observing (*aiming*, seeing) a set of data in order to envisage a future situation. This abuse of language and the spread thereof seems to be explained by a transfer of the term from the "hard" sciences, where it refers to a variety of data science techniques derived from mathematics, statistics and game theory that analyse present and past facts to make hypotheses about the content of future events.

---

29. https://www.nature.com/articles/nature16961

61. It should also be noted that the logic of the operation of predictive justice software is essentially based on either generative (commonly referred to as Bayesian) or discriminative methods which eventually try to estimate the current or future range of values of a variable (e.g. the outcome of a trial) from the analysis of past examples.

62. Generally speaking, it is also important to keep in mind the anthropomorphic notion that computing machines are intelligent and that their designers have managed to slip a mind inside their mechanisms.[30] Unfortunately, this idea still permeates many analyses of predictive justice that lend these devices immediate or future capabilities for the near replication of human intelligence. This context, fuelled every day by a further series of revolutionary AI advances,[31] therefore leads us all to approach these predictive tools with a certain dose, conscious or not, of mysticism, sometimes stating that what is not entirely possible today will inevitably be possible tomorrow.

63. The "strong" AIs of science-fiction literature do not exist. This type of AI, which would be equipped not only with intelligence but also with conscience, remains purely fictional. The machine learning systems currently being developed are described as "weak" AIs and **are capable of extracting complex patterns and learning from large volumes of data efficiently and often with high levels of predictive accuracy.**

64. To overcome any instinctive or fantasised considerations, these information processing and analysis technologies must be understood on the basis of the following three concepts.

- **AI is not a single, homogeneous object**: it is actually an assemblage of sciences and techniques (mathematics, statistics and computer science) capable of processing data to design very complex computer processing tasks.

- **AI engines do not produce intelligence per se but proceed using an inductive approach:** the idea is to associate in a nearly automated way a set of observations (inputs) with a set of possible results (outputs) using various preconfigured properties. Specifically for predictive justice, the engine builds links between the different lexical groups composing judicial decisions. These groups are correlated between those identified

---

30. Dominique Cardon, *A quoi rêvent les algorithmes, nos vies à l'heure des big data*, La République des idées, Editions du Seuil, 2015, p. 58.
31. "AlphaZero : l'IA de DeepMind apprend et maîtrise le jeu d'échecs en moins de 4 heures", *Génération NT*, article published on 8 December 2017, [Online], https://www.generation-nt.com/alphazero-ia-deepmind-maitrise-echecs-4-heures-actualite-1948919.html (page accessed on 14 December 2017).

at the input stage (facts and reasoning) and those at the output stage (the operative part of the decision) then classified.

- The reliability of the model (or function) built strongly depends on the quality of the data used and the choice of machine learning technique.

65. It is worth returning briefly to the ancestors of these systems – expert systems – which for their part relied on processing rules written by a computer scientist. Expert systems (ES) developed rapidly in the late 1980s and 1990s, especially in medicine and finance.[32] These systems were able to answer specialised questions and reason using known facts, executing predefined encoding rules in an engine. Despite the success of ESs, such as Deep Blue against Garry Kasparov in a series of chess games in 1997, these systems ended up failing, notably because they were unable to interpret "the infinite variety of situations and contexts"[33] and became ineffective beyond 200 to 300 encoding rules, both in terms of execution performance and maintenance (the reasoning followed by the system became almost impossible to apprehend for its designers).

Fig. 2: The old expert systems were programmed with rules reproducing the logic of legal reasoning

Rules/Templates

Data → **Expert system** → Results

----
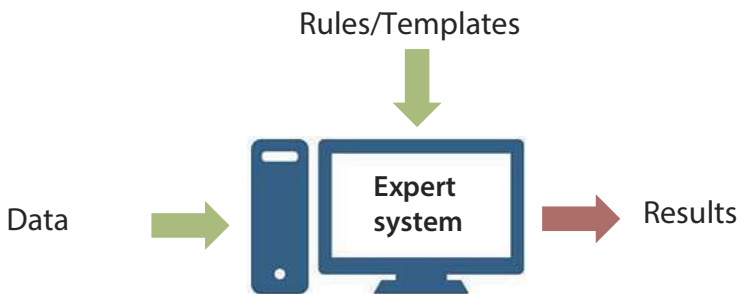
32. An example is High Frequency Trading (HFT), which is a type of financial transaction carried out at high speed by software based on algorithms. In 2013 and in Europe, 35 % of transactions were already carried out with HFT. In the USA, HFT represents 70 % of trading volumes on the equity market. In view of the automation of transactions, the fall is even faster in the event of crashes, as was the case during the 2007 financial crisis.
33. Hubert Dreyfus, *What Computers Still Can't Do. A Critique of Artificial Reason*, Cambridge, The MIT Press, 1992 cited by Dominique Cardon in *A quoi rêvent les algorithmes, nos vies à l'heure des Big Data*, p.59.

66.   Today, the idea is no longer to write reasoning rules mirroring human reasoning, as with older expert systems,[34] but to let machine learning systems themselves identify existing statistical models in the data and match them to specific results.

Fig. 3: Machine learning alone produces models by automatically searching for correlation results

Rules/Templates

Data   Machine learning   Results
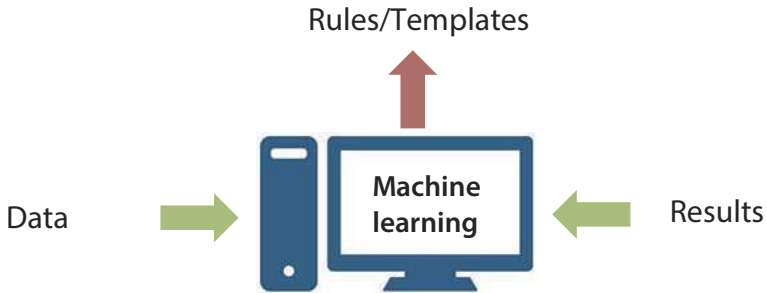
67.   In doing so, these systems do not reproduce or model reasoning (such as legal reasoning). For example, today's online translators do not carry out abstract reasoning. They infer a probable estimate of the best match between groups of lexical structures and translations already done. Users' actions obviously contribute to the improvement of the match search, but the machine learning algorithm does not actually perform a translation by understanding the meaning of the sentences processed.[35]

68.   This example shows how the approach of lawyers and researchers can be different: a lawyer will seek to understand the effectiveness of pre-existing rules, which the researcher can only provide through the language of his science, and this is sometimes limited to interpreting thousands of parameters from a machine learning system.

69.   Machine learning is a field of computer science in which computer programmes learn from experience. Algorithms make a machine perform a

_____

34. Written processing rules could be based on specific code in different programming languages such as LISP or editors to model rules. See for example Radboud Winkels, "CLIME : Un projet de développement de serveurs juridiques intelligents", in Danièle Bourcier, Patricia Hasset and Christophe Roquilly (eds.), *Droit et intelligence artificielle,* 2000, Romillat, p. 59.

35. Li Gong, "La traduction automatique statistique, comment ça marche ?", *Interstices.info*, published on 29 October 2013, [Online], https://interstices.info/jcms/nn_72253/la-traduction-automatique-statistique-comment-ca-marche (page accessed on 14 December 2017).

training process, like a child learning in its environment. In summary, these learning techniques may or may not be supervised by a human. The most prominent category of machine learning is currently that of reinforcement learning: the machine alone reaps virtual "rewards" if action produces the expected result in a given environment. Machine learning methods include neural networks (or their more complex version known as deep learning), which have been quite widely publicised in view of their autonomy and their quite striking applications, such as the one succeeding in obtaining high scores on old Atari 2600 video games[36] (based solely on pixel positioning and scores). Other practical applications for these technologies are already affecting our daily lives and are beginning to appear in the professional world of justice.[37]

70. However, these algorithms remain highly specialised in one particular task and present discernment problems when faced with chaotic situations or with insufficient data to allow prediction (such as the actual understanding of natural language). In the social sciences, to which law and justice belong, failure would even appear inevitable in the absence of a convincing model of cognition. For Pierre Lévy, artificial intelligence is in fact content to provide a heterogeneous toolbox (logical rules, formal syntaxes, statistical methods, neural or socio-biological simulations, etc.) that does not offer a general solution to the problem of mathematical modelling of human cognition.[38] Thus, real predictive learning should in reality be based on a good systemic representation of the world, which AI researcher Yann LeCun believes is a fundamental scientific and mathematical issue, not a question of technology.[39]

71. Moreover, the uniqueness of current big data processing systems is that they do not attempt to reproduce our cognition model but produce context statistics on an unprecedented size of data, without any real guarantee of excluding false correlations.[40]

_____

36. https://www.nature.com/articles/nature14236
37. Artificial intelligence by IBM Watson providing a service to the medical field and providing a search tool called "Ross", presented as a virtual lawyer – Roos, "Do more than humanly possible" [Online], http://rossintelligence.com (page accessed on 14 December 2017).
38. Pierre Lévy, "Intelligence artificielle et sciences humaines", *Pierre Levy's blog*. 8 October 2014. [Online], https://pierrelevyblog.com/2014/10/08/intelligence-artificielle-et-sciences-humaines/ (page accessed on 30 December 2017).
39. Yann LeCun, "Qu'est-ce que l'intelligence artificielle", research published on the Collège de France website, [Online], https://www.college-de-france.fr/media/yann-lecun/UPL4485925235409209505_Intelligence_Artificielle__Y._LeCun.pdf (page accessed on 14 December 2017).
40. Dominique Cardon, *op. cit.*, p.60.

72.   While it is commonly accepted that statistical and probabilistic models are improved by increasing the data feeding them, some mathematicians have warned of the risks of increasing false correlations (i.e. links between factors having absolutely no causal link) in big data. The mathematicians Cristian Sorin Calude and Giuseppe Longo point to the risk of a deluge of false correlations in big data: the larger a database used for correlations, the greater the chances of finding recurrent patterns and the greater the chances of making errors.[41] What may appear as regularities for an AI (recurrent links between different data, concepts, contexts or lexical groups) may actually be random. Even if the argument of the two mathematicians should not be generalised too hastily, they note that in certain vast sets of numbers, points or objects, regular random patterns appear and it seems impossible to distinguish them algorithmically from patterns revealing causalities.[42]

73.   Lastly, algorithms, whatever their current level of sophistication, still automatically boil down to the interactions established in the learning phase and hence to their past. The content and quality of the data flows used in the composition of the calculations is therefore fundamental in understanding the results obtained and in identifying possible analytical biases. There are many challenges here again since, in an exhaustive approach, the analysis of the greatest possible amount of data relating to an activity will produce results whose meaning has to be clarified with regard to all the factors that have had an influence. In a more restrictive approach in which incoming data is sampled, risks will also arise from the trade-off biases required to select one data over another.

4. Can artificial intelligence model legal reasoning in advance?

Artificial intelligence circumvents the difficulties encountered with older expert systems: they do not attempt to manually mirror legal reasoning, whose reproduction is not in itself an objective for them. Machine learning leads to categorisations between the different parameters identified by the designers or those discovered by the machine.

74.   Continental legal systems are far removed from the ideal of rationality that embodied, for example, the 1804 Civil Code in France. There are a multitude of sources that do not fit together perfectly and which relate to

____

41. Cristian Sorin Calude, Giuseppe Longo, "Le déluge des corrélations fallacieuses dans le big data", in Bernard Stiegler (dir.) (2017), *La toile que nous voulons*, FYP editions, p.156.

42. Theory of "Ramsey", Ronald L. Graham, Joel H. Spencer – "Ramsey Theory", *Scientific American*, vol.263, No.1, July 1990, p112-117 cited by Cristian Sorin Calude, Giuseppe Longo, "Le déluge des corrélations fallacieuses dans le big data", *ibid*.

a set of rules whose meaning remains undetermined, which the legal theorist Herbert L. A. Hart called the "open texture of law".[43] Common law systems, although considered more economically efficient because they adapt dynamically to new legal needs,[44] are also evolving and offer no more legal certainty. Legal rules therefore do not evolve in a linear fashion, distinguishing them from empirical laws (those of the "exact sciences"), where each new rule generally complements the previous ones and does not invalidate a complete set of reasoning.

75.   It should be stressed that, faced with this general complexity of legal systems, the old IT expert systems quickly reached their limits once 200 to 300 logical rules were nested. The division of the law into rules of production was not sufficient to provide a valid representation of the body of knowledge and methods that guide a lawyer.

76.   Consequently, it has proved just as impossible to model the reasoning of judges on a computer as it has to model a positive legal system. As the theory of law has highlighted, judicial reasoning is above all a matter of assessment and interpretation, of the proven and relevant facts of a case, of the applicable rules of law (textual or precedents) – the meaning of which remains, as has been said, indeterminate –,[45] and of the subjective interpretation by judges of the concept of equity, which should undergo new changes in Europe with the requirement of a proportionality review encouraged by the European Court of Human Rights.

77.   Xavier Linant de Bellefonds has stressed that the complexity of the law lies in its teleological and contentious nature: two coherent arguments can lead to different judgments according to two different priorities.[46]

78.   This is because the famous legal syllogism is more a way of presenting legal reasoning than its formal translation. It does not reflect the full reasoning of the judge, which is in fact made up of a multitude of decision-making factors, cannot be formalised a priori, and is sometimes based on his discretion: what are the relevant facts? Are these facts proven? Which rule applies to them? What is the meaning of this rule with regard to the case to be decided? Which source should prevail between a range of conflicting

_____

43. Herbert L. A. Hart (1976), *Le concept de droit*, Saint-Louis university departments, Brussels.
44. https://www.contrepoints.org/2014/08/15/177160-common-law-contre-droit-civil-lexperience-francaise-de-lancien-regime
45. On these questions, Michel Troper (2001), *La théorie du droit, le droit, l'Etat*, PUF, Paris, spec. p. 69-84.
46. Xavier Linant de Bellefonds (1994), "L'utilisation des systèmes experts en droit comparé", *Revue internationale de droit comparé*, Vol. 46, No. 2, p. 703-718.

sources? The overall consistency of judicial decisions is never achieved and is more a matter of an a posteriori account which judges use in their reasoning, having more of a function of convincing themselves of the validity of a specific solution than of describing strictly and objectively all the stages that resulted in the decision made.

79.   However, this work of interpretation is precisely what machine learning tecnhiques do not do – and do not attempt to do – today, since they carry out, as we have seen, automated processing where the correlation of vast amounts of information is supposed to replace the understanding of the real causalities of a decision. They make not attempt to formalise legal reasoning but hope that the models captured by them can anticipate the likely decisions of a judge in similar situations.

80.   The results achieved by AIs are in reality unrelated to the question of the legal conformity of a particular solution and cannot discriminate between legal and illegal arguments.

81.   A review of the work of the University College of London (UCL) on the case-law of the European Court of Human Rights confirms this diagnosis. The UCL study assumed that a simple automatic learning model could predict the outcome of a case with 79 % accuracy for that particular court. The machine learning model proved to be more accurate with regard to the descriptive part of the facts of the decisions studied than with regard to the reasoning relating to the application of the Convention to the case in question.[47]

82.   The examination of the terms correlated with the finding of a violation (with a positive weighting) and those correlated with a non-violation of the European Convention on Human Rights (with a negative weighting) sheds unambiguous light on the mechanics at work, which is in no way comparable with legal reasoning.

---

47. Work on a sample of 584 decisions of the European Court of Human Rights: Nikolaos Aletras, Dimitrios Tsarapatsanis, Daniel Preoţiuc-Pietro, Vasileios Lampos, "Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective", published on 24 October 2016, [Online], https://peerj.com/articles/cs-93/ (page accessed on 14 December 2017).

Fig. 4: Illustration of UCL's work – Theoretical weight allocated to words or terms according to their link with findings of violation or non-violation of the European Convention on Human Rights

Positive State Obligations: + 13,50

Treatment by state officials: + 10,20

Detention conditions: + 11,70

Enforcement of domestic judgments and reasonable time: + 11,70

VIOLATION

Evidence : - 15,20

Sentencing: - 17,40

Previous violation of article 2 : - 11,40

Property rights and company claims: - 9,08

NON VIOLATION
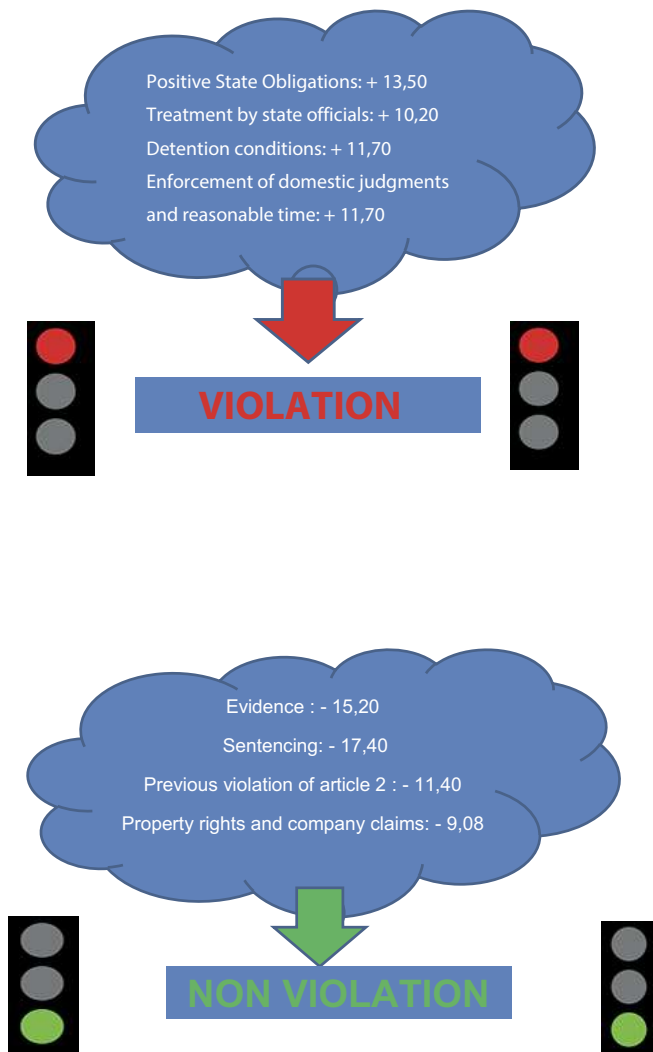
83.   The choice of the relevant facts and their interpretation constitute one of the elements of the judge's decision. In other words, the UCL study was in reality only able to produce a probability with lexical material largely derived from the reasoning and motivation of the judge and not with that assembled by the applicant on the basis of frequencies alone. Their AI established thus a high probability of correspondence between groups of words and a decision that had already been formalised and could only give rise to a limited number of possible outcomes. Under no circumstances can it alone reproduce the reasoning of European judges nor, above all, predict an outcome, on the basis, for example, of a future applicant's raw account before the Strasbourg court, whose application will be subject to a very strict admissibility examination (nearly 70,356 applications were declared inadmissible or struck from the list in 2017)[48] based largely on the application of standards of assessment (importance and seriousness of the complaint, etc.) leaving considerable leeway in decision-making.

5. Can AIs explain judges' behaviour in retrospect?

An a posteriori explanation of a judge's behaviour, in particular the revelation of bias, would require all the potentially causative factors to be identified through an interpretative framework and a contextualised analysis. The fact that, statistically, childcare is more often entrusted to mothers than fathers does not demonstrate a bias on the part of judges but reveals the need to mobilise different disciplines from the social sciences to shed light on this phenomenon.

84.   From a scientific point of view, explaining a phenomenon or, as far as we are concerned, a piece of human behaviour, amounts to determining the causal mechanisms that led to this behaviour using a certain amount of contextual data.

85.   This requires, in a very schematic way, the preliminary constitution of an interpretative framework, itself derived from the repeated observation of this type of event or behaviour in the presence of certain factors or elements. The interpretative framework is made up of the hypotheses or points of view adopted by the different social science disciplines. This is an additional analytical step that can be fed into algorithms, but which they cannot perform alone.

86.   Some legal tech companies went further and thought they could identify possible personal biases of judges and feed suspicions of bias. The open data of the names of certain presiding judges of administrative courts and

_____

48. Source: 2017 Statistical analysis of CEDH, January 2018.

administrative courts of appeal in France have made it possible to develop an indicator of the rejection rate of appeals against obligations to leave French territory taken by the administrative authorities. Some commentators have argued fervently that the alleged impartiality of judges was therefore cast into doubt by artificial intelligence.

87.   But can such interpretations really be achieved on the basis of an algorithmic processing of court decisions? For there to be personal bias in judges' decision-making processes (differing from their personal and public statements in the case concerned) their behaviour, or in this case their decision, needs to be determined by their personality traits, opinions or religion. However, as has been said, such a causal explanation cannot simply be deduced from the probabilistic result provided by algorithms. On the contrary, it requires additional analytical work in order to isolate, among the many correlated factors (including the identity of members of the panel of judges), those that are truly causative. For example, the fact that a family court statistically decides more often that children should live with their mother does not necessarily reflect the judge's bias in favour of women, but rather the existence of psychosocial, economic and even cultural factors specific to the jurisdiction, such as the working time of each of the parents, their income, the local availability of collective childcare, whether or not the child is in school, whether or not one of the parents is in a new relationship or even simply the lack of interest by either parent to take care of a young child.

88.   Similarly, decisions on expulsion from a country given by an administrative court located near a large detention centre cannot be compared fairly with those of a court that deals with such disputes only occasionally.

89.   In addition, regardless of the court's location, the issue of the case-law of a single duty judge who only occasionally deals with a certain type of litigation, but who uses (or ignores) the case law of his colleagues, is particularly interesting and may legitimately raise the question of equality of citizens in judicial proceedings. However, the focus must remain on the remodelling or preservation of the collegial nature of the judicial system rather than classification or stigmatisation via machine learning tools.

90.   What can be deduced from the personality of the president of a panel of judges in a collegial court, when his name is the only nominative information visible in open administrative court decisions?

91.   Furthermore, how can we account for two distinct philosophical and cultural approaches to judicial decisions, whereby, in some European countries, including France, there is a culture of precedent and a detailed knowledge by judges of the factual databases of all 1st and 2nd instance decisions

(Ariane database) in the field of administrative justice, while other countries or systems favour the intellectual independence of each court, along with a desire to deal with each situation on a case-by-case basis?

92.   Providing an accurate explanation of a court decision therefore requires a much more detailed analysis of the contingent data in each case and the applicable rules of law, rather than sustaining a vain hope that the mass of links will make sense.

6. How is AI to be applied in civil, commercial and administrative justice?

The state of development of machine learning techniques does not allow today to reach reliable results regarding the "prediction" of judicial decisions. On the other hand, their application in the field of civil, commercial and administrative justice is to be considered for the creation of scales or the pre-litigation resolution of disputes online, when a later appeal to the judge remains possible.

93.   As we can see, the first question raised by such use of artificial intelligence is not so much whether it is beneficial or harmful, desirable or otherwise, but whether the algorithms proposed can achieve the type of result sought. The conclusions from the experiments conducted in the Douai and Rennes appeal courts in France clearly demonstrate that in the presence of a thought-provoking discourse promoting a product of AI can hide unacceptable design flaws and totally erroneous analysis results.

94.   Regardless of the quality of the software tested, anticipating judges' decisions in civil, commercial and administrative matters would appear to be a potentially desirable benefit, albeit sometimes for very different reasons, both for those responsible for judicial public policy and for private legal professionals.

95.   Whatever the legal tradition of the country, legal uncertainty, i.e. the risk of having one's legal claim validated or rejected, prompts the desire to be able to quantify these factors using these new technological applications.

96.   Lawyers (or a company's legal department) see the possibility of using this technology to provide their clients with better informed advice by empirically and systematically assessing the chances of a procedure's success, as well as encouraging the conclusion of transactions that, if necessary, make it possible to avoid a long and costly trial. Some insurers already offer the use of predictive systems to their clients to evaluate the merits of their business.[49]

_____
   49. The Predictive software is offered to Allianz policyholders, for example.

97.   At the same time, public decision-makers see this as an opportunity to better regulate the flow of new proceedings through the courts and provide themselves with a lever to reduce judicial operating costs. This is thought to encourage litigants to use alternative dispute resolution methods (conciliation, mediation or arbitration).[50]

98.   The approach that already exists in many judicial systems of harmonising decisions in many matters by using scales (divorce, dismissal, compensation for personal injury) could be revitalised through a probabilistic or actuarial approach.[51] Alternative online dispute resolution services have even been created to help assess the amount of compensation for small disputes, *inter alia*. However, these interesting approaches are not unbiased and must not deprive citizens of access to a judge or call into question the adversarial principle.

Experiments conducted in France

At the initiative of the Ministry of Justice, the two courts of appeal in Rennes and Douai agreed to test predictive justice software on various litigation appeals in spring 2017, which in reality was an analysis of civil, social and commercial decisions of all French courts of appeal.

Although these internal and exhaustive case law data had already been available to them free of charge for many years (JURICA database), the Ministry made them specially available to the publishing company when they offered to assess the value of a quantified (innovative) analysis of the sums allocated by the two courts, in addition to a geographical classification of the discrepancies noted for similar applications and trials.

The stated objective of the software was therefore to create a decision-making tool in order to reduce, if necessary, excessive variability in court decisions, in the name of the principle of equality of citizens before the law. The result of the experiment, contradictorily debated between the two courts of appeal, the Ministry of Justice and the legal tech company who designed the product unfortunately stated the absence of added value of the tested version of the software for the work of reflection and decision-making of the magistrates.

More significantly, software reasoning biases were revealed that led to aberrant or inappropriate results due to confusion between mere lexical occurrences of judicial reasoning and the causalities that had been decisive in the judges' reasoning.

---

50. In this respect see information report No. 495 (2016-2017) prepared on behalf of the Senate Law Commission, and tabled on 4 April 2017, by Senator Philippe Bas
51. With respect to the actuarial approach, Case Law Analytics' offer announces more of a risk assessment than a prediction of a solution to a dispute.

6.1. A new computing tool to calculate scales

99.   The procedures for calculating scales in various civil matters (for example, compensation for bodily injury, compensatory allowance and severance pay) appear to be considerably enhanced when combined with other processing techniques, subject to numerous design measures and uses (performative effect).[52]

100. It is important to highlight what Jean-Paul Jean, chair of the CEPEJ evaluation working group, described as the qualitative challenge during a conference on open data held in 2016 in France: the procedure carried out by machine learning or any other method of processing should use certified originals, whose integrity has been verified, and which have been enriched to distinguish the important from the insignificant.[53]

101. The other risk already encountered by the designers of these tools is that of "data-snooping", namely selecting upstream-only data that is significant to predetermined analysis grids, for example by excluding from samples decisions that lend themselves poorly to correlations of linguistic sequences through machine learning or any other method (for example, decisions without presentation of the litigation or little reasoning).

102. But if we calculate a scale, are we not already doing a little prediction? The borderline may seem relatively sketchy unless we clearly distinguish the aim of the process: the objective here is not to deliver prescriptive information but to provide information about a state of affairs.

103. Subject to these methodological and operational precautions, the scales are available for handling certain disputes and are a powerful tool for harmonising case-law. Scales were previously calculated on more or less limited samples of decisions, but AI tools make it possible to survey more decisions and are liable to produce more accurate scales, applying a standard and hence giving more weight to the outcomes.

_____

52. The performative or self-realisation effect is the risk that a system will produce the same output progressively by influencing the producers of input information; this effect is often mentioned with regard to judicial scales which, when informed by decisions based on these scales, tend to be representative only of themselves.

53. J-P. Jean, "Penser les finalités de la nécessaire ouverture des bases de données de jurisprudence", Conference of 14 October 2016 held at the Court of Cassation, https://www.courdecassation.fr/IMG///Open%20data,%20par%20Jean-Paul%20Jean.pdf, page accessed on 17 March 2018.

6.2. Online dispute resolution

104. All European courts face to different degrees repetitive low-value civil litigation. The idea of facilitating the procedure through information technology and/or outsourcing them from the courts is quite widely shared. Great Britain, the Netherlands and Latvia are examples of countries that have already implemented or are about to implement these types of more or less automated solution[54]s. For cross-border claims, the European Union has set up, by means of Regulation n°524/2013, a common framework available on the Internet (European small claims).

105. However, the scope of these online dispute resolution (ODR) services seems to have gradually extended. They have gone from restricted online services to alternative dispute resolution measures before the complaint is brought before the court, and are now being introduced

increasingly into the court process itself, to the point of offering electronic court services'.[55] They do not only concern low-value disputes, but also tax disputes or disputes relating to social security services, or divorce proceedings.

106. For those who advocate such solutions, which are of interest to a number of legal professions and the private sector, access to justice could be significantly improved by a broad solution combining ODR and AI (or at least expert systems, see section 3 above for the distinction). The idea is to take complainants through an automated diagnosis of the dispute by putting a number of questions, which are then processed by the machine, resulting in proposals for a solution. The work of the Cyberjustice de Montréal laboratory, which assembles the various pre-litigation and litigation phases into a computer-based process for low-intensity disputes (for example, small claims courts in Quebec), is a good illustration of hybridisation.[56] According to the designers, there are clear benefits in terms of efficiency and quality.

---

54. See to this end the online dispute resolution available in the UK – https://www.judiciary. gov.uk/wp-content/uploads/2015/02/Online-Dispute-Resolution-Final-Web-Version1.pdf. See also the system PAs in the Netherlands, which issues automated decisions based on previously granted permits, and which has given rise to litigation at the national level and before the ECJ: there are two cases (c-293/17 en c-294/17) brought to the Council of State of the Netherlands (farmers/ nature conservation against the Netherlands) to determine whether or not a system (Programme Regulating Nitrogen) is allowed to decide whether or not farmers a.o. are violating the Habitat directive. Recently the Court of Justice of the European Union in Luxembourg has answered on the requests for a preliminary ruling concerning these joined cases (ECLI:EU:C:2018:882)

55. Darin Thompson, "Creating new pathways to justice using simple artificial intelligence and online dispute resolution", Osgoode Hall Law School of York University.

56. http://www.cyberjustice.ca/projets/odr-plateforme-daide-au-reglement-en-ligne-de-litiges/

107. But on what basis would any compensation proposed by such a system be calculated? What method? Does the algorithm process information fairly? Is the proposal intended to be discussed on an adversarial basis with the help of a trained and certified third party? Is access to a judge always possible? Some authors even see the widespread use of these dispute resolution methods as a new manifestation of digital "solutionism", i.e. the systematic use of technologies to try to solve problems that do not necessarily fall within their scope.[57] It should also be noted that in Europe, a more protective regulatory framework that is binding on the member States has recently been put in place: Article 22 of the Data Safety Monitoring Plan explicitly provides for persons to be able to refuse to be the subject of a decision based exclusively on automated processing, with certain exceptions.[58]

---

57. Evgeny Morozov, "Pour tout résoudre cliquez ici", FYP editions, cited by David Larrousserie, "Contre le 'solutionnisme' numérique", Le Monde, 6 October 2014, https://www.lemonde.fr/sciences/article/2014/10/06/contre-le-solutionnisme-numerique_4501225_1650684.html

58. Article 22(1) of EU Regulation 2016/679: "The data subject shall have the right not to be subject to a decision based solely on automated processing"; exceptions are provided for (such as the data subject's consent) but "suitable measures to safeguard the data subject's rights and freedoms and legitimate interests " must be implemented by the controller, including "the right of the data subject to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision".
See, in the same sense, the Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, as amended by the Protocol adopted in May 2018 when the latter enters into force. Article 9(1)(a) provides the principle that "Everyone has the right not to be subject to a decision affecting him significantly, which shall be taken solely on the basis of automatic processing of data, without his point of view being taken into account. Notwithstanding this principle of prohibition, Article 9(2) states that "paragraph 1(a) shall not apply if the decision is authorised by a law to which the controller is subject and which also provides for appropriate measures to safeguard the rights, freedoms and legitimate interests of the data subject".
The explanatory report states (§75): "It is essential that any person likely to be subject to a purely automated decision should have the right to challenge that decision by effectively putting forward his point of view and his arguments. In particular, the data subject must have the possibility to prove the possible inaccuracy of personal data before their use, the inadequacy of the profile to be applied to his particular situation or other factors which will have an impact on the outcome of the automated decision. This is particularly the case when the application of an algorithmic reasoning, by leading to the limitation of a right, the refusal of a social benefit or the evaluation of their borrowing capacity on the sole basis of the software, has the effect of stigmatizing individuals. However, the data subject may not exercise this right if the automated decision is provided for by the law to which the controller is subject, which provides for appropriate measures to safeguard the rights and freedoms and legitimate interests of the data subject".
See also Article 9(1)(c) of the modernised Convention 108, which provides for the right of the data subject "to obtain, at his request, knowledge of the reasoning underlying the processing of data, when the results of such processing are applied to him". The explanatory

108. The potential benefits of an ODR system, its degree of integration into a complete judicial process (from pre-litigation to actual litigation) and the almost decisive role of AI in the execution of the process must therefore be properly assessed on a case-by-case basis.

109. ODR already offers upstream knowledge of judicial processes. Its role is clearly to contribute to the implementation of conciliation, mediation and arbitration services outside the courtroom. These services can also be used during contentious proceedings under the supervision of judges before they decide on the outcome of disputes based on the merits (for some disputes this phase is considered compulsory).

110. On the other hand, the actual contribution of AI should be assessed. Is it only a question of using machine learning to establish indicative scales or prescribe a solution? Is it really AI that is being used or an expert system or just the logical rule chain? In any event, it should be possible to combine these systems with the requirements of transparency, neutrality and loyalty.[59]

111. Finally, it is also necessary to examine the way in which complainants are encouraged to use the system: is there potential confusion in the very name of what is on offer? If one speaks of a court, it must be the form of organisation defined by the European Convention on Human Rights and not simply a private justice institution with the mere appearance of state justice[60]. Is recourse to a judge clearly possible? In the Netherlands, private health insurance contracts seem to provide automatically for recourse to an ODR before any legal action is brought..

112. The CEPEJ Mediation Working Group (CEPEJ-GT-MED), launched in 2018, has offered its first thoughts on the contribution of information technology towards alternative dispute resolution methods. The CDCJ

———

(cont'd) dreport of the modernised Convention (§77) states: "Data subjects have the right to obtain knowledge of the reasoning underlying the data processing, including the consequences of this reasoning and the conclusions which may have been drawn from it, in particular when using algorithms for automated decision-making, in particular in the context of profiling. For example, in the case of a credit rating system, borrowers have the right to know the logic behind the processing of their data that leads to the decision to grant or refuse credit, rather than simply being informed.
of the decision itself. Understanding these elements contributes to the effective exercise of other essential safeguards such as the right of objection and the right of appeal to the competent authority". This "must obtain useful information concerning the underlying logic" is also found in the GDR (Article 13(1)(f); Article 14(2)(g); Article 15(1)(h)).

59. Charlotte Pavillon, "Concerns over a digital judge", nrc.nl, https://www.nrc.nl/nieuws/2018/01/19/zorgen-om-populaire-digitale-rechter-a1588963
60. Scarlett-May Ferrié, Algorithmes tested against fair trial, document consulted on http://lexis360.fr, downloaded on 09/07/2018,§ 27-38

is currently doing in-depth work on ODRs to identify the potential of these tools but also any problem points amounting to possible violations of Articles 6, 8 and 13 of the European Convention on Human Rights.

6.3. The main guarantees to be reaffirmed in civil, commercial and administrative proceedings

Right of access to a court

113. The provision of online dispute resolution tools should not affect the right of access to a court within the meaning of Article 6, even if this right is not absolute and lends itself to implicit limitations.[61] In civil matters, for example, every litigant has the right to submit to a court any dispute relating to his "civil rights and obligations" heard by a court.[62] In 2015, the Parliamentary Assembly of the Council of Europe adopted a resolution on "Access to justice and the Internet: potential and challenges" in which it called to ensure that "parties engaging in ODR procedures retain the right to access a judicial appeal procedure satisfying the requirements of a fair trial pursuant to Article 6 of the Convention".[63]

Adversarial principle

114. It seems imperative to make a certain amount of quantitative information (for example, the number of decisions processed to obtain the scale) and qualitative information (origin of decisions, representativeness of selected samples, distribution of decisions between different criteria such as the economic and social context) accessible to citizens and, above all, to the parties to a trial in order to understand how scales have been constructed, to measure their possible limits and to be able to debate them before a judge.

Equality of arms

115. The use of technological means should not cause imbalances between parties, since the use of digital means could indeed facilitate proceedings

---

61. Art.6 §1 « 1. everyone is entitled to a fair and public hearing .. by an independent and impartial tribunal… who shall decide (...) the merits of any criminal charge against it..; for the limitations see Deweer c. Belgique, § 49 ; Kart c. Turquie [GC], § 67.
62. CEDH, Golder c. Royaume-Uni, §§ 28-36
63. Resolution 2054 (2015) of the Parliamentary Assembly of the Council of Europe (PACE), 10 November 2015, http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-EN.asp?fileid=22245&lang=en

for certain operators (institutions, companies with means, computer literate persons) and, on the contrary, pose difficulties for certain population types that are more uncertain or less familiar with computers. It is important that no individuals are left alone in front of their screens, and that they are informed that they can seek legal advice and are assisted where necessary.

Impartiality and independence of judges

116. It has been posited that the norm derived from the majority trend referred to above in section 2.2. may have indirect effects on the independence and impartiality of the judiciary, particularly in systems where the independence of the judiciary is not fully achieved. In these systems, we cannot rule out the risk that such norms will place indirect pressure on judges when decisions are taken and prompt their approval, or that the executive will monitor those who depart from the norm.

Right to counsel

At the beginning of this chapter, we mentioned the advantages derived from the application of predictive justice tools for lawyers and, in particular, the possibility of providing their clients with better informed advice by empirically and systematically assessing the chances of a procedure's success. However, let us imagine a case where the chances of success for the litigant are extremely poor: could this affect the lawyer's decision to assist his client? Professional practice should aim to minimise the risk that persons requiring legal advice may ultimately be deprived of it.

7. Issues specific to criminal justice: prevention of offences, risk of recidivism and assessment of the level of danger

Even they are not specifically designed to be discriminatory, the use of statistics and AI in criminal proceedings has shown a risk of prompting the resurgence of deterministic doctrines to the detriment of doctrines of individualisation of the sanction, which have been widely acquired since 1945 in most European judicial systems.

117. The use of AI science and technology in criminal matters poses specific challenges as its application may reflect some current public debates about the alleged predictability of offending behaviour. However, this debate seemed to have been thoroughly settled for some thirty years in a number of European countries. In Italy for example, Article 220, paragraph 2, of the Code of Criminal Procedure expressly rules out the use of an expert opinion

to establish habitual or professional criminal features, the tendency to commit a crime, the character and personality of the accused and, in general, the psychological qualities of the accused, regardless of the pathological causes. In France, for example, the doctrine of "new social defence" developed by Marc Ancel was the foundation of criminal law: instead of a merely punitive and deterministic approach, a system of social rehabilitation was introduced to prevent the commission of an offence by avoiding the conditions for crime. This approach is shared by a number of European criminal policy instruments that focus on the objectives of re-educating and reintegrating offenders.[64]

118. Criminal justice tools should therefore be designed in accordance with these fundamental principles of rehabilitation, [65]including the role of the judge in the individualisation of the sentence, based on objective elements of personalities (training, employment, regular medicals and social care) without any other form of analysis than that carried out by specifically trained professionals, such as probation officers. Big data analytics techniques could be used by these professionals to centralise and collect information on the person accused of a crime or misdemeanour, which could then be stored by various institutions and agencies and would then need to be examined by a judge, sometimes within a very short time frame (for example in the context of accelerated trial procedures).

7.1. Tools used by investigative authorities before the criminal trial

119. Instruments described as "predictive policing" (before the judicial process or before a court referral) are already growing rapidly and are beginning to be known by the general public (for example, think of the no fly list, which is actually a big data analytics application that collects and analyses data on potential terrorists in order to prevent the commission of acts, or algorithms used to detect fraud or money laundering).

120. In general, a large number of computer tools are commonly used to prevent the commission of criminal acts (by identifying possible places where this could happen or their authors) or prosecute them more effectively.[66] The

---

64. See European Court of Human Rights, Grand Chamber, *Vinter and Others vs. United Kingdom*, paras. 114-118

65. On the other hand, the use of AI for treatment and rehabilitation purposes (e.g. to collect data on treatment administered or methods of reintegration in prison) should be encouraged.

66. See Ales Zavrsnik, *Big Data, crime and social control*, page 194 et seq., which lists in detail a series of instruments used by police services in Europe and the United States.

first category includes "predictive policing" instruments that are used to prevent certain types of offences with elements of regularity in their occurrence such as burglary, street violence, theft from/of vehicles. The designation of these tools derives from their ability to determine precisely where and when these offences could be committed and to reproduce this information on a geographical map in the form of hot spots that are monitored in real time by police patrols. This process is called predictive criminal mapping. Most of the software used in this area is based on historical crime location evidence, such as police reports, but even more powerful new technologies combining various data and from different sources are also being tested.[67] These instruments, which have very persuasive rates of effectiveness, are also claimed to have deterrent effects on the commission of offences in areas surrounding hot spots, leading to a positive opinion of public policies.[68]

121. However, the predictive capabilities of these tools, which show their limitations with regard to crimes of a less regular nature or targeting different locations, such as terrorism, must be put into perspective. In addition, one of their weaknesses is the effect of "vicious circles" and "self-fulfilling prophecies": neighbourhoods considered at risk attract more police attention and police detect more crime, which leads to excessive police surveillance of the communities living in them.[69] Lastly, questions about a possible "tyranny of the algorithm" that could minimise or even progressively replace human judgment are not totally absent within the police services themselves, even

----

67. For example, as part of the project "E-Security – ICT for knowledge-based and predictive urban security" (http://www.esecurity.trento.it/), which was conducted in the Italian city of Trento between November 2012 and May 2015, a database gathering information on crimes reported to the police, the results of surveys conducted by the city hall on victimisation and real and perceived security by citizens, information on physical and social urban disorder from the police, as well as other variables relating to "SmartCity" (e.g. information relating to the socio-demographic context, urban setting, night lighting, the presence of surveillance cameras and public transport). It was created to better equip crime prevention and urban safety improvement work. The project managers testified to the reliability of the techniques used, which were said to make it possible to predict criminal acts with a success rate of approximately 60-65 % and which were said to help to better fight crime when limited resources were available. In addition, tests conducted in the United Kingdom as part of a pilot project to predict possible burglary, theft, and assault locations show that the software projections used, called PREDPOL, were accurate in 78 % of cases, compared to 51 % using traditional techniques.
68. Indication of a geographical concentration of the crime would help police forces to better consider the environmental factors making the crime more likely in the identified area (lighting, presence of shops, etc.) and to plan adequate responses in consultation with other partners.
69. "Predicting crime, LAPD style", The Guardian, 25 June 2014.

if, for the time being, technology is presented as remaining at the service of human beings in order to better equip them for decision-making.[70]

122. In addition, big data analytics are increasingly being applied in the prosecution of crime. Tools such as Connect, which is used by the UK police to analyse billions of data generated in financial transactions to find correlations or patterns of operations, or the International Child Sexual Exploitation Database (ICSE DB), managed by Interpol, which helps identify victims and/or perpetrators through the analysis, for example, of furniture and other objects in abusive images, or the analysis of background noise in videos, have proven particularly effective in fighting crime. With Connect, for example, searches that previously required months of investigation can now be performed in minutes, with a very high level of complexity and volume of data.

123. The doctrine nevertheless questions the managerial logic of the crime response provided by these predictive tools, in which a thorough analysis of the reasons for the crime becomes less important than doing something here and now. This is occurring at a time when available budgets are shrinking and the police must provide the same level of public protection, but with limited personnel, equipment and resources.[71]

7.2. Tools during the criminal trial

124. The use of predictive tools[72] by judges in criminal trials is very rare in Europe.

125. HART (Harm Assessment Risk Tool) was developed in partnership with Cambridge University and is now being tested in the UK. This technology based on machine learning was trained using Durham Police archives dating from 2008 to 2012. By learning from decisions made by police officers during this period, and whether or not certain suspects reoffended, the machine is expected to be able to assess the risk – low, medium or high – of suspects reoffending, based on around thirty factors, some of which are not related to the crime committed (for example, postal code and gender).

------

70. "How technology is allowing police to predict where and when crime will happen", The Independent, 7 October 2017.
71. Ales Zavrsnik, *Big Data, crime and social control*, page 196.
72. In the literature, these tools are often referred to as "algorithmic justice" or "automated justice", or "simulated justice".

126. In tests initially conducted in 2013, during which suspect behaviour was observed over a two-year period after commission of the crime, HART predictions were found to be 98 % effective at predicting low risk and 88 % effective at high risk of recidivism. In this experimental phase, HART will have a purely advisory value for the judge. In addition, audits of HART's functioning and the reliability of its conclusions will be regularly conducted by the police.

127. Even if it is the only predictive tool identified in Europe to date, it provides the opportunity to consider the challenges that public decision-makers could face in the near future if this type of application is tested on a larger scale, particularly in the light of findings in America.

128. In the United States,[73] the NGO ProPublica revealed the discriminatory effects of the algorithm used in COMPAS software (Correctional Offender Management Profiling for Alternative Sanctions), which aims to evaluate the risk of recidivism when the judge must determine the sentence for an individual.

129. This algorithm, which was developed by a private company and which must be used by judges in certain American federal states, includes 137 questions answered by either the defendant or information pulled from criminal records. The questions are quite varied and include the presence of a telephone at home, difficulty paying bills, family history, criminal history of the accused, etc.[74] The algorithm rates the person on a scale from 1 (low risk) to 10 (high risk). It is an aid to judicial decision-making, its conclusions being only one of the variables considered by the judge when deciding on the sentence.

130. African-American populations were assigned a recidivism high-risk rate twice that of other populations within two years of sentencing – without this effect being naturally sought by the designers[75]. In contrast, the algorithm considered that other populations appeared much less likely to repeat an

---

73. A 2015 study had identified about sixty predictive tools in the United States.
74. There are other algorithms that have been developed using critical observations expressed by the doctrine (see next chapter) that are based on smaller variables, more directly related to the crime committed and less related to race, gender or socio-economic status. An example is the Public Safety Assessment Tool used in 30 American jurisdictions.
75. This purely discriminatory effect can in fact be explained by the relatively permissive "calibration" of the algorithm model, which creates many "false positives".

offence[76]. It must of course be noted that this type of misleading interpretation in reality reveals only the social and economic fragility of certain groups of populations which are obviously not criminogenic by nature. Researchers at Dartmouth College have also shown that this type of algorithm does not produce added value since people without any criminal history can reproduce exactly the same assessment simply by answering the questionnaire.

131. In addition, the lack of transparency in the algorithm operation processes designed by private companies (which claim intellectual property) was another cause for concern. If we take into account the fact that they take their source data from the state authorities themselves, their lack of accountability to citizens poses a major democratic problem. Accounts have shown that the public is informed about big data operations accidentally, sporadically and when there are leaks or errors: an example of this is when ProPublica revealed the flaws in the COMPAS algorithm following the owner company's refusal to share it. The NGO had to appeal to public authorities to access the data and hire its own scientist to examine the algorithm.

7.3. The challenges of "prediction" in criminal matters

132. We have seen in the previous sections that the degree of development of predictive tools in Europe in the criminal field is very varied. Although instruments described as "predictive policing" are growing rapidly and are even beginning to come to the attention of the general public, the situation is not the same when it comes to the application of this type of tool by

---

76. Black populations were more frequently classified as high risk, while they did not reoffend within two years of conviction; white populations, on the other hand, who were more frequently classified as low risk, committed offences within the next two years. In short, the algorithm overestimated the risk of recidivism for blacks and underestimated it for whites (the "false positives" were mostly blacks, while the "false negatives" were mostly whites). In response to ProPublica's allegations, NorthPointe (now Equivant since this controversy) replied that white and black populations were equally represented when considering the "true positives", i.e. those who had actually reoffended. **The question of how to reconcile both the accuracy of the algorithm in detecting recidivism and the need to avoid discriminatory effects towards black populations has been the source of intense debate in the literature;** see in particular Chouldechova A (2016), "A fair prediction with a disparate impact: a study on bias in recidivism prediction instruments", available at http://arxiv.org/abs/1610.07524; and also "Bias in criminal risks scores is mathematically inevitable, Researchers say", available at https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say. **This debate also reflects the question of the legitimacy of a private company, without any institutional control, to arbitrate between two opposing requirements: that of defending society, on the one hand, and that of respecting the rights of individuals, on the other.**

judges in criminal trials. As regards the instruments available to prosecution services, thoughts about their advantages and disadvantages have already been expressed. Let us now study the tools specific to criminal trials.

133. First of all, it is important to rule out arguments based solely on the efficiency or inefficiency of these tools. The examples given above show that there can be tremendous opportunities but also real risks in the application of new technologies that are used without the necessary precautions. Public decision-makers and judicial stakeholders must be particularly vigilant and play an active role in the development of these technologies; continuous monitoring is necessary to determine their real effectiveness and efficiency, and to avoid unforeseen consequences. This is even more important in criminal proceedings because of their direct impact on the individuals' personal freedoms.[77]

134. This implies that both the benefits and the drawbacks of the application of such tools in the judicial field should be carefully measured.

135. Supporters often argue that they are neutral and that they rely on factual and objective methods that help to make justice more accurate and transparent. Another great asset, it is claimed, is their efficiency, which sometimes exceeds human capacities and can only be extremely valuable in a general context of reduced public funds or even a shortage of resources.

136. The inclusion of algorithmic variables such as criminal history and family background means that the past behaviour of a certain group may decide the fate of an individual, who is, of course, a unique human being with a specific social background, education, skills, degree of guilt and distinctive motivations for committing a crime[78]. They also argue that human decisions can be based on values and considerations (e.g. societal) that would not be retained by the machine. For example, a judge could decide to order the bail of a female offender who has a risk of recidivism, on the basis of a hierarchy of values, for example by setting greater store by her role as a mother and protector of her children, whereas the algorithm would be able to determine the risk of reoffending more accurately but would not be able to operate such a hierarchy of priorities.

----

77. An extract from the Wisconsin Supreme Court decision in *Wisconsin v. Loomis* can also provide inspiration at European level: "It is important to consider that tools such as COMPAS continue to change and evolve. The concerns we address today may very well be alleviated in the future. It is incumbent upon the criminal justice system to recognize that in the coming months and years, additional research data will become available. Different and better tools may be developed. As data changes, our use of evidence-based tools will have to change as well. The justice system must keep up with the research and continually assess the use of these tools."

78. Aleš Završnik, "Big Data, crime and social control", page 196.

137. In criminal matters, there are also potential risks of discrimination when one considers that these tools, which are constructed and interpreted by humans, can reproduce unjustified and already existing inequalities in the criminal justice system concerned; instead of correcting certain problematic policies, technology may end up legitimising them. As already mentioned, the NGO ProPublica[79] clearly revealed the discriminatory effects of the algorithm used in COMPAS,[80] which predicted that black populations were twice as likely to reoffend as white populations within two years of sentencing while considering that white populations were much less likely to repeat the offence. They could, however, help reveal the errors in decision-making so that they may be corrected.[81] Moreover, the lack of transparency in algorithm construction processes by proprietary companies and their accountability to the public is a cause for concern, all the more so if they are part of steps taken by state authorities to make data available to the public.

138. In the light of the foregoing, when algorithms are used in the context of a criminal trial it seems essential to fully guarantee respect for the principle of equality of arms and presumption of innocence established by Article 6 of the ECHR. The party concerned should have access to and be able to challenge the scientific validity of an algorithm, the weighting given to its various elements and any erroneous conclusions it comes to whenever a judge suggests that he/she might use it before making his/her decision. Moreover, this right of access is also covered by the fundamental principle of personal data protection. All people have the right not to be subject to decisions affecting them significantly made solely on the basis of automated data processing, without their point of view having been taken into account beforehand.

139. In this respect, there is a difference between Europe and the United States with regard to the right of access to algorithms: while in the United States the judicial authorities are still reluctant to recognise this right fully and weigh private interests (particularly the protection of intellectual property) against the rights of defence, in Europe the framework is more protective because of the GDPR, which establishes a right to information on the underlying logic of decisions made using algorithms.[82]

----

79. www.propublica.org/article/technical-response-to-northpointe.
80. Other algorithms focus on other elements more directly related to the offence committed.
81. Mojca M. Plesnicar and Katja Sugman Stubbs, "Subjectivity, algorithms and the courtroom".
82. Article 15, 1. (h) of EU Regulation 2016/679: "The data subject shall have the right to obtain from the controller"... "the following information:"..." the existence of automated decision-making, including profiling, as referred to in Article 22, paragraphs 1 and 4, and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject".

140. The considerations expressed earlier regarding the potentially negative effects of these tools on the impartiality of the judge are also valid in criminal matters: a judge who decides against the prediction of an algorithm is likely to take risks as he assumes greater responsibility. It does not seem unrealistic to imagine that judges would be reluctant to take on this additional burden, particularly in systems where their terms of office are not permanent but subject to popular vote,[83] or in which their personal liability (disciplinary, civil or even criminal) is likely to be incurred, especially if their statutory guarantees in disciplinary matters are insufficient.

8. Specific questions relating to the protection of personal data

The use of algorithms raises the question of the protection of personal data when being processed. The precautionary principle should be applied to risk assessment policies.

141. For the full potential of algorithms to be exploited while complying with data protection principles, the precautionary principle should be applied and preventive policies should be put in place to counter the potential risks associated with the use of the data processed by these algorithms and the impact of their use on individuals and society in general.

142. The principle of lawfulness in the processing of personal data and the obligation to prevent or minimise the impact of data processing on the rights and fundamental freedoms of data subjects should induce prior risk assessment. This should make it possible to implement the appropriate measures, particularly during the design stage (and hence by design) and by default, in order to mitigate the risks identified.

143. Since personal data must be processed for specified and legitimate purposes, they must not be used in a way that is incompatible with those purposes and must not be further processed in a way that the data subject may consider unexpected, inappropriate or questionable (principle of loyalty). The issue of re-using personal data, making them widely accessible, therefore needs to be handled with the utmost caution.

144. The design of the data processing methods used by algorithms should minimise the presence of redundant or marginal data and avoid any potential

------

83. Mojca M. Plesnicar and Katja Sugman Stubbs, "Subjectivity, algorithms and the courtroom".

hidden bias and any risk of discrimination or negative impact on the fundamental rights and freedoms of the data subjects.

145. When artificial intelligence is used, the rights of data subjects are of especial importance, and the control that each of us must have over our personal information implies that it must be possible to exercise the following rights: the right of data subjects not to be subject to automated decisions significantly affecting them without their point of view being taken into account, the right to obtain information about the reasoning underlying the data processing carried out by algorithms, the right to oppose such processing and the right to a legal remedy.

9. The potential and limitations of predictive justice tools

The term predictive justice should be dismissed because it is ambiguous and misleading. These tools are based on methods of analysis of case-law, using statistical methods that do not in any way reproduce legal reasoning but may attempt to describe it. Analytical biases, if they cannot be totally eliminated, must be identified. The design process and the use of the tool must be embedded in a clear ethical framework.

146. In section 3, we already highlighted the ambiguity and fallacy of the concept of predictive justice and how it operates a slow shift in the collective mind, leading us to believe that machines, devoid of any emotion, will one day be better able to make the act of judging more reliable. More than ever, its promises need to be examined in an objective and scientific manner, based on solid foundations of fundamental research, in order to identify possible limitations. In this connection, it should be noted that the risks of distorted interpretations of the meaning of court decisions are extremely high when based on statistical modelling alone. This observation is borne out further by the lack of precise understanding of the links between the data and the obvious presence of false correlations that cannot be discerned in large masses of data.

147. Moreover, the neutrality of algorithms is a myth, as their creators consciously or unintentionally transfer their own value systems into them. The philosopher, Eric Sadin, noted that, behind their efficient and impersonal facade, algorithmic systems imperceptibly reflect the intentions of their designers or sponsors, inducing a functioning and asymmetric power over the lives of other people. Similarly, researcher Aurélien Grosdidier considers that an algorithm, in itself, is capable of nothing other than allowing us – at best – to grasp part of the designer's intention and extends the questioning

to the entire information processing chain (designer's intention, production of computer code, execution of computer code and context of execution then maintenance). This observation is also shared by the criminologist Aleš Zavrsnik, who underlines how the stages of construction and interpretation of algorithms are carried out by man, for man, and cannot escape errors, prejudices, values, human interests and a human representation of the world, however they are conceived.

148. Despite these significant limitations, should we overlook the contribution of a technology with unequalled power? Mathematicians C. S. Calude and G. Longo themselves stress in their study on big data that the restrictive or negative scope of their results, as often happens, does not destroy the science of data but paves the way for greater thought, including the challenge of a new, more extensive scientific method capable of incorporating both new algorithmic instruments and classical tools by accompanying processing with a rigorous evaluation of evidence. As highlighted in the introduction, the use of AI is likely to offer extremely significant support for professionals, including judges and lawyers, but also the general public, especially if one day they make it possible to construct unparalleled research and documentary analytical tools in legislative, regulatory, jurisprudential and doctrinal matters and create dynamic links between all these sources. But this type of application goes beyond the scope of this article since it is not designed to predict the outcome of a dispute but to analyse case-law in a given time and space.

149. As discussed in section 6, subject to the representativeness of selected and processed samples, AI has helped to devise much more precise scales of the average or median amounts of money allocated, *mutatis mutandis*, in various fields (financial support, compensatory benefits, compensation for bodily injury, severance pay, etc.). These scales, which are based more on a consensus than on an average analysis of what already exists, already provide significant support for decision-making and guidance without being able to replace the law itself. As mentioned earlier, the risk is that, in the absence of a statistical representation of reality or of being able to predict anything, the results of predictive justice software will be set as standards without any validation by the legal system and in conflict with it.

150. Lastly, let us consider the idea of being able to step back at will from predictive systems. Rather than locking users into a probability (or set of probabilities), the idea would be to allow them to navigate through the correlations that led the system to propose its assessment and to be able to distance themselves by selecting other more relevant concepts or groups of words or to exclude false correlations. To use the UCL example, this would

consist of proposing a graphic representation of the different terms retained by the system (with their respective weightings) to find that there has been a violation (or non-violation) and to authorise other paths to be taken by proposing the selection of other terms or lexical groups.

151. However bold and seductive this proposal may be, it presupposes that the professionals themselves (judges, lawyers, universities) take it over collectively to test its feasibility and that they do not allow private operators alone, save for a few unchecked scientists, to design software and abstruse or locked modes of reasoning or calculation.

152. The ambitious (and unfulfilled) promises of some legal tech companies must not hide the immense potential of technologies and the need for applications adapted and built in directly with scientific and academic research environments, as well as with all legal professionals, such as magistrates, clerks, lawyers, notaries, bailiffs and field experts. A number of measures would appear to be able to take full advantage of these new tools through applications adapted and devised in direct association with research workers and all legal professionals, including judges, prosecutors, clerks, lawyers, notaries, bailiffs and experts in the field.

153. In this dynamic context, it seems essential, firstly, not to make hasty decisions and to take time to debate in advance the risks and practical applications of these instruments for judicial systems, and to test them in the first stage. A judicial system in keeping with its time would be one that is capable of establishing, administering and guaranteeing genuine **cyberethics** for both the public and private sectors, and insisting on total transparency and fairness in the functioning of algorithms, which may contribute one day to judicial decision-making.

10. The need for an in-depth public debate on these tools prior to the implementation of public policies for their development. The urgent need for cyberethics to provide a framework for the development of artificial intelligence algorithms while respecting fundamental rights

The challenge of integrating these tools into a judicial decision-making process justifies simplifying the concepts for the target group concerned. An ethical framework must be set up to promote the rapid development of a form of AI that includes mechanisms preventing bias and discrimination in its very design processes.

10.1.The importance of debating, testing and continually reviewing the application of these tools prior to the implementation of public policies

154. The issues connected with the implementation of predictive justice tools are so numerous and multifaceted that they call for a balanced approach by public decision-makers.

155. Firstly, it is essential to hold a public debate around these questions, bringing together both the designers of the tools and legal professionals. Judicial councils, professional associations of judges and bar associations can undoubtedly contribute to this and help to identify opportunities and more controversial aspects. In addition, judicial training and law schools can play a key role in raising awareness among justice professionals on these issues, so that they can better understand and practically contribute to current developments.

156. It is also essential to carry out research on the proposed applications and test them, both to understand their potential and their weaknesses and to be able to develop them further and adapt them to our needs. A right to examine the components and characteristics of the instruments proposed by the private sector (or those developed by independent and specialised public institutes, a solution which should be encouraged) seems equally important so that the justice service can effectively carry out its mission. A rigorous evaluation of test results should be conducted before wider deployment and integration into public policy. It also seems strongly advisable to regularly assess the impact of these tools on the work of justice professionals.

10.2. The establishment of an ethical framework

157. First of all, merely adopting a legislative or regulatory framework for AI seems vain in a digital context, which is inherently transnational in its scope. On the other hand, scrupulous attention to the nature and quality of open data is likely to minimise the risks of inadequate cross-referencing and to reinforce the relevance of the results of automated processing. With regard to the names of professionals, a simple precaution would be to prohibit their public dissemination in structured raw databases given the risks of misuse. It would not be a question of limiting access to the information already processed (for example, the composition of a panel of judges) but of filtering the raw data made freely available. In short, a distinction needs to be made between access to information and access to databases, which can be manipulated at will.

158. Researchers Buttarelli and Marr stressed how big data need to be closely controlled and protected. Other researchers (Pasquale and Morozov) have stressed the need to establish transparent procedures for the deployment of big data and more generally AI in the judicial field, because the solutions proposed can never represent life in its complexity.

159. The development of cyberethics rules to guide the activity of stakeholders in the sector and to promote the above-mentioned principles of transparency, fairness and neutrality of the tool is essential. Regular monitoring by independent experts should ensure that the drivers of artificial intelligence used to assist judges in their decisions are not biased. It is not inappropriate to anticipate the implementation, discreet or not, of paid referencing systems (based on Google's Search Engine Advertising model) allowing certain operators to give less weight to decisions that are unfavourable to them. These rules will play a key role in increasing citizens' confidence in their judicial systems.

160. In this connection, the quality of the best systems could be recognised by the award of a label or certification. In particular, they must guarantee total transparency and perfect fairness in the way information is processed, both for professionals and for citizens, to prevent the repetition of errors such as the aforementioned COMPAS algorithm. Justice professionals must be closely involved to be able to properly assess the risks and the impact of these applications on judicial systems.

161. Nowadays, all the experts involved in the development of AI, including researchers, engineers and computer developers, have quite exceptional and unprecedented responsibilities. Their work could be accompanied by an even greater strengthening of the humanities. The example of some innovative schools of computer developers shows that behind the will to "hack the system" hides, according to some observers, in reality a pragmatism devoid of any contextualisation of the responsibility that now hangs over technicians with quasi-demiurge powers. The Hippocratic Oath certainly has its limits in medicine but ritualises responsibility and provides an ethical framework.

162. Finally, cyberethics must be accompanied by large-scale training of stakeholders, from algorithm designers and legal tech companies to their users. New transdisciplinary humanities should be made available to all so that AI becomes a vector of positive development for humankind.

Appendix II
Which uses of AI in European judicial systems?

This Annex to the Charter reviews different uses of AI in European systems and encourages to a different degree their application in the light of the principles and values set out in the Ethical Charter.

The use of machine learning to constitute search engines for case-law enhancement is an opportunity to be taken up for all legal professionals. Additional applications (drawing up of scales, support for alternative dispute settlement measures, etc.) should be considered, but due care must be taken (in particular, the quality of the data source and not mass processing of the entire dispute in question). Other applications ("predictive justice") should be assigned to the field of research and further development (in consultation with legal professionals in order to ensure that they fully tie in with actual needs) before contemplating use on a significant scale in the public sphere.

In criminal matters, this is a very sensitive issue but it should not be ignored. In the light of the many existing questions as to their compatibility with a certain number of fundamental rights, the use of algorithms to calculate the potential risks of recidivism of an individual brought to justice should be considered with the most extreme reservations. On the other hand, the processing of global quantitative data for crime prevention is an avenue to be further explored with these new techniques, taking into consideration known biases (performative effects, data quality, etc.). Similarly, the use of algorithms to form a better link between the type of community service available and an individual's personality may be a factor in the effectiveness of a measure of this kind.

Uses to be encouraged

▶ **Case-law enhancement**: machine Learning techniques have been increasingly deployed in the field of natural language processing in the past years (this includes initial efforts in natural language understanding) and are a considerable asset for finding search options to complement current keyword or full-text search. These tools could link various sources (e.g. constitutions and conventions, laws, case law and legal theory). Data visualisation techniques could illustrate search results.

▶ **Access to law**: without replacing human intervention, chatbots could be set up to facilitate access to the various existing sources of information using natural language. Document templates (court applications, lease agreements, etc.) could also be generated online.

▶ **Creation of new strategic tools**: the use of data science and artificial intelligence techniques on court activity data can help improve the efficiency of justice by making it possible, for example, to carry out quantitative and qualitative evaluations and to make projections (e.g. future human and budgetary resources). Key performance indicators could be drawn up on this basis. It is recommended that legal professionals, especially judges, be involved in the implementation of these tools, in terms of taking ownership of these tools and of analysing the results in conjunction with factors relating to the specific features of the court in question or the quality of justice (for example, the need to preserve access to justice).

Possible uses, requiring considerable methodological precautions

▶ **Help in the drawing up of scales in certain civil disputes**: an analysis of all judicial decisions is not statistically meaningful if all the causative factors (explicit and implicit in the decisions) are not identified. Knowing that the average compensation awarded in a certain geographical area is higher than in another can be explained not due to the behaviour of judges, but in the light of the characteristics of the area in question. Machine learning can therefore be useful in identifying decisions (see *Case-law enhancement* above), but the automated processing of data alone cannot produce meaningful information. An essential prerequisite is the compilation of a relevant sample of decisions to be processed (for example, via surveys).

▶ **Support for alternative dispute settlement measures in civil matters:** in some European countries, "predictive justice" tools are used by insurance companies to evaluate the chances of success of a dispute and to steer the litigant towards another method of dispute resolution when it is felt that there is little chance of success. Furthermore, some systems abroad offer compensation amounts without any real transparency as to the rules of calculation. However, these systems cannot be considered as impartial and reliable (see the section on machine learning techniques). Decisions are made about a citizen using truncated bases. In other cases, a litigant may be advised, by means of a virtual agent (chatbot), to opt for an alternative dispute settlement measure after a preliminary examination of the criteria entered by the litigant himself or herself, when visiting a court's website or searching legal information online. The virtual agent may, where appropriate, also recommend that the litigant seek advice from a mediation service or a lawyer. In all these cases, the presence of a trained third party (mediator using not only techniques but maybe scales as calculated above, or a lawyer) would appear to be the most appropriate solution at this stage.

▶ **Online dispute resolution**: when litigants go onto an online dispute resolution platform, they should be informed in a clear and comprehensible manner whether the processing of their dispute is done in an entirely automated way or with the involvement of a mediator or arbitrator. In addition, the information given to litigants must be honest and must avoid giving them the impression that a court is involved (in this connection, the term "online court" is often used for this type of platform, whereas technically their purpose is to provide alternative dispute resolution services). These are two essential factors in enabling litigants to make an informed choice, possibly disagreeing with the advice and deciding to go to a real court within the meaning of Article 6 of the ECHR. Furthermore, in view of the requirements of Articles 6 and 13 of the ECHR, forms of review of the online dispute resolution procedure and its outcome by state courts should always be considered, especially where the litigant has consented to fully automated online dispute resolution.

▶ **The use of algorithms in criminal investigation in order to identify where criminal offences are being committed**: this type of application could concern not only the police but also prosecutors in the crime prevention bodies of which they are a part. Systems have been used in the United States to guide police patrols in real time towards possible locations where offences are being committed. However, this type of

quantitative approach can generate a strong "performative effect" (in a given location, there is a greater chance of discovering an offence and this then reinforces the system). Criminal analysis through approaches combining geographic information systems (GISs) and large amounts of data on procedures could be better shared with prosecutors and could certainly benefit from a significant machine learning contribution. Anti-money laundering units already use "predictive" systems to identify suspicious financial flows, but in the case of quantitative (financial) information, machines are more able to produce reliable results. Researchers should also have better access to these data to produce relevant studies for policy-makers.

Uses to be considered following additional scientific studies

► **Judge profiling**: quantifying a judge's activity will reveal less about any possible biases than about any external factors influencing his or her decisions. The judge himself or herself is not the reason why judicial activity in an impoverished area does not produce the same results as in another territory, whatever his personality may be. When the decision is rendered in a collegial manner and without the possibility for a judge to express a divergent opinion, it is pointless to profile each of the judges of the chamber. On the other hand, offering judges a more detailed quantitative and qualitative assessment of their activities, thanks to new tools, but with a purely informative aim of assisting in decision-making and for their exclusive use, could be encouraged.

► **Anticipating court decisions:** statistical processing of lexical groups alone reveals the frequency of the use of certain groups of words but does not identify the real reasons for a decision and does not carry out a legal analysis (see the study carried out on ECHR decisions by the University College of London which produced better results on the facts than on the analysis of the law). Hybrid systems, founded on the construction of mathematical models that are supposed to represent the diverse range of judges' reasoning, are not any more efficient because they are still limited by bias in the data sample that they have processed and have to start again from square one if a law is amended or if there is a reversal in case law.

Uses to be considered with the most extreme reservations

► **Use of algorithms in criminal matters in order to profile individuals**: experiments in other countries (COMPAS in the United States

and HART in the United Kingdom) have been criticised by NGOs (see work by ProPublica in the United States and Big Brother Watch in the UK). Because of the limitations of the methodology used, this pure statistical approach has led to wrong result: the finding that some African-American individuals are more often involved in criminal acts has led to a higher risk factor for the entire African-American population. Thus, even for minor offences, these systems have negatively weighted African-American defendants, with the result of unfairly increasing the quantum of their sentences. This approach, which has discriminatory and deterministic effects, must be replaced by one that is more respectful of European standards concerning criminal sanctions and which must offer the individual the possibility of rehabilitation and reintegration. If algorithmic systems manage to help improve the collation of information for probation services, for example, and make it possible for the relevant information to be collected more quickly for subsequent human processing, then progress would definitely be made (particularly in expedited proceedings). Any other use is prone to biases which will come into conflict with certain national and supra-national fundamental principles.

► **Quantity-based norm:** it is not only a question of producing scales, which could be legitimate, but of providing each judge with the content of the decisions produced by all the other judges and claiming to lock his future choice into the mass of these "precedents". This approach should be rejected because this large number cannot add to or act in place of the law. For the reasons provided above (*Help in the drawing up of scales*), a quantity-based approach is not the way to go. The CEPEJ study also highlighted the dangers of the crystallisation of case law and the potentially negative effects on the impartiality and independence of judges.

Appendix III
Glossary

This glossary provides a definition of the terms used by the Ethical Charter and study document. Preference has been given to providing a narrow definition for all the vocabulary used. All the documents must be read and understood in the light of these definitions.

A

ALGORITHM Finite sequence of formal rules (logical operations and instructions) making it possible to obtain a result from the initial input of information. This sequence may be part of an automated execution process and draw on models designed through machine learning.

ANONYMISATION Method to process personal data in order to completely and irreversibly prevent the identification of a natural or legal person. Anonymisation therefore implies that there is no longer any possible link between the information concerned and the person to whom it relates. Identification then becomes completely impossible.[84] As the principles relating to data protection apply to all information relating to an identified or identifiable individual, they do not apply to anonymised data.

ARTIFICIAL INTELLIGENCE (AI) A set of scientific methods, theories and techniques whose aim is to reproduce, by a machine, the cognitive abilities of human beings. Current developments seek to have machines perform complex tasks previously carried out by humans.

---

84. Working Party on Article 29 Opinion 05/2014 on Anonymisation Techniques. See also No. 26 of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016.

However, the term artificial intelligence is criticised by experts who distinguish between "strong" AIs (yet able to contextualise specialised and varied problems in a completely autonomous manner) and "weak" or "moderate" AIs (high performance in their field of training). Some experts argue that "strong" AIs would require significant advances in basic research, and not just simple improvements in the performance of existing systems, to be able to model the world as a whole.

The tools identified in this document are developed using machine learning methods, i. e. "weak" AIs.

B

BIG DATA (*metadata, large data sets*) The term big data refers to large sets of data from mixed sources (e.g. open data, proprietary data and commercially purchased data). For data derived from judicial activity, big data could be the combination of statistical data, records of business software connections (application logs), court decisions' databases, etc.

C

CHATBOT (*conversational agent*) Conversational agent which converses with its user (for example, empathy robots used to help those who are ill, or automated conversation services in customer relations).[85]

D

DATA Representation of information for automatic processing. When it is said that algorithms can be "applied" to the most diverse realities in the legal world or elsewhere, one presupposes the "digitalizability" of any reality in the form of "data". But it is clear from physics that nothing tells us that physical processes can be adequately translated in terms of "data" (and integrated into the input/output cycle of algorithms). If this is already the case in physics, there is no reason why this should not also be the case in social relations. We must therefore be cautious with the idea of "data", which always assumes that the reality we are trying to describe has a format such that it is naturally algorithmically processable.

---

85. CNIL Report December 2017: How can humans keep the upper hand? The ethical matters raised by algorithms and artificial intelligence.

DATABASE A database is a "container" that stores data such as numbers, dates or words, which can be reprocessed using a computer to produce information, for example, gathering and sorting numbers and names to form a directory.

DATA MINING Data mining makes it possible to analyse a large volume of data and highlight models, correlations and trends.

DATA SCIENCE A large field that groups together mathematics, statistics, probabilities, data processing and data visualisation in order to gain understanding from a mixed set of data (images, sound, text, genome data, links between social networks, physical measurements, etc.).

Methods and tools derived from artificial intelligence fall within this category.

DEEP LEARNING See *Machine Learning* and *Neurons*

E

EXPERT SYSTEM This is one of the ways to achieve artificial intelligence. An expert system is a tool capable of reproducing the cognitive mechanisms of an expert in a particular field. More precisely, it is software capable of answering questions, by reasoning based on known facts and rules. It consists of 3 parts:

> ► a fact base;
> ► a rule base;
> ► an inference engine.

The inference engine is able to use facts and rules to produce new facts, until it reaches the answer to the expert question asked.

Most existing expert systems are based on formal logic mechanisms (Aristotelian logic) and use deductive reasoning.

L

LEGAL TECH Companies using information technology in the field of law in order to offer innovative legal services. These companies are start-ups specialised in law. Other terms derived from business sectors have also appeared such as *Fintechs* for start-ups offering financial services and *Medtechs* in the medical field.

M

MACHINE LEARNING Machine learning makes it possible to construct a mathematical model from data, incorporating a large number of variables that are not known in advance. The parameters are configured gradually during the learning phase, which uses training data sets to find and classify links. The different methods of machine learning are chosen by the designers depending on the nature of the tasks to be completed (grouping). These methods are usually classified into three categories: (human) supervised learning, unsupervised learning and reinforcement learning. These three categories group together different methods including neural networks, deep learning, etc.

The graph below illustrates the different categories of machine learning:
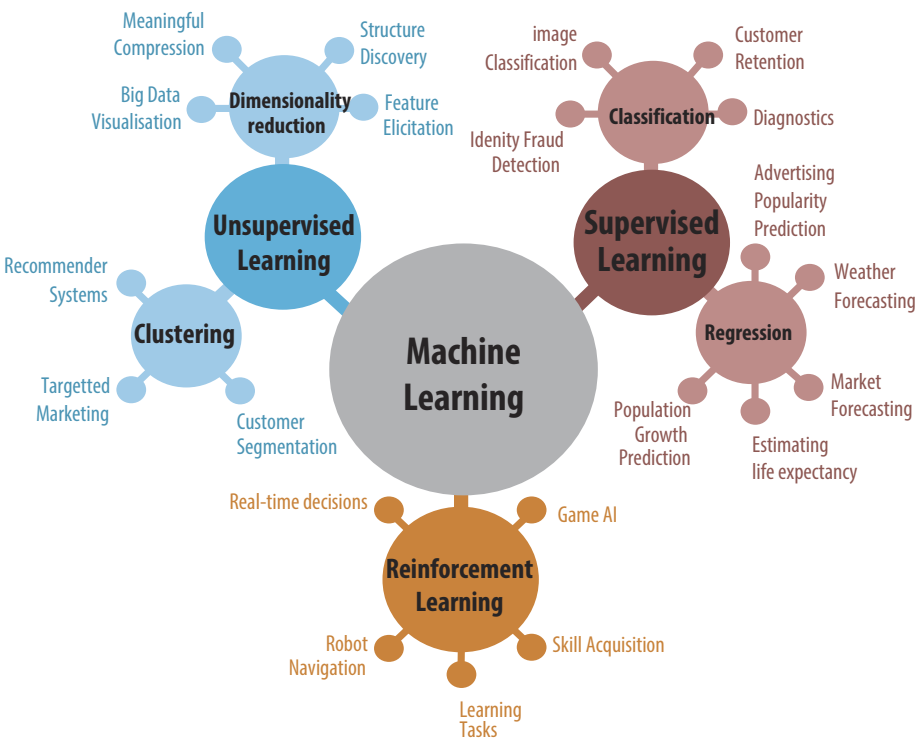
METADATA Data that make it possible to define, contextualise or describe other data. In most of its computer uses, the meta prefix means "reference definition or description".

Metadata synthesize basic information about data, they facilitate the search and manipulation of particular data instances. The author, creation date, modification date and file size are examples. Metadata and its corollary, data filtering, help to locate a specific document.

N

NEURONS/NEURAL NETWORK Neural networks are computing systems vaguely inspired by the biological neural networks that constitute animal brains.[1] Such systems "learn" to perform tasks by considering examples, generally without being programmed with any task-specific rules. For example, in image recognition, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any prior knowledge about cats, e.g., that they have fur, tails, whiskers and cat-like faces. Instead, they automatically generate identifying characteristics from the learning material that they process.

An ANN is based on a collection of connected units or nodes called artificial neurons which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal additional artificial neurons connected to it.

The original goal of the ANN approach was to solve problems in the same way that a human brain would. However, over time, attention moved to performing specific tasks, leading to deviations from biology. Artificial neural networks have been used on a variety of tasks, including computer vision, speech recognition, machine translation, social network filtering, playing board and video games and medical diagnosis.

O

OPEN DATA The term refers to making structured databases available for public download. These data can be inexpensively re-used subject to the terms of a specific licence, which can, in particular, stipulate or prohibit certain purposes of re-use.

Open data should not be confused with unitary public information available on websites, where the entire database cannot be downloaded (for example, a database of court decisions). Open data do not replace the mandatory publication of specific administrative or judicial decisions or measures already laid down by certain laws or regulations.

Lastly, there is sometimes confusion between data (strictly speaking open data) and their processing methods (machine learning, data science) for different purposes (search engines, assistance in drafting documents, analysis of trends of decisions, predicting court decisions, etc.).

OPEN SOURCE SOFTWARE Software for which the source code is available to everyone. The software can therefore be freely used, modified and redistributed.

P

PERSONAL DATA Any information concerning an identified or identifiable natural person (the "person concerned"), directly or indirectly.

These include sensitive data relating to genetic data, biometric data uniquely identifying an individual, data relating to offences, criminal proceedings and convictions and related security measures, and any data for information they reveal on racial or ethnic origin, political opinions, trade union membership, religious or other beliefs, health or sex life.

PREDICTIVE JUSTICE Predictive justice is the analysis of large amounts of judicial decisions by artificial intelligence technologies in order to make predictions for the outcome of certain types of specialised disputes (for example, redundancy payments or alimentary pensions).

The term "predictive" used by legal tech companies comes from the branches of science (principally statistics) that make it possible to predict future results through inductive analysis. Judicial decisions are processed with a view to detecting correlations between input data (criteria set out in legislation, the facts of the case and the reasoning) and output data (formal judgment such as the compensation amount).

Correlations deemed to be relevant make it possible to create models which, when used with new input data (new facts or precisions described as a parameter, such as the duration of the contractual relationship), produce according to their developers a prediction of the decision (for example, the compensation range).

Some authors have criticised both the form and substance of this approach. They argue that, in general, the mathematical modelling of certain social phenomena is not a task comparable to other more easily quantifiable activities (isolating the really causative factors of a court decision is infinitely more complex than playing the game of Go or recognising an image for example):

here, there is a much higher risk of false correlations. In addition, in legal theory, two contradictory decisions can prove to be valid if the legal reasoning is sound. Consequently, making predictions would be a purely informative exercise without any prescriptive claim.

PROFILING An automated data processing technique that consists of applying a "profile" to a natural person, in particular in order to make decisions about him or her or to analyze or predict personal preferences, behaviours and attitudes.

PROCESSING OF PERSONAL DATA According to Article 2 of the revised Convention 108, "data processing" means any operation or set of operations performed on personal data, such as the collection, storage, preservation, alteration, retrieval, disclosure, making available, erasure, or destruction of, or the carrying out of logical and/or arithmetical operations on such data.

PSEUDONYMISATION Pursuant to Article 4 of the GDPR, this is the processing of personal data in such a manner that they can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.[86]

---

86. Article 4 of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016.

Appendix IV

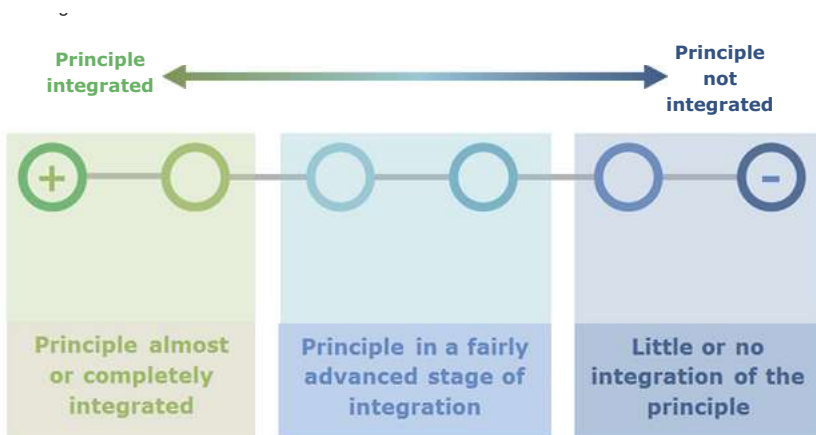# Checklist for integrating the Charter's principles into your processing method

In order to assess the compatibility of your processing method with the Charter, there is a self-evaluation scale available for each of the principles listed.

For each principle, tick the box that corresponds to your processing methods.

The box furthest to the left indicates complete integration, the box furthest to the right indicates no integration.

At the bottom of the evaluation sheet (in the line marked "Total"), add up the number of boxes ticked. The column with the highest score indicates your processing method's level of compatibility with the Charter.

This evaluation is of course purely informative and by no means equates to any certification.

Principle integrated		Principle not integrated
Principle almost or completely integrated	Principle in a fairly advanced stage of integration	Little or no integration of the principle

Checklist for evaluating your processing methods

Principle of fundamental rights:
Ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights, including the right to protection of personal data

Checklist for evaluating your processing methods

Principle of non-discrimination:
Specifically prevent the development or intensification of any discrimination between individuals or groups of individuals

Checklist for evaluating your processing methods

Principle of quality and security:
With regard to the processing of judicial decisions and data, use certified sources and intangible data with models elaborated in a secure technological environment
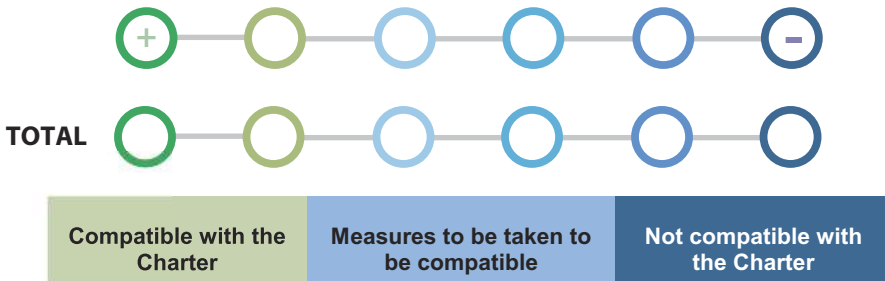
Checklist for evaluating your processing methods

Principle of transparency, impartiality and fairness:
Make data processing methods accessible and understandable, authorise external audits

Checklist for evaluating your processing methods

Principle "under user control": Preclude a prescriptive approach and ensure that users are informed actors and in control of their choices

TOTAL

Compatible with the Charter	Measures to be taken to be compatible	Not compatible with the Charter

The Charter provides a framework of principles that can guide policy makers, legislators and justice professionals when they grapple with the rapid development of Artificial Intelligence in national judicial processes.

The CEPEJ's view is that the application of Artificial Intelligence in the field of justice can contribute to improve the efficiency and quality. It must be implemented in a responsible manner which complies with the fundamental rights guaranteed in particular in the European Convention on Human Rights (ECHR) and the Council of Europe Convention on the Protection of Personal Data. It is essential to ensure that Artificial Intelligence remains a tool in the service of the general interest and that its use respects individual rights.

The Charter defines five core principles to be respected in the field of Artificial Intelligence and justice: respect of fundamental rights; non-discrimination; quality and security; transparency, impartiality and fairness; "under user control". The Charter is accompanied by a n in-depth study on the use of Artificial Intelligence in judicial systems.

www.coe.int/cepej

ENG

www.coe.int

The Council of Europe is the continent's leading human rights organisation. It comprises 47 member states, including all members of the European Union. All Council of Europe member states have signed up to the European Convention on Human Rights, a treaty designed to protect human rights, democracy and the rule of law. The European Court of Human Rights oversees the implementation of the Convention in the member states.

c e p e j

European Commission for the Efficiency of Justice

Commission européenne pour l'efficacité de la justice

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE