



# Legal Challenges of AI Supported Legal Services: Bridging Principles and Markets

Giulia Schneider\*

### Abstract

In light of the persisting regulatory gaps in the field of artificial intelligence-driven legal services, this study questions which are the legal tools that are relevant to govern the current expansion of the correspondent market in a way that is consistent with ethical declarations.

We move from the acknowledgment that machine learning models are being increasingly applied to textual data contained in legal materials for the prediction of outcomes regarding the legal position of citizens, in terms, for example, of discovery review, contract analytics and legal research. In this respect, the analysis gives account of ongoing transformations in the market of Artificial Intelligence (AI)-supported legal services, with the aim of rooting in the market reality the relevant regulatory framework. In our understanding, the analysis related to the risks connected to the employment of AI-driven legal decision-making tools delivered by the market triggers the question whether the applicable ethical-legal framework provides sufficient tools for addressing the current developments in the market of AI-assisted legal services or whether additional sector-specific solutions need to be introduced.

The analysis identifies a gap it intends to fill between the blooming market reality and the ethical and legal perspectives.

The uncertainties stemming from a vague ethical and legal framework must be overcome so as to better operationalise and protect fundamental ethical values and fundamental rights in the market of artificial intelligence-driven legal services. Against this backdrop, the study demonstrates how possible solutions against ethics/market mismatches are provided by the legal system, which can work as a bridge vehiculating into the market practice of AI-based legal decision-making tools declared ethical principles, while preventing eventual chilling effects on the market. It thus shows how these need to be adequately matched and integrated with legal design requirements to maximise the resulting positive synergies within the market and thus avoid risks of ethical dilution. In this respect, a layered regulatory regime is proposed for artificial intelligence-driven legal services, of both public and private destination. This framework is meant to operationalise general ethical values and fundamental legal liberties within the more specific regulatory framework given by the European data protection, the Open Data, the European competition framework and the European Commission's newly proposed rules for artificial intelligence.

## I. Introduction

The penetration of artificial intelligence-based tools in the legal sector is

\* Assistant Professor of Law and Economics, Università Cattolica del Sacro Cuore, Faculty of Banking, Finance and Insurance Sciences, Milan.

moving forward, further accelerated by the exceptional needs brought about by the pandemic.<sup>1</sup> As a result of the occurring digitisation patterns, some strand of the literature has declared the end of the law, as we know it.<sup>2</sup> Although this statement may sound a drastic conclusion, it is certainly a provocation to be taken seriously.

Promoters<sup>3</sup> of these technologies support these developments, expecting far better access to justice, previously constrained by the prohibitive cost of legal advice<sup>4</sup> and highlighting inclusion effects on those strands of the population that would be ordinarily left outside the privilege of legal consultancy.<sup>5</sup> The major advantages that are identified are related to the lowering of operating expenses for legal research, time savings, and, as a result of these, the creation of greater opportunities of meeting and supporting citizens' and clients' needs. Moreover, the probabilistic computation of litigation success could determine a reduction in the cases that go to court.<sup>6</sup>

It is worth noting from the outset that the European Commission for the Efficiency of Justice (CEPEJ)<sup>7</sup> has expressed concerns on these tools and France has already outlawed some of them, for instance, banning and punishing the use of predictive litigation Artificial Intelligence (AI) for the purpose or effect of assessing, analysing, comparing or predicting judges' real or supposed professional practices.<sup>8</sup> Similarly, the proposed AI Act<sup>9</sup> adopts a deeply asymmetric approach,

<sup>1</sup> For an overview see A.F. Mainini, 'Il futuro immediato della Giustizia dopo il 12 maggio 2020 - Gli effetti della crisi sanitaria determinata dalla pandemia Covid-19' (12 May 2020), available at <https://tinyurl.com/3jupb7e7> (last visited 30 June 2022). See even before the Covid-19 outbreak, the considerations by B. Monarch, 'The Promise and Perils of Legal Technology in a Period of Economic Uncertainty' (8 May 2015), available at <https://tinyurl.com/5n82jwxz> (last visited 30 June 2022).

<sup>2</sup> X. Labbé, 'Robot. La fin du monde, la fin du droit ou la transition juridique' 2 *Recueil Dalloz*, 78 (2019).

<sup>3</sup> M. Juetten, 'The Future of Legal Technology: It's Not as Scary as Lawyers Think' *Forbes*, available at <https://tinyurl.com/3d2py9xv> (19 February 2015) (last visited 30 June 2022).

<sup>4</sup> K.N. Kotsoglou, 'Subsumtionsautomat 2.)- Über die (Un-)Möglichkeit einer Algorithmisierung der Rechtserzeugung' *Juristenzeitung*, 451 (2014); M. Engel, 'Erwiderung: Algorithmisierte Rechtsfindung als Juristische Arbeitshilfe' *Juristenzeitung*, 1096 (2014).

<sup>5</sup> M. Fries, 'Man Versus Machine: Using Legal Tech to Optimize the Rule of Law', available at <https://tinyurl.com/bdh4mwcx>, 8 (2016) (last visited 30 June 2022).

<sup>6</sup> This could thus counterbalance that what some strand of the literature has observed as an excessive optimism regarding litigation outcomes. O. Bar-Gill, 'The Evolution and Persistence of Optimism in Litigation' 22 *Journal of Economics & Organisation*, 490 (2006).

<sup>7</sup> Council of Europe-European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment' 64, 65 (2018).

<sup>8</sup> See Art 33, Loi no 2019- 222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (so-called Justice Reform Act), available at <https://tinyurl.com/bdjdf4kw> (last visited 30 June 2022). In particular: 'Les données d'identité des magistrats et des membres du greffe ne peuvent faire l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées. La violation de cette interdiction est punie des peines prévues aux articles 226-18, 226-24 et 226-31 du code pénal, sans préjudice des mesures et sanctions prévues par la loi no 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés'.

displaying some concerns about the use of AI systems by the judiciary while remaining silent on their use by attorneys or, more generally, other legal decision makers:<sup>10</sup> Annex III of the Proposal for an European Artificial Intelligence Act indeed lists under the high risk tools referred to in Art 6, para 2, systems that are employed in ‘administration of justice and democratic processes’ defined as ‘AI systems that are intended to assist a *judicial authority* in researching and interpreting facts and the law and in applying the law to a concrete set of facts’, along with several ‘AI systems intended to be used by law enforcement authorities’.<sup>11</sup> While attracting to a stricter regulatory regime those artificial intelligence tools directly used by the judiciary, the Proposal for a Regulation on AI leaves some severe regulatory uncertainties regarding those AI-based legal services that are addressed to the wider array of other private stakeholders active in the processes of legal decision-making, first of all law firms and legal consultant businesses but also independent administrative authorities.

In this way, the Act casts a shadow of suspects only onto the use of AI by public actors as judges and law enforcement authorities.<sup>12</sup> Strangely as it might sound, the same AI systems not used to assist a judicial authority ‘in researching and interpreting facts and the law and in applying the law to a concrete set of facts’ would not be considered as high risk. In other words, the same tools used by attorneys or a private arbitration centres would not be considered as high risk. The qualification evoked by ref. 40 that

‘such qualification [as high risk] should not extend, however, to AI systems intended for purely ancillary administrative activities that do not affect the actual administration of justice in individual cases, such as anonymisation or pseudonymisation of judicial decisions, documents or data, communication between personnel, administrative tasks or allocation of resources’

does not reduce the actual sharp limits only to judicial use of AI. Actually, such a

<sup>9</sup> European Commission, ‘Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts’, COM(2021) 206 final (21 April 2021).

<sup>10</sup> European Commission, ‘Annexes to the Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts’, COM(2021) 206 final (21 April 2021).

<sup>11</sup> *ibid*, Annex III, no 6 and 8, emphasis added.

<sup>12</sup> This is not surprising, since the overall framework of the proposed Artificial Intelligence Act, at least until the recent Council’s amendments, judges the use of artificial intelligence tools by public authorities as riskier than that of private players. This is evident in respect to the case of social scoring that has been at first banned by the European Commission only in the hands of public authorities under Art 5, para 1, lett. c), and then banned also in the case of private social scoring under Art 5, para 1, lett. c) of the Council’s version of the proposal. See Council of the European Union, ‘Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts - Presidency compromise text’, 2021/0106(COD) (29 November 2021).

limitation in itself creates a number of concerns because it unbalances the tools available to the judiciary versus legal practitioners and potentially creates, within the realm of the private sphere, a significant market unbalance between those professionals that can afford the use of AI systems and those that cannot do it.

In light of the persisting regulatory gaps in the field of artificial intelligence-driven legal services, this study questions which are the legal tools that are relevant to govern the current expansion of the correspondent market in a way that is consistent with ethical declarations. To these ends, it generally intends legal services run by artificial intelligence as every application based on machine learning techniques, which are destined to directly provide or more broadly support the delivery of legal assistance. As known, machine learning enables to predict results based on the identification of statistical patterns within a given datasets.

The analysis moves from the acknowledgment that machine learning models are being increasingly applied to textual data contained in legal materials for the prediction of outcomes regarding the legal position of citizens, in terms, for example, of discovery review, contract analytics and legal research. In this respect, the study gives account of ongoing transformations in the market of AI-supported legal services, with the aim of rooting the relevant regulatory framework in the market reality. In our understanding, the analysis related to the risks connected to the employment of AI-driven legal decision-making tools delivered by the market triggers the question whether the applicable ethical-legal framework provides sufficient tools for addressing the current developments in the market of AI-assisted legal services or whether additional sector-specific solutions need to be introduced.

In this respect, a gap is found between the blooming market reality and the ethical and legal perspectives.

To date, the debate has been harping mostly the tune of ethical constrains. This can be easily derived from documents that have been issued at EU and international level, as the Guidelines on a Trustworthy AI by the European Commission's High Level Expert Group on Artificial Intelligence<sup>13</sup> and, with specific regard to the legal sector, the Council of Europe's European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment.<sup>14</sup>

Also, more general documents, as the recent European Commission's European Strategy for Data (hereafter 'EU strategy for data')<sup>15</sup> and the White Paper on AI,<sup>16</sup>

<sup>13</sup> European Commission-High Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' (8 April 2019), available at <https://tinyurl.com/2p8z5dkb> (last visited 30 June 2022).

<sup>14</sup> Council of Europe-European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter on the Use of Artificial Intelligence' n 7 above.

<sup>15</sup> European Commission, 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of the Regions, A European Strategy for Data', COM/2020/66 final, (19 February 2020).

<sup>16</sup> European Commission, 'White Paper on Artificial Intelligence - A European Approach to Excellence and Trust', COM(2020) 65 final (19 February 2020).

place a particular emphasis on the need to consolidate an ethical framework for the employment of artificial intelligence. At national level, the experience of the German Datenethikkommission is a good example to recall.<sup>17</sup> In this landscape, the European Ethical Charter Charter on the Use of Artificial Intelligence in Judicial Systems and their environment is meant to provide a comprehensive guidance to justice professionals in the process of assimilation of artificial intelligence applications within the judicial system.<sup>18</sup>

In order to sharpen the normative strength of ethical declarations, these mentioned documents often rely on the concept of fundamental rights. Yet, most of the concerns expressed within the Charter do not seem to be addressed by the spontaneous drive of market forces nor by the European Union (EU) legislator.

In our opinion, the overemphasis on the ethical concerns while leveraging fundamental rights to define the ethics boundaries of AI in justice administration risks to mix up different levels of normativity, that is the ethical and the legal dimension, where only the latter provides enforceable rules.<sup>19</sup> On the legislative side, since the proposed Regulation on artificial intelligence addresses only the use by the judiciary, there is a high risk of undermining the regulatory needs for the private use of AI system in legal services especially if ethical considerations do not go along with an accurate knowledge of the reality of the developing markets in the legal services domain. At a third level, it seems that from the market side, producers of these AI-driven legal decision-making tools, in their race for reaching the competitive edge, do not take into adequate consideration ethical standards.

It thus appears that the emerging economic and technical reality of new technologies for legal decision making and the theoretical policy debates regarding the legitimacy of such applications are silently developing at parallel but non-communicating levels.

As here argued, the emerging gap between the three realities, that is the politico-ethical, the legislative and the market one, is destined to result in what has been defined in the literature as the phenomenon of ‘ethical dilution’<sup>20</sup> or ethical ‘washout’.<sup>21</sup> The concrete result of this is the lack of regulatory certainties

<sup>17</sup> See the Ethical Guidelines recently issued by the German Data Ethics Committee, Datenethikkommission, ‘Gutachten der Datenethikkommission’, available at <https://tinyurl.com/y51ra7x3f> (last visited 30 June 2022).

<sup>18</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above.

<sup>19</sup> G. Comandé, ‘Unfolding the Ethical Component of Trustworthy AI: a Must to Avoid Ethical Dilution’ *Annuario di diritto comparato e di studi legislativi*, 39, 62 (2020).

<sup>20</sup> *ibid*

<sup>21</sup> E. Bietti, ‘From Ethics Washing to Ethics Bashing - A View on Tech Ethics From Within Moral Philosophy’ *Journal of Social Computing*, 2 (2021). K. Hao, ‘In 2020, Let’s Stop Ethics Washing and Actually Do Something’, available at <https://tinyurl.com/5cz9b5bz> (last visited 30 June 2022). See also K. Yeung et al, ‘AI Governance by Human Rights-Centred Design, Deliberation and Oversight: An End to Ethics Washing’, in M. Dubber and F. Pasquale eds, *The Oxford Handbook of AI Ethics* (Oxford: Oxford University Press, 2019), available at <https://tinyurl.com/2s4x333b> (last visited 30 June 2022). See also B. Wagner, ‘Ethics as an Escape From Regulation: From Ethics Washing to Ethics

for stakeholders, which either remain inactive or turn to abuses, directly given by the exploitation of regulatory gaps.<sup>22</sup>

By placing emphasis on the ethical dimension, businesses may be left free to conceal themselves behind compliance with a vague ethical framework, reassuring users, while perpetrating their abuses. At the same time, they expose themselves to the risk of fines and liability actions,<sup>23</sup> causing a spill-over harm to society, as an end effect.<sup>24</sup>

Although these considerations may apply to every AI-driven market sector, the particular area of AI-based legal services is, overall, at risk of a doubled-edged ethical washing outcome, given on the one hand by the fact that ethical declarations do not accurately identify the legal provisions that shall substantiate relevant ethical principles in this specific context, and on the other hand by the fact that current legal provisions applicable to our case – first of all the Proposal for a European Regulation on Artificial Intelligence – are shaped in a way that do not adequately address ethical concerns.

The sensitiveness of the market for legal services requires a prompt realignment between a clear legal and ethical framework and the market practice in the field of AI-driven legal services. Such a realignment is urgently needed to avoid substantial risks for both citizens that are the addressees of AI-driven legal services and legal operators who come to interact with these tools in their legal practice. An unbalanced development of the market for these services might result not only in competitive hurdles and market abuses but also to undermine the basic tenets of the administration of and access to justice.

For these reasons, the uncertainties stemming from a vague ethical and legal framework must be overcome so as to better operationalise and protect fundamental ethical values and fundamental rights in the market of artificial intelligence-driven legal services. Against this backdrop, the analysis demonstrates how possible solutions against ethics/market mismatches are provided by the legal system regulating evolving digital markets in the legal sector. If properly implemented, the rules that govern internal market developments in the field of digital technologies can work as a bridge vehiculating into the market practice of AI-based legal decision-making tools declared ethical principles, while preventing eventual chilling effects on the market.

The perspective of the general framework that is consolidating for the

Shopping?', in E. Bayamlioglu et al eds, *Being Profiled: Cogitas Ergo Sum* (Amsterdam: Amsterdam University Press, 2018), 84, 88.

<sup>22</sup> G. Comandé, n 19 above, 39.

<sup>23</sup> For example, under data protection law. See Arts 82-83 of the General Data Protection Regulation. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119 (4 May 2016), hereafter GDPR.

<sup>24</sup> G. Comandé, n 19 above.

regulation of digital markets offers a starting standpoint so as to adequately match and integrate ethical principles with legal design requirements, maximise the resulting positive synergies within the market and thus to avoid risks of ethical dilution. In this respect, the enquiry proposes a layered regulatory regime for artificial intelligence-driven legal services, of both public and private destination. This framework is meant to operationalise general ethical values and fundamental legal liberties within the more specific regulatory framework given by the European data protection, the Open Data, the European competition framework and the European Commission's newly proposed rules for artificial intelligence.<sup>25</sup>

From a methodological standpoint, the proposed integrated approach between ethics and law aims to preserve their respective intrinsic roles without collapsing into a functional overlap. Nonetheless, it spots functional synergies between the high level of ethical and fundamental legal principles and existing market specific rules, so as to objectify general precepts into concretely actionable legal rights.

Under these premises, the study is structured as follows: a first section illustrates the more recent trends in the market of AI-related legal services and recalls applicable ethical principles that have been issued so far at European policy level; the second section detects the risks originating from the misalignment between the two levels; ultimately, the third section identifies the relevant provisions in the legal system for an ethically-sound development of AI-supported legal decision-making tools.

Overall, the study sets the analytical framework for future enquiries: open issues related to the difficult match between the market, legal, and ethical dimensions are ultimately highlighted, unveiling the challenges of further research in this field.

## II. Mapping the Policy Landscape and the Market of AI-Assisted Legal Decision-Making Tools

The development of a market of products and applications designed for the legal sector has been set as a goal by the European Commission in its recent Action Plan for e-Justice,<sup>26</sup> where a list of projects for implementation in the time frame 2019-2023 is considered. These projects concern, *inter alia*, the consolidation of a criminal court database,<sup>27</sup> the improvement of the EUR-lex search-engine,<sup>28</sup> the advancement of court decisions' accessibility<sup>29</sup> and the interconnection and

<sup>25</sup> European Commission, 'Proposal', n 9 above.

<sup>26</sup> European Commission, '2019-2023 Action Plan European e-Justice, 2019/C 96/05', available at <https://tinyurl.com/3536uxhm> (last visited 30 June 2022).

<sup>27</sup> *ibid* 12.

<sup>28</sup> *ibid* 13. For the literature see M. Ovádek, 'Facilitating Access to Data on European Union Laws' 3 *Political Research Exchange*, 1 (2021).

<sup>29</sup> European Commission, '2019-2023 Action Plan' n 26 above, 15.

interoperability of legal information published in EU websites.<sup>30</sup> In this perspective, also the Legivoc system is worth to be mentioned: it is a database of terms that should help Member States understand European Union laws and intended to ‘promote the semantic alignment of the vocabularies of EU Member States along with third States’.<sup>31</sup> As announced, the database constitutes a lexicon of legal terms that are readily usable for legal informatics projects developed for the purposes of improving accessibility to Member States laws and of advancing exchanges of legal information in the context of judicial and legal cooperation.<sup>32</sup>

For these last purposes, technologies for automatising the anonymisation and pseudonymisation of legal documents and especially court decisions are included in the agenda.<sup>33</sup>

In addition to these fields of action, the Commission also considers AI-based solutions for the analysis of Court decisions,<sup>34</sup> and the definition of use cases for blockchain technologies in the e-justice domain.<sup>35</sup> The development of chatbots, assisting the user and directing her/him in legal research, and especially in the research of relevant case law, is further envisaged.<sup>36</sup> It is worth noticing from the outset that these last fields of intervention are considered as high risk AI system in the proposed AI act. Other areas of proposed intervention regard the development of digital means for a faster communication between citizens, judicial and practitioners<sup>37</sup> which, to the contrary are not considered as high risk under the proposed Regulation on artificial intelligence.

As the proposed initiatives show, the European Commission is taking into account and promoting the digital transformations of the EU legal system(s), at both European level – for example through the proposed enactment of new EU portals<sup>38</sup> and the improvement of existing ones<sup>39</sup> – and national level, for example through the proposed interconnection of national legal information systems<sup>40</sup> and the planned automatization of national court decisions’ analysis.

Overall, the European planned lines of action aim to lay down the political foundations for a developing ‘e-justice’ market, which is to be fuelled by the sharing and aggregation of legally relevant data. Exactly for the purposes of enabling ‘innovative ‘gov-tech’, ‘reg-tech’ and ‘legal tech’” tools to support practitioners and

<sup>30</sup> *ibid* 14, 22.

<sup>31</sup> *ibid* 24.

<sup>32</sup> See Legivoc, available at <https://tinyurl.com/3rt7ufxn> (last visited 30 June 2022).

<sup>33</sup> European Commission, ‘2019-2023 Action Plan’ n 26 above, 15.

<sup>34</sup> *ibid* 17.

<sup>35</sup> *ibid* 19.

<sup>36</sup> *ibid* 17.

<sup>37</sup> *ibid* 21.

<sup>38</sup> See for example the proposed common search engine on the European e-Justice Portal, for advertisements of judicial sales published in the Member States. *ibid* 13.

<sup>39</sup> See the proposed development of new features for the e-Justice Portal, such as a central query tool. *ibid* 10.

<sup>40</sup> *ibid* 14.

other services of public interests, the European Strategy for Data envisages an outright ‘Common European Data Spaces for Public Administration’.<sup>41</sup>

In the intention of the European regulators this market shall have the very core objective of increasing data literacy in the legal sector, relating to the integration of traditional legal reasoning methodologies and services with new technological tools meant to support the former.<sup>42</sup> In this perspective, the mentioned programs are meant to variously assist both decision-makers, as judges or arbitrators, and the parties to a dispute. This means that the subjects targeted by the set plan are not only legal practitioners but also citizens without a legal expertise, whose access to legal services and thus, more in general, to justice is intended to be improved thanks to the disintermediation of legal knowledge the proposed tools offer.<sup>43</sup>

Nevertheless, as anticipated, an inherent tension can be observed between these shared goals and the suspect enshrined in the proposed AI act for the use of AI tools by the judiciary itself. This is why the considered European plan is structured around three different objectives: 1) the expansion of access to legal services facilitating as well the disintermediation of legal services; 2) the improvement of legal services by way of AI systems; and 3) the use of AI systems by the judiciary.

The sustained policy comes in tension when the last aim is target of actual legislation that encumbers with regulatory burdens specific uses of artificial intelligence in the legal sector, rendering them more difficult, if not impairing them at all.

Another mismatch can be identified between the European strategic program over e-justice and its high-level institutional perspective on the one hand and the actual targets of AI-driven legal services’ markets on the other hand: indeed, the market perspective provides additional insights on the deep ongoing transformations in the legal service domain, which in some cases goes beyond what is perceived at regulatory level. A whole array of new start-ups is offering technologies for the improvement or the augmentation of legal services.<sup>44</sup> Overall, emerging applications in the context of both private and public legal services reflect a tendency towards a legal system of ‘predictive justice’ using data mining methods and approaches perfectly fitting the definition of AI system in the AI Act.<sup>45</sup> Unlike in the movie ‘Minority Report’, the examples that will be provided below do not

<sup>41</sup> European Commission, ‘Communication’ n 15 above, 22-23.

<sup>42</sup> As well highlighted in the European Strategy for data, data literacy is closely related to a shift in the competences needed in order to correctly implement and understand the results of employed technologies. *ibid.*, 10,11, 20,21.

<sup>43</sup> On the disintermediation of legal services, see P. Heudebert and C. Leveneur, ‘Blockchain, Disintermediation and the Future of the Legal Professions’ 4 *Cardozo International & Comparative Law Review*, 275 (2020).

<sup>44</sup> See J. Armour et al, ‘Augmented Lawyering’ *ECGI Working Paper Series in Law n. 558/2020*, available at <https://tinyurl.com/a7shkvxn> (last visited 30 June 2022).

<sup>45</sup> European Commission, ‘Proposal’ n 9 above, Art 3 and annex I.

predict any outcome but only provide a ‘forecast’ of what could happen based on a number of analysed variables.

To date no actual legal reasoning analysis is permitted by technology. Thus, predictive justice relates to tools that anticipate what could be an outcome leveraging various forms of ‘statistical’ and knowledge discovery methods. Of course, this shift of emphasis does not only illustrate the vanity of the hope of replacing judges and lawyers but rather makes it clear that more than predictive justice we should speak of ‘predictable’ justice capable of analysing and ‘imagining’ possible legal solutions as the aimed goal. The result is that of a justice system that enhances the collective intelligence of actors through the tools of artificial intelligence.

However, to date no marketed technology is capable of autonomously reproducing a human decision based on a ‘real’ legal reasoning. Hence, the mentioned technologies can only provide a support to legal professionals or more in general to citizens that need to be integrated with solid ‘traditional’ domain knowledge. Moreover, without proper legal analysis the actual ‘predictions’ not only can be misleading, resulting in over or under litigation for instance but can actually be manipulated to drive legal actions even purposely in the wrong direction. In this perspective, future lines of development of the considered market should perhaps move from the persisting needs to integrate automated-driven tools – as search tools or information aggregation tools –, with applications that automatise the representation of sectorial domain knowledges.

As the examples below show, AI-based applications in the context of both private and public legal services are capable of structurally innovating and changing the legal profession, overturning in many cases the competences traditionally required in the legal sector.<sup>46</sup> At a deeper level, these applications *de facto* force changes also at the education level and at the institutional level, since in many countries legal profession is regulated/protected, granting a certain amount of exclusivity in providing legal services. In these cases, innovations must also face these regulatory hurdles: if automated legal analysis is offered mostly with the interaction of data scientists, software engineers reserving legal advice to lawyers might appear anachronistic but needed.

Among the various applications offered on the market, a distinction needs to be made between those artificial intelligence-based tools destined to private purposes – eg for the support of law firms’ activities or of citizens’ legal queries – and those designed for public purposes, eg for the automatization of specific tasks in judicial decision-making. The mentioned distinction is relevant because the different private or public interests involved in the use of AI-based applications in the legal sector raise different legal and ethical issues and are treated radically differently by the proposed AI Act. Below we sketch a possible, although non exhaustive, categorization.

<sup>46</sup> J. Armour et al, n 44 above, 57.

### III. AI-Based Tools for Private Legal Services

1. *Assistance to Law Firms and Legal Consulting Businesses Through Predictive Coding*: digital tools are transforming the law office management, through user-friendly interfaces and electronic communication means with courts or other attorneys. The particular situation of the Covid-19 pandemic and the spread of smart working solutions has accelerated these developments. Also, digital support tools for contract or asset analysis are becoming more powerful. Early support tools in this sense are search engine tools as Westlaw; LexisNexis; Beck-online, which have had a substantial influence on legal advice and on state jurisdiction.<sup>47</sup> Developments in this sense are related to the use of AI-driven artificial intelligence software to conduct legal research, as occurs with advanced case-law search engines and predictive analytic tools.<sup>48</sup> An example in this respect is given by Ross AI, which uses natural language processing to find relevant results and to provide meaningful ranking of results.<sup>49</sup> Dorothei AI uses natural language processing to search patent filings.<sup>50</sup> These softwares enable to support legal advice, both in terms of fast retrieval of guiding principles of case precedents, and of interpretation and application of all the cases.<sup>51</sup>

Apart from legal research, automated driven tools are also changing the way legal advice is delivered. An increasing number of startups is offering automated online legal consultations services, as Justia<sup>52</sup> and Avvo.<sup>53</sup> Other services are designed to match lawyers with clients, without the expensive intermediation of a law firm, as UpCounsel,<sup>54</sup> Lawgives<sup>55</sup> and LegalHero.<sup>56</sup> The software eBrevia<sup>57</sup> is structured for document review, ‘contract analyser’ and ‘diligence accelerator’, specifically designed for lawyers to perform due diligence review for mergers and acquisitions. Similarly, Wevorce<sup>58</sup> is meant to simplify divorce processes, through personalised algorithms that seek to streamline asset division, form completion and other divorce-related work.<sup>59</sup> All the mentioned services fall under the category

<sup>47</sup> M. Fries, n 5 above, 8.

<sup>48</sup> These two categories of AI-based legal services are mentioned by the Council of Europe-European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 17.

<sup>49</sup> See Ross Intelligence, available at <https://tinyurl.com/y3t8kh7m>.

<sup>50</sup> Dorothy AI, available at <https://tinyurl.com/5dejdwcw>.

<sup>51</sup> The Portal Geblitz, available at <https://tinyurl.com/we9j8hpj>, has collected a substantial amount of information about sporadic measurement errors of individual speed cameras, and can challenge any overspeed fines originating from these cameras.

<sup>52</sup> Justia, available at <https://tinyurl.com/42pj2x9x>.

<sup>53</sup> Avvo, available at <https://tinyurl.com/3r999d85>.

<sup>54</sup> UpCounsel, available at <https://tinyurl.com/58fwzlnk>.

<sup>55</sup> Lawgives, available at <https://tinyurl.com/4y5bdtmb>.

<sup>56</sup> LegalHero, available at <https://tinyurl.com/yeywker5>.

<sup>57</sup> eBrevia, online available at <https://tinyurl.com/2pfv8bky>.

<sup>58</sup> Wevorce, online available at <https://tinyurl.com/4utud8ap>.

<sup>59</sup> The recalled AI-driven programs are listed by A. McPeak, ‘Disruptive Technology and the Ethical Lawyer’ 50 *University of Toledo Law Review*, 457, 461 (2019).

of so-called 'e-discovery' or 'technology-assisted review' technologies,<sup>60</sup> which have the distinctive features of quickly retrieving relevant information from a vast number of documents on the basis of predetermined classifications.<sup>61</sup> It is worth highlighting that these technologies are not only used in the private sector by law firms and legal consulting businesses, but also by public agencies. In the United States, for example, the Antitrust Division of the DoJ is already making use of e-discovery technologies in the course of mergers and acquisitions investigations.<sup>62</sup>

2. *Simple Serial Litigation*: technology is also used by institutional actors, as insurance companies, which employ analytical systems to collect facts of a case before getting in touch with the policy owner and cut in this way their legal expenses. By collecting facts of a case before getting in touch with the policy owner and fund the expenses of an attorney or even the courts. There are also specialized businesses, which process information collected from understandable online questionnaires to assess cases and litigate for a low fee, as in the case of the challenging of speeding fines and of the claim of lump-sum damages for flight delays.<sup>63</sup> These predictive systems could also be displayed in review to the parties, allowing them to decide whether they want to stick with their claim or withdraw it without bearing the court expenses. Moreover, they are being employed by lawyers for the purposes of calculating the probabilities of success of a certain litigation; as well as for the purposes of identifying and selecting the aspects of a case upon which it is convenient to work on for a successful outcome

3. *Assessment of Cases by Non-Lawyers* are equally being facilitated by technology advancements: new softwares are directly addressed to end-customers, regardless of whether these are a legal experts, or consumers, or a small business

<sup>60</sup> S. Gobbato, 'Procedure di e-discovery e tutela dei dati personali: una questione di metodo' *Media Laws*, available at <https://tinyurl.com/5e2xdr4z> (last visited 30 June 2022).

<sup>61</sup> A definition of these technologies has been given by Judge A.J. Peck in the ruling *Da Silva Moore v Publicis Groupe et al*, no 1:2011cv01279 – Document 96 (S.D.N.Y. 2012), para 3-4 where the Judge defines 'predicting coding' technologies as 'tools (different vendors use different names) that use sophisticated algorithms to enable the computer to determine relevance, based on interaction with (ie, training by) a human reviewer. Unlike manual review, where the review is done by the most junior staff, computer-assisted coding involves a senior partner (or [small]team) who review and code a 'seed set' of documents. The computer identifies properties of those documents that it uses to code other documents. As the senior reviewer continues to code more sample documents, the computer predicts the reviewer's coding (Or, the computer codes some documents and asks the senior reviewer for feedback). When the system's predictions and the reviewer's coding sufficiently coincide, the system has learned enough to make confident predictions for the remaining documents'. See also A.J. Peck, 'Search, Forward. Will Manual Document Review and Keyword Searches Be Replaced by Computer-assisted Coding?' *Law Technology News*, available at <https://tinyurl.com/yc8fhc5t> (last visited 30 June 2022).

<sup>62</sup> T. Greer, 'Electronic Discovery at the Antitrust Division: An Update', available at <https://tinyurl.com/2p8xb92t> (last visited 30 June 2022).

<sup>63</sup> This is the case of the services offered by the firms *EUclaim*, online available at <https://tinyurl.com/2z55rkjd>; *flightright*, online available at <https://tinyurl.com/2p8kr545>; *Fairplane*, online available at <https://tinyurl.com/49bnzzcm>, which growingly facilitate their case assessment on the basis of the analysis of the information retrieved from flight tracking or the automatic analyses of weather reports.

without a legal department, independently helping the client to analyse and prepare legal documents, as for example offered by the start-ups Catalystsecure<sup>64</sup> and Leverton.<sup>65</sup> The services provided by these companies digitise legal documents and display online forms ready to be downloaded and used by users, as judicial demands or tenancy agreements. Furthermore, these technologies can help prepare deeds and automate parts of the legal case assessments, as offered by Lexalgo;<sup>66</sup> drafting contracts, wills or other legally relevant statements, as enabled by Legal/Zoom;<sup>67</sup> RocketLawyer<sup>68</sup> or Janolaw.<sup>69</sup>

Other softwares operate a categorisation of contracts according to different criteria, detecting divergent or incompatible contractual clauses, or providing ‘chatbots’ informing litigants or supporting them in their legal proceedings.<sup>70</sup> Startups as Legalsifter,<sup>71</sup> Seal Software<sup>72</sup> and Exigent Group<sup>73</sup> employ AI for the purposes of helping clients to understand and assess drafted contracts. Similarly, the Claudette system developed by the European University Institute in Florence is an automated detector of potentially unfair clauses.<sup>74</sup> Other AI-driven tools assess the risks of success or defeat, as well as the litigation risks for the client:<sup>75</sup> Robot lawyer Lisa,<sup>76</sup> provides legal expertise automation, and is capable of issuing basic legal advice, creating legal documents as contracts. Ultimately, some services based on blockchain technologies offer to conduct automated transactions without the presence of lawyers, in the form of smart contracts.<sup>77</sup>

#### IV. AI-Based Tools for Public Legal Services

##### 1. *Judicial Rights Enforcement*: technology tools for judicial decision-making

<sup>64</sup> Catalystsecure, available at <https://tinyurl.com/5ak36f23>.

<sup>65</sup> Leverton, available at <https://tinyurl.com/bdhnd6j6>.

<sup>66</sup> Lexalgo, available at <https://tinyurl.com/2p8r6v5r>.

<sup>67</sup> LegalZoom, available at <https://tinyurl.com/vucek3hj>.

<sup>68</sup> RocketLawyer, available at <https://tinyurl.com/zprv6z8n>.

<sup>69</sup> Janolaw, available at <https://tinyurl.com/2p9y8bm6>.

<sup>70</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 17.

<sup>71</sup> Legalsifter, available at <https://tinyurl.com/4pf4afa5>.

<sup>72</sup> Seal Software, available at <https://tinyurl.com/yckjm2py>.

<sup>73</sup> Exigent Group, available at <https://tinyurl.com/2p9adtjc>.

<sup>74</sup> Caludette, available at <https://tinyurl.com/zw2j59nb>. For an overview of Claudette’s features see G. Sartor et al, ‘Claudette Meets GDPR: Automating the Evaluation of Privacy Policies Through Artificial Intelligence’ *Study Report, Funded by The European Consumer Organisation (BEUC)* (2 July 2018).

<sup>75</sup> The issue of the calculation of risks in the context of judicial proceedings had been anticipated and assessed by a strand of the literature well before the wave of digital transformations. See H. Eidenmüller, ‘Prozeßrisikoanalyse’ 113 *Zeitschrift für Zivilprozess*, 5 (2000).

<sup>76</sup> Robot Lawyer Lisa, available at <https://tinyurl.com/27m6u33d>.

<sup>77</sup> J. Eyre, ‘Blockchain ‘Smart Contracts’ to Disrupt Lawyers’ *Financial Review*, available at <https://tinyurl.com/pr42bx5> (last visited 30 June 2022). See more in general, K. Werbach and N. Cornell, ‘Contracts *ex Machina*’ 67 *Duke Law Journal*, 313 (2017).

are to be contextualised in the broader topic of developments regarding the digitization of the public administration.<sup>78</sup> Judicial analytics imply the analysis of docket entries, case opinions, oral argument text, or other inputs to gain insights into judicial decision-making.<sup>79</sup> The more sophisticated technologies in this respect provide predictions regarding judicial decisions, making it possible to predict a case outcome from a judge's standpoint and to assess faster evidence.<sup>80</sup> An example of judicial analytics is given by the software developed by the company Gavelytics, which detects whether a judge would be favourable for a particular litigant, using data of precedents, judicial workloads and biographical information.<sup>81</sup>

These kinds of platforms can be classified as tools for what is increasingly known as 'predictive justice',<sup>82</sup> in which data mining techniques are employed for the purposes of classifying decisions or subjects based on their specific features, targeting them through a specific variable upon which the outcome of a litigation or the behaviour of a certain individual is calculated.<sup>83</sup> These predictive systems are based on statistical elaborations of employed terms, revealing the frequency of the occurrence of specific groups of terms.<sup>84</sup>

In the criminal law sector, these tools can be employed for the prediction of crimes; the prediction of the risk of recidivism; the identification of future criminals or victims.<sup>85</sup> In this regard, predictive systems may help mapping the elements of an investigation, supporting human experience with an integrated

<sup>78</sup> See D. Freeman Engstrom and D.E. Ho, 'Algorithmic Accountability in the Administrative State' 37 *Yale Journal on Regulation*, 800 (2020); G. Schneider, 'The Algorithmic Governance of Administrative Decision-Making: Towards an Integrated European Framework for Public Accountability' *Eurojus- Special Issue Big Data and Public Law: New Challenges Beyond Data Protection*, 126 (2019); C. Benetazzo, 'Intelligenza artificiale e nuove forme di interazione tra cittadino e pubblica amministrazione', available at *federalismi.it*, 27 May 2020; G. Tuzet, 'L'algoritmo come pastore del Giudice? Diritto, tecnologie, prova scientifica' *Media Laws*, available at <https://tinyurl.com/wek3uzct> (last visited 30 June 2022).

<sup>79</sup> A. McPeak, n 59 above, 464. These technologies have been also object of European projects. See F. Romeo et al, 'CREA Project – Conflict Resolution Equitative Algorithms', available at <https://tinyurl.com/4jbc3b82> (last visited 30 June 2022).

<sup>80</sup> We refer for instance to services such as Lex Machina, available at <https://tinyurl.com/4t99j72s>, which provides an analysis of parties, judges and counsel, the French Case Law Analytics, available at <https://tinyurl.com/57k2mbd2>.

<sup>81</sup> Gavelytics, available at <https://tinyurl.com/ykxfes4t>. For the literature on this point, see S.B. Starr, 'Evidence-based Sentencing and the Scientific Rationalization of Discrimination' 66 *Stanford Law Review*, 803 (2014) where the Author defines criminal justice predictive systems as 'evidence-based methods'.

<sup>82</sup> As the CEPEJ Ethical Charter explains, a predictive system is a tool that announces what will happen in advance of future events. Council of Europe-European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment', n 7 above, 29-30.

<sup>83</sup> F. Romano et al, 'The Challenges of Legal Analysis Between Text Mining and Machine Learning' *JADT 2020: 15es Journées internationales d'Analyse statistique des Données Textuelles*, available at <https://tinyurl.com/y4zv6vua> (last visited 30 June 2022).

<sup>84</sup> Predictice, available at <https://tinyurl.com/2dcfcfw>.

<sup>85</sup> European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter' n 7 above, 48.

analysis of collected and available data. Examples of such tools are given by the Compas algorithm,<sup>86</sup> and the Hart Assessment Risk Tool (Hart).<sup>87</sup> In the field of civil law, conversely, predictive systems could be employed to measure the separation and divorce alimony.<sup>88</sup>

The employment of artificial intelligence for the ‘prediction of judicial decision-making is well suited for civil law countries, where the structure of the argumentation is well defined by the law. This structure is more easily replicated by computational systems, allowing judges to quickly spot the issues and legal questions underlying the case to be decided.

Overall, these tools can provide an important analytical support for judges, offering quantitative or qualitative insights over their decision-making processes. Through these systems judges and lawyers could be facilitated in finding the cases with identical or similar arguments and in using text modules. This could give advantages in terms of uniformity in case law, especially in respect to the case law of lower jurisdictions. These last considerations clearly sustain the European Commission’s policy favouring investment in these tools, although the proposed AI regulation regards with suspicion their use by the judiciary in actual cases.

2. *Alternative Dispute Resolution Systems*: the digitization of alternative dispute resolution mechanisms is becoming particularly relevant in the field of consumer services. In this case, the EU has pioneered the ODR platform, recording consumer complaints online, forwarding them to a dispute resolution body, and enabling the parties to conduct the negotiation process completely online.<sup>89</sup> According to a strand of the scholarship, the European consumer dispute resolution platform could be considered as a forefather of an outright online court.<sup>90</sup> Yet, once again there is a tension between the opening to the Online Dispute Resolution (ODR) market using AI solutions and the impossibility to use similar mechanisms by the judiciary itself. Note that, for instance, an AI system that would analyse the case at hand to advise the judge to send the case for a mediation attempt would be considered as a high risk one if used by a court<sup>91</sup> while it could seamlessly be fostered in contractual clauses.<sup>92</sup> Incidentally,

<sup>86</sup> Eg the Compas algorithm. J. Larson et al, ‘How we Analyzed the COMPAS Recidivism Algorithm, Pro Publica’, available at <https://tinyurl.com/3pdrpkye> (last visited 30 June 2022).

<sup>87</sup> See for example the Hart algorithm employed by the Durham Police and Cambridge University. For the literature see M. Oswald et al, ‘Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality’ 27 *Information & Communications Technology Law*, 223, 250 (2018).

<sup>88</sup> F. Romano et al, n 83 above.

<sup>89</sup> European Commission, ‘Online Dispute Resolution’, available at <https://tinyurl.com/mryuwbkf>.

<sup>90</sup> M. Fries, ‘Verbraucherrechtsdurchsetzung’ (Tübingen: Mohr Siebeck, 2016), 200, assessing the broader claim regarding whether ADR systems contribute to effectively pursue consumer rights enforcement.

<sup>91</sup> D. Thompson, ‘Creating New Pathways to Justice Using Simple Artificial Intelligence and Online Dispute Resolution’ *Osgoode Legal Studies Research Paper Series*, available at <https://tinyurl.com/5xxftkj2> (last visited 30 June 2022).

the latter would come at odds with the concern of the CEPJ about ‘possible violations of Arts 6, 8 and 13 of the European Convention on Human Rights’<sup>93</sup> for the risk of confusion between a court assessment and an alternative (out of court) dispute resolution mechanism. For these reasons, these particular systems are considered by the CEPEJ as ‘possible uses, requiring considerable methodological precautions’.

## V. Mapping the Ethical Principles for AI-Assisted Legal Services

The ethical framework applicable in the EU to AI-driven legal tools can be found at a general level in the Guidelines on a Trustworthy AI by the European Commission’s High Level Expert Group on Artificial Intelligence and, with specific regard to the legal sector, in the Council of Europe’s European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment. The Charter is meant to provide guidance to justice professionals in the process of assimilation of legal technologies within the legal system. Although specifically designed for guiding policy makers and justice professionals in the development of AI in national judicial processes, the ethical framework is believed to be applicable in analogy also to automated-driven technologies applied in the private sector of legal services.

An accurate analysis of the two charters shows the existence of a common set of principles, which are directly substantiated in i) the principle of quality and security of employed datasets; ii) principle of non-discrimination and equality; iii) the principle of fairness; iv) the principle of transparency; v) principle of ‘under user control’.

As a general premise it can be said that the first three principles assure that the considered legal technologies structurally embed specific values, especially in terms of non-discrimination and equality. The other two principles, conversely, assure that these values are externally verifiable and supervised by human subjects.

Ultimately, all the mentioned principles point to the overarching principle of human-centrism and autonomy in the use of AI-assisted legal decision-making tools.<sup>94</sup> This principle has a central importance in respect to artificial intelligence systems for legal decision making. It indeed requires that legal professionals maintain an autonomous judgment in respect to what is suggested by the automated system. This means that the subjects that interact with these technologies need to keep full and effective control over the final determinations, and let technologies

<sup>92</sup> This is well illustrated by the Cyberjustice project in Quebec, available at <https://tinyurl.com/pdp6fwza> (last visited 30 June 2022).

<sup>93</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 46, 47.

<sup>94</sup> See L. Floridi et al, ‘AI4People- an Ethical Framework for a Good AI Society: Opportunities, Risks, Principles and Recommendations’, available at <https://tinyurl.com/2s3avau3> (last visited 30 June 2022).

complement and empower their decision making without losing their autonomy. In this positive perspective, the CEPEJ Ethical Charter underlines how automated-driven systems should amplify legal professionals' decisional space, by constructively supporting the conduction of legal tasks through the provision of analytical evidence.<sup>95</sup> A different deployment of the same principle can be envisaged when the technologies are offered to non-experts: in this case, the principle of autonomy shall safeguard the decision-making space of subjects that are not familiar with the outputs rendered by the machine.

In order to guarantee citizens' autonomy in using AI-assisted legal decision-making tools, the considered principle also requires that results obtained are interpretable and understandable to assure control by the user over the machine. From this further perspective, the principle of autonomy requires the constant supervision of humans over the functioning of employed technologies. This relates also to the actual ability of reviewing issued decisions and in particular the data that have grounded a specific outcome by overcoming the so-called automation bias.

Interestingly, the Guidelines link the said principles to the fundamental rights of equal access to justice and to a fair trial in the changing legal system.<sup>96</sup> In this perspective, the High-Level Expert Group presents the mentioned principles as the ethical formants of automated legal decision-making processes in which rule of law, due process and equality before the law are cherished.<sup>97</sup> The assumption is indeed that the protection of fundamental rights is not only a legal but also a moral entitlement.<sup>98</sup>

The Guidelines do not however provide practical guidance as to how effectively secure listed ethical values and thus protect related fundamental rights in the considered artificial intelligence technologies. Thus, they set the general goal of pursuing the mentioned ethical values and connected fundamental rights, without tracing any patterns for the achievement of such objectives.<sup>99</sup> In the absence of more elaborated methodological instructions, crucial ethical problems related to artificial intelligence, as those related to 'trolley dilemmas',<sup>100</sup> the algorithmic decision-making superiority or inferiority to human decisions routines,<sup>101</sup> or

<sup>95</sup> European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter' n 7 above, 12.

<sup>96</sup> This is highlighted by D.B. Wilkins and M.J. Esteban, 'Taking the 'Alternative' out of Alternative Legal Service Providers: Re-mapping the Corporate Legal Ecosystem in the Age of Integrated Solutions' 5 *The Practice*, available at <https://tinyurl.com/2p86tp3u> (last visited 30 June 2022).

<sup>97</sup> European Commission-High Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' n 13 above, 11.

<sup>98</sup> *ibid*

<sup>99</sup> M. Veale, 'A Critical Take on the Policy Recommendations of the EU High-Level Expert Group on Artificial Intelligence' 1 *European Journal of Risk Regulation*, 10 (2020).

<sup>100</sup> J. Cows, 'AI and the 'Trolley Problem'', available at <https://tinyurl.com/2p9ndf33> (last visited 30 June 2022).

<sup>101</sup> J. Zerilli et al, 'Algorithmic Decision-Making and the Control Problem' 29 *Minds and Machines*, 555, 578 (2019).

the ‘hidden’ social and ecological costs of AI systems<sup>102</sup> remain unsolved.

The establishment of these ethical principles has thus moved in disconnection with the law,<sup>103</sup> and without taking into account what a strand of the literature has referred to as ‘the question of problem framing’.<sup>104</sup> This relates to the identification, at the same policy level in which ethical principles have been issued, of the ‘problems’ given by the unethical design and practical employment of machine learning-based technologies in specific sectors, as the ones employed in the legal sector.

In this respect, the technical assessment and understanding of those tools surely offers important insights to understand the functioning of AI-based legal decision-making tools.<sup>105</sup> The relevance of such an assessment has been well highlighted in the case of the Compas algorithm, in relation to which the 2016 Propublica Investigation revealed the discriminatory evaluation of African American defendants’ recidivism rate.<sup>106</sup>

To advance the awareness over the effects on society of automated prediction models, technical tools, as the so-called ‘Ethical Explorer’<sup>107</sup> or Facebook’s ‘Fairness Flow’,<sup>108</sup> have been elaborated for guiding developers and product managers in ‘building solutions that avoid the potential downsides of technology’ and thus in developing ‘responsible tech’ solutions.<sup>109</sup>

Moreover, businesses themselves have started to decline general ethical principles into their own corporate realities in the form of ethical charters, as the one released by Microsoft<sup>110</sup> and Google,<sup>111</sup> or codes of conduct.<sup>112</sup> In other cases, *ad hoc* ‘AI ethical committees’ have been directly established within the internal organisation of AI producers, with monitoring and supervisory tasks over the

<sup>102</sup> T. Hagendorff, ‘The Ethics of AI Ethics- An Evaluation of the Guidelines’ 30 *Minds and Machines*, 30, 104 (2020).

<sup>103</sup> Talking about a first ‘wave of movement’ focusing on ‘ethics over law’, C. Kind, ‘The Term ‘Ethical AI’ is Finally Starting to Mean Something’ *Venturebeat*, available at <https://tinyurl.com/2n6kr2wc> (last visited 30 June 2022).

<sup>104</sup> M. Veale, n 99 above, 1-10.

<sup>105</sup> C. Kind, n 103 above, talks about a second wave of ethical AI in which data and computer scientists ‘sought to promote the use of technical interventions to address ethical harms’.

<sup>106</sup> J. Angwin et al, ‘Machine Bias’ *ProPublica* available at <https://tinyurl.com/fvwx68rh> (last visited 30 June 2022), where it was found that the Compas algorithm was rating black defendants worse than white ones.

<sup>107</sup> See Ethical Explorer, available at <https://tinyurl.com/46p66r36>.

<sup>108</sup> See D. Gershgorin, ‘Facebook Says It Has a Tool to Detect Bias in Artificial Intelligence’ *Quartz*, available at <https://tinyurl.com/4ah5n3xb> (last visited 30 June 2022). Similarly see IBM, ‘Introducing AI Fairness 360’, available at <https://tinyurl.com/2p8Suaxyk> (last visited 30 June 2022).

<sup>109</sup> Ethical Explorer, n 107 above.

<sup>110</sup> Microsoft, ‘Microsoft AI Principles’, available at <https://tinyurl.com/2rbxmv45> (last visited 30 June 2022).

<sup>111</sup> See <https://tinyurl.com/yck2w5kt> (last visited 30 June 2022).

<sup>112</sup> See on the issue, H. Hilligoss and J. Fjeld, ‘Introducing the Principled Artificial Intelligence Project’ *CyberLaw Clinic*, available at <https://tinyurl.com/y3tdyvph> (last visited 30 June 2022). It is a project conducted by Harvard Berkman Klein Center that has mapped Ethical Principles and Guidelines issued by both public and private stakeholders between 2016 and 2019.

ethical countenance of developed AI tools.<sup>113</sup>

Nonetheless, all the mentioned examples rely on self-regulation for the purpose of conforming market efforts to ethical principles and, as stressed for example by Google, the referred ethical principles do not have a universal value but reflect those of the self-regulated company: ‘we will incorporate *our privacy principles* in the development and use of our AI technologies’.<sup>114</sup> However, self-regulation in the field of ethical AI is more related to concerns regarding producers’ ethical reputation than to those related to an effective implementation of ethical values.<sup>115</sup> Indeed, those initiatives are in most of the cases not overseen by any public agency and thus lack of a fundamental feature, that of enforceability.<sup>116</sup> In this perspective, they could encourage, rather than mend, ‘ethical washout’ outcomes.

A merely apparent compliance with ethical principles entails substantial risks in all sectors in which AI-driven tools are adopted. In the field of legal decision-making, nonetheless, these risks take up a particular shape, which is worth to be enquired more in depth. The acknowledgment of the peculiar risks related to that what we have defined as the ‘ethics/market mismatch’ in the development of AI-assisted legal services, suggests the urge to find viable solutions for the practical implementation of ethical principles.

After having mapped the risks resulting from ‘ethical dilution’ threats in the market for legal technologies, we will delve into the identification of patterns relevant for bridging market and ethics. Contrary to what some strand of the literature<sup>117</sup> and corporations themselves<sup>118</sup> are lately suggesting, we will demonstrate how these bridges do not rest on a more accurate socio-technical assessment of AI’s functioning, but rather on the enforceable rules provided by the European legal framework regarding emerging digital technologies, as artificial intelligence, and the data that fuels these.

<sup>113</sup> In this respect, it is worth to recall that Google announced the establishment of an external advisory council for the responsible development of AI in March 2019. The council was nonetheless removed just after one week. See E. Bietti, ‘From Ethics Washing’ n 21 above, 1.

<sup>114</sup> S. Pichai, ‘AI at Google: Our Principles’, available at <https://tinyurl.com/2kmlhn3p2> (last visited 30 June 2022).

<sup>115</sup> E. Bietti, ‘From Ethics Washing to Ethics Bashing’ n 21 above, 6. In respect to ethics committees, the Author highlights how these are mostly influenced by the management and also dependent on company funding. Moreover, no disclosure requirements regarding these council’s decision-making processes are in place. On the issue, see A. Papazologou, ‘Silicon Valley’s Secret Philosophers Should Share Their Work’ *Wired*, available at <https://tinyurl.com/52r96a32> (last visited 30 June 2022).

<sup>116</sup> G. Comandè, n 19 above. Of course, the lack of enforceability holds true as long as the ethical reference is not understood by regulators (such as the American FTC) as actual binding policies whose violation triggers its intervention.

<sup>117</sup> C. Kind, n 103 above.

<sup>118</sup> K. Johnson, ‘Microsoft Researchers Create AI Ethics Checklist With ML Practitioners From a Dozen Tech Companies’, available at <https://tinyurl.com/hdf4ywnm> (10 March 2020).

## VI. Mapping Market/Ethics Mismatches and Risks in AI-Assisted Legal Services

The illustrated developments in the market and the growing employment of these technologies in the legal practice, come along with some risks, which need to be carefully considered. Before engaging in the effort of identifying these risks, it is worth recalling that the same Council of Europe's Ethical Charter appears to consider artificial intelligence technologies employed in judicial systems in accordance with a risk-based approach.<sup>119</sup> Following this approach, the Council welcomes the improvement of legal decision making through the employment of what are considered as low-risk technologies: this is the case of visualisation techniques displaying data in a more efficient way;<sup>120</sup> or of the application of machine learning techniques in the field of natural language processing that operates based on key words or by linking various sources, as constitutional and conventional sources, case law and scholarship.<sup>121</sup> Among low-risk applications there are also those tools, which enlarge the scope of accessibility to legal information and to legal expertise, as chatbots and all those technologies that have the effect of disintermediating legal knowledge. Furthermore, the Council promotes the developments of those tools that provide indicators in respect to the performance of judicial systems, and that are thus strategically relevant for the conduction of qualitative and quantitative evaluations, which can potentially guide systemic reforms or, even before, address justice departments' re-organisation plans.<sup>122</sup>

Note, however, that these very same tools lend themselves to an extensive control over the judiciary and can be easily unfold in a sort of chilling effect on judges by way of stimulating conformity to previous judgments to boost statistical outcomes. In a sense they are precursors or enablers of tools actually profiling judges (and attorneys), AI based tools which are deemed<sup>123</sup> 'uses to be considered following additional scientific studies'.

Other artificial intelligence-based applications, conversely, need to be approached taking methodological precautions, in terms of technical structure and legal compliance, assuring a full protection of fundamental rights to subjects involved, both on the side of litigants and of legal professionals.<sup>124</sup> Among these 'riskier' tools, the CEPEJ lists applications that automatise the liquidation of

<sup>119</sup> European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter' n 7 above, 16-18 and 59-63.

<sup>120</sup> *ibid* 64. Highlighting the persuasive power of visualisation techniques, R. Ducato, 'De iurisprudencia picturata: Brief Notes on Law and Visualisation (editorial)' 7 *Journal to Open Access to Law*, 1, available at <https://tinyurl.com/y2s5ycnp> (last visited 30 June 2022).

<sup>121</sup> European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter' n 7 above, 64.

<sup>122</sup> *ibid*

<sup>123</sup> *ibid*

<sup>124</sup> *ibid* 64-65.

damages in civil proceedings; tools providing alternatives to the judicial resolutions of controversies, as online dispute resolution tools. The fully automated nature of many of those applications, as well as the frequent lack of information regarding the absence of involvement of a real court, pose substantial threats to the protection of the right to a fair trial under Art 6 ECHR and of the right to an effective remedy enshrined in Art 13 ECHR. Ultimately, as anticipated, also those tools that come to profile legal practitioners, be it judges, lawyers or consultants, are to be included in the high-risk category.<sup>125</sup> As already stated this categorisation well echoes what has been finally enshrined in the proposed AI Act, which includes in the high risk categories those ‘systems intended to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts’.<sup>126</sup>

Based on these premises, the following section provides an overview of the common risks associated to the considered technologies across them and due to technical reasons.

### 1. Biases and Due Process Guarantees

The performance of artificial intelligence tools is primarily related to the nature of data employed for the training and the functioning of automated-driven models. In this respect, integrity and quality of the datasets represent the fundamental prerequisites for a well-functioning design of technologies for the legal domain.<sup>127</sup> Biases in training data and proxy discrimination are the two major biases potentially affecting datasets.<sup>128</sup> One common source of biased training data is given by sampling bias. This bias emerges when some strands of the population are misrepresented, because there is not a sufficient representation of the features of these strands of the population in the used datasets. Sampling bias leads to misrepresentation distorting the evidence drawn from the same training data. The bias is in turn incorporated into the statistical model that originates from the training data and propagates into the output, eventually producing misleading results.<sup>129</sup>

Another bias potentially affecting training data is related to what data scientists call ‘historical bias’, resulting from sociological and/or historical misconceptions

<sup>125</sup> *ibid* 65.

<sup>126</sup> European Commission, ‘Annexes’ n 10 above, Annex III, 8 a.

<sup>127</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 9-10.

<sup>128</sup> P. Hacker, ‘Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination under EU Law’ *Common Market Law Review*, 1143, 1148 (2018); G. Comandé, ‘Regulating Algorithms Regulation? First Ethico-Legal Principles, Problems and Opportunities of Algorithms’, in T. Cerquitelli et al eds, *Towards glass-box data mining for Big and Small Data* (New York: Springer International, 2017), 169-207.

<sup>129</sup> J.A. Kroll et al, ‘Accountable Algorithms’ *University of Pennsylvania Law Review*, 633, 680 (2017).

that are reflected into the datasets, likewise skewing their representativeness. Nonetheless, the objective of achieving representativeness of the data, may lead itself to additional biases, for it *per se* forces the designers of the technology to enact stricter surveillance and classification methods needed exactly for the measuring of the targeted data representativeness.<sup>130</sup> This could in turn expose minorities to additional harms.<sup>131</sup> Moreover, an excessive focus of accuracy and quality of employed datasets could have the paradoxical outcome of stimulating developers' extensive – and unlawful under data protection or competition laws – data collection, under the façade of the design of an ethical artificial intelligence tool.<sup>132</sup>

Biases in training data are also likely to generate biases in the subsequent moment of the analytical processing. If the algorithmically calculated risk scores are distributed in an untruthful way among protected groups, then the employed dataset is affected by a bias called 'unequal ground truth'. Such bias causes a 'proxy discrimination', that is a statistical discrimination,<sup>133</sup> given by 'untrue' statistical associations and subsequent scientific inferences.<sup>134</sup>

Also, with reference to those tools assessing the inclinations to decide in a certain way (for judges) or to win and /or move in a case in a certain way (for attorneys), biases, lack of accuracy in the data, quality of the training data, and so on, might result in erroneous or biased predictions leading to both discrimination and harm to all individuals involved, namely the 'evaluated' individual and the end users (eg clients), while producing a stigmatization on the decisionmaker with consequent chilling effect and a serious harm to judicial independency.

The detection of the mentioned biases both in employed datasets and in subsequent processing patterns is impaired by two major obstacles, a technical and a legal one.

The technical impairment relates to the difficulty of designing technologies, and thus also those designed for legal decision purposes, in a manner that renders their functioning transparent, interpretable and thus explainable,<sup>135</sup> mainly due to their adaptive and unpredictable nature.

From the legal standpoint, technologies are often developed by private corporations that are eager to protect their newly developed technologies through intellectual property rights and mainly through trade secrets.<sup>136</sup> The shielding of

<sup>130</sup> J. Powles, 'The Seductive Diversion of 'Solving' Bias in Artificial Intelligence' *Onezero*, available at <https://tinyurl.com/bddazcm7> (last visited 30 June 2022).

<sup>131</sup> *ibid*

<sup>132</sup> C. Kind, n 103 above.

<sup>133</sup> J.A. Kroll et al, n 129 above, 680-681.

<sup>134</sup> Highlighting the risks of these bias in the construction of an autonomous-driven and personalised legal system, H. Eidenmüller and G. Wagner, *Law by Algorithms* (Tübingen: Mohr Siebeck, 2021), 47-79.

<sup>135</sup> See V. Chiao, 'Fairness, Accountability and Transparency: Notes on Algorithmic Decision-Making in Criminal Justice' 15 *International Journal of Law in Context*, 135, 138 (2019).

<sup>136</sup> This is well highlighted by R. Brauneis and E.P. Goodman, 'Algorithmic Transparency for the Smart City' 20 *Yale Journal of Law & Technology*, 103 (2018).

legal algorithms' internal functioning through intellectual property protection upholds businesses' competitive advantage deriving from the investments in the collection and production of information.

Intellectual property tools contribute to obscuring of AI-based technologies' functioning,<sup>137</sup> rendering data collection and processing activities opaque and exclusive.<sup>138</sup> This leads in turn to opaque and exclusive quantification and categorization practices,<sup>139</sup> which come to sustain legally binding decisions.

In this perspective, the growing use of privately-developed algorithms for the purposes of legal decision-making, means greater influence of private corporate power in those same decisional processes. This raises also in the legal sector, the risk of 'private capture' that the literature has generally observed in respect to administrative decision-making.<sup>140</sup> The direct corollary of this capture is the difficulty to externally control the actual functioning of – and thus the existence of – eventual biases within- corporations' AI legal tools. As a result, legal practitioners and citizens making use of these technologies may find it difficult to challenge and object to the decisions determined by those technologies.<sup>141</sup> The already mentioned *Compas* case well illustrates these shortcomings.

As apparent, the risk of biases, matched with the technical and legal hurdles to algorithms' accessibility, brings about substantial concerns regarding due process guarantees.<sup>142</sup> It is important to note that from a formal point of view these AI based tools might not trigger the intervention of legal safeguards, such as the right to not being subject to a solely automated decision-making process (Art 22 GDPR), since technically (as in COMPAS, eg) it is a human being responsible of the final decision with actual power to overrule the AI indication. And yet, the automation bias (the subjection of the human decisionmaker to the suggested super performance of the AI tool) can easily kick in creepingly substituting a potentially biased machine decision to the officially human one.

## 2. Automation Bias and Machine Dependence

<sup>137</sup> F. Pasquale, *The Black Box Society – The Secret Algorithms that Control Money and Information* (Cambridge: Harvard University Press, 2015), 3,11.

<sup>138</sup> D. K. Citron and F. Pasquale, 'The Scored Society: Due Process for Automated Predictions' 89 *Washington Law Review*, 1 (2014).

<sup>139</sup> *ibid* 10, 13.

<sup>140</sup> C. Coglianese and D. Lehr, 'Regulating by Robot: Administrative Decision Making in the Machine Learning Era' 105 *Georgetown Law Journal*, 1147, 1151 (2017), assessing the problem of 'cyberdelegation' by governments to private corporations developing algorithms.

<sup>141</sup> This is underlined by R. Yu and G. Spina Ali, 'What's Inside the Black Box? AI Challenges for Lawyers and Researchers' 19 *Legal Information Management*, 5, 6 (2019), underlining problems of unpredictability of AI-based tools. See also F. Pasquale, 'Secret Algorithms Threaten the Rule of Law' *MIT Technology Review*, available at <https://tinyurl.com/3wuw7s3z> (last visited 30 June 2022).

<sup>142</sup> Stressing this link J. Balkin, 'The Three Laws of Robotics in the Age of Big Data' 78 *Ohio State Law Journal*, 5, 1217, 1239 (2017). Assessing due process guarantees in algorithmic decision-making, D. K. Citron and F. Pasquale, n 138 above, 2; K. Crawford and J. Schultz, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' *Boston College Law Review*, 55, 93 (2014).

The last considerations lead us to a different concern raised by the increasing reliance by legal professionals on automated-driven applications for the performance of their tasks. In this respect, the risks emerge of increasing dependence on AI's evidence by decision-makers, if not of an outright capture of legal decision-making processes by employed AI tools.

In addition to the consequences of the automation bias (not questioning the automated outcome but uncritically endorsing it) on the autonomy and independence with which legal decisions are taken, in the legal domain the automation bias can have a troublesome impact on the quality of legal reasoning itself, increasingly relying on statistical computations and favouring standardization, and ultimately undermining the creative component of legal reasoning, which is the gist of every hermeneutic exercise by jurists.<sup>143</sup> In other words, since the legal systems evolve by differentiating dissimilar cases and by incorporating emerging legal needs and solutions, the excessive reliance on previous patterns can lead to a crystallization of law halting the evolution of legal rules and their ability to adapt to different problems and/or mutations in the social understanding of them.

Further concerns relate to the difficulties of incorporating professional standards into employed technologies. This raises an array of largely unexplored issues concerning the risks concealed in automatizing a 'human science' as the legal science, because of the replacement of legal practitioners by automated/intelligent agents. It is not only a matter of de-humanising the way in which the law is interpreted, applied and enforced but a possible significant limit to the way in which legal systems evolve, by small judicial changes, by challenging the *status quo* with new arguments, by slowly internalizing new social needs.<sup>144</sup>

This concern related to what is defined in technical terms as 'automation bias', is well known in other domains experiencing AI applications, where machine-driven results are a-critically applied without an autonomous evaluation of whether the rendered solution is suitable or correct in respect to the case at stake.<sup>145</sup>

Even when the most sophisticated automated systems develop actual predictive capabilities, which can imagine legal outcomes beyond established legal courses, for structural reasons, they remain bound to the criteria upon which the model

<sup>143</sup> For a theorization of the distinction between 'human' and 'computational' law see G. Zaccaria 'Figure del giudicare: calcolabilità, precedenti, decisione robotica' *Rivista di diritto civile*, 277, 280 (2020), where it is highlighted how uncertainty, in the sense of a dynamic evolutive nature, is a core feature of a human-based legal system in which the human conscience transforms into legal constructs the underlying social developments.

<sup>144</sup> For some philosophical reflections on the distinction between 'human being' and 'autonomous system' see M. Ferraris, 'Anima e automa' *La ricerca*, 12 (2020) where the Author identifies the distinctive feature of the human being in its corporeality and social nature, respectively related to its being attached to a human body and to a complex social context. Change of the social context in which the human acts, stimulates a change in its way of thinking and thus in its decision-making process. For a theorization of the distinction between 'human' and 'computational' law.

<sup>145</sup> G. Comandé, 'Intelligenza artificiale e responsabilità tra «liability» e «accountability»'. Il carattere trasformativo dell'IA e il problema della responsabilità' *Analisi Giuridica dell'Economia*, 169, 188 (2019).

has been built, or that the same model has autonomously generated. This means that also the most sophisticated computational models, are affected by a certain degree of ‘data’ or better said ‘pattern dependency’, which, despite the adaptiveness of the model, risk to harness automatised legal reasoning into an analytical determinism that blurs consistent patterns with the unpredictable factor characterising the human decision making and that risks making the legal system increasingly self-referential and stagnant.

Certainly, also human reasoning suffers from a certain degree of ‘path dependency’, in terms of what is known as the legal interpreter’s ‘pre-understanding’.<sup>146</sup> After all, the myth of the hyper-rationality of decision makers has been disowned in various fields of study, not only theoretically<sup>147</sup> but also empirically.<sup>148</sup> However, despite being inherently influenced by external factors, the human legal reasoning is still characterised by elements of irrationality and discretion, which generates some costs and maybe also some harms to third parties or society as a whole, but which is equally contributing to the evolutive development of the legal system.

Nonetheless, the adaptive nature of the artificial intelligence-based tool may potentially mitigate the threats of excessive path dependence and thus of monolithic AI-driven legal decision-making processes. Also in this case, however, an excessive adaptiveness of employed tools could entail additional risks, first related to the difficulties of identifying above-illustrated biases. Moreover, an inscrutable path-change of the correlations on the basis of which the legal decision-making process is conducted, could impair the fundamental relation between predictability and certainty on which the human-generated legal systems rely on. A fast-changing decision-making machine could indeed render it very difficult for the addressees of the legal verdict to predict the outcome of the decision-making process.<sup>149</sup>

<sup>146</sup> See D. Canale, ‘La precomprensione dell’interprete è arbitraria?’ *Ethics & Politics*, 1 (2006); P.G. Monateri, ‘Sub voce Interpretazione del diritto’ *Digesto Discipline Privatistiche, Sezione Civile* (Torino: UTET, 1993, 4<sup>a</sup> ed), 53.

<sup>147</sup> Interesting in this regard are the theories developed by Nobel prize Herbert Simon, in his work, H. Simon, *Administrative Behaviour: A Study of Decision Making Processes in Administrative Organisations* (New York: Free Press, 1997, 4<sup>th</sup> ed) where the Author distinguishes between routine decisions that are of ‘structured nature’, thus also subjectable to predictive and automated decision-making, and strategic and managerial decisions, where the decision maker exercises his very own ‘creative’ discretion and which can thus hardly be standardized into a machine-readable format. See as well D. Kahneman et al, *Judgment Under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press, 1982).

<sup>148</sup> A.J. Wistrich et al, ‘Heart versus Head: Do Judges Follow The Law or Follow Their Feelings?’ 93(4) *Texas Law Review*, 855 (2015). See also D. Lieberman, ‘Reflective and Reflexive Judgment Processes: A Social Cognitive Neuroscience Approach’, in J.P. Forgas et al eds, *Social Judgments: Implicit and Explicit Processes* (Cambridge: Cambridge University Press, 2003), 44; J.St.B.T. Evans, ‘Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition’ *Annual Law Review of Psychology*, 255, 278 (2008).

<sup>149</sup> G. Zaccaria, n 143 above, 279.

### 3. Translational Bias

The highlighted threat of *over-reliance* on machine-driven outputs needs to be further distinguished from that of *over-use* of the considered technologies, in turn connected to the risk of ‘translational biases’, given by the employment of an artificial intelligence tool outside of the context or scenarios for which it was trained for.<sup>150</sup> The prevention of these biases is particularly important in respect to technologies employed for legal decision-making purposes, whose application in a context that is different from that for which the system was trained could have dramatic consequences. Think, for example, about the distortive, if not harmful, effects of applying a system designed for the purposes of criminal law assessments to predict the probabilities to commit a crime, for example, to the calculation of the damages deriving from a moral distress. Think also about the simple use of analytical tools conceived to assist legal experts (able to doubt machine-driven outputs and critically assess them) and the opening of its use to lay people or less experienced ones. The consequences of the over-use of these applications also beyond their original purpose, are manifold and directly trigger civil liability concerns, as well as the need to enact specific supervisory mechanisms related to the right combination between the specific types of applications and the professionals that come to handle them in a specific field of expertise.

From a yet further perspective, it is worth highlighting how law evolves along deviant lines of reasoning from precedent cases.<sup>151</sup> Such ‘deviating lines’ offer innovative arguments that the IA could disregard as such – since outside the detected usual pattern – or classify as errors impairing the evolving nature of the legal system. Of course, there is a positive side of the coin in letting a machine learning algorithm spotting important variables in the decision-making process. It can unveil the existence of meta-legal decisive factors – not necessarily biased or forbidden ones – thus contributing to the evolution of the rule of law. The very same tool in the hands of domain experts can actually sparkle innovation instead of hampering it.

### 4. Inequality, Discrimination and New Vulnerabilities

The European Commission’s Ethics Guidelines on Trustworthy Artificial Intelligence highlight how equality in data-driven decision-making processes requires that ‘the system’s operations’ do not ‘generate unfairly biased outputs’, this implying particular attention towards ‘vulnerable persons and groups’, which are ‘at risk of exclusion’. As the CEPEJ Ethical Charter suggests,

<sup>150</sup> On automation and translational biases see G. Comandé, n 145 above, 176.

<sup>151</sup> G. Comandé, ‘Legal Comparison and Measures: It’s Logic to Go Beyond Numerical Comparative Law’, in *Studi in onore di Aldo Frignani, Nuovi orizzonti del diritto comparato europeo e transnazionale* (Napoli: Jovene, 2011), 173. See also ECHR, *Greek Catholic parish Lupeni and Others v Romania* [GC] no 76943/11, 29/11/2016, § 116; ECHR, *Z. v Finland* no 22009/93, §§95.

‘public and private stakeholders must ensure that the methods used do not reproduce or aggravate such discrimination and that they do not lead to deterministic analyses or uses’.<sup>152</sup>

In this respect, particular attention needs to be given to those cases in which sensitive data are processed, as data related to

‘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’.<sup>153</sup>

Indeed, when these data are processed by technological infrastructures that are affected by the above recalled biases, the risk of discriminatory legal decisions becomes higher.

The proposed AI regulation is even sharper. While referral 44 requests in general that

‘training, validation and testing data sets should be sufficiently relevant, representative and free of errors and complete in view of the intended purpose of the system’,

referral 40 specifies:

‘in particular, to address the risks of potential biases, errors and opacity, it is appropriate to qualify as high-risk AI systems intended to assist judicial authorities in researching and interpreting facts and the law and in applying the law to a concrete set of facts’.

Overall, the untenable aim of ‘free of errors’ datasets approach boils down in a specific legal rule (that is Art 10.3 AI Act) established only for high-risk AI systems.

As a wide strand of the literature has acknowledged, indeed, when data subjects are classified on the basis of particular features, as health conditions or religious beliefs, the resulting decision-making processes are likely to impact on related fundamental rights, as the right to health or the freedom of speech and religion,<sup>154</sup> for example when access to a specific medical treatment is impaired to a patient that has been erroneously calculated by the machine as being low risk, or when access to a banking service is denied to a citizen that has been identified as high risk. The possibility to be enclosed in pre-determined algorithmic

<sup>152</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 9.

<sup>153</sup> *ibid*

<sup>154</sup> This is well highlighted last by S. Wachter, ‘Affinity Profiling and Discrimination by Association in Online Behavioral Advertising’ 35 *Berkeley Technology Law Journal*, 366 (2020).

patterns is destined to generate stigmatization outcomes,<sup>155</sup> as well as new vulnerabilities,<sup>156</sup> related, for example, to the emotional distress of being evaluated by non-challengeable predictions.<sup>157</sup>

These general shortcomings of automated-driven decision-making mirror also on the legal domain. From the perspective of citizens, the reliance on AI-driven tools in the domain of private legal services, may generate discriminations resulting for example in a biased calculation of contracts risks,<sup>158</sup> or, in the domain of public legal services, in a biased prediction of recidivism rates.<sup>159</sup>

Legal professionals, conversely, risk to be exposed to profiling activities and their legal reasoning may be ‘captured’ by the machine along the lines of an outright machine-dependency. These two aspects will be better explored below.

#### 4. Profiling

A risk entailed in AI-driven tools for legal services concerns the profiling of legal decision makers. This risk is particularly relevant in respect to judges and had already been noted by the Commission in the *Green Paper on public sector information in the information society*.<sup>160</sup> Here, it was noted that the enlargement of judicial databases could lead to the creation of outright individual dossiers on decision-makers.<sup>161</sup> In the artificial intelligence era the fast aggregation of information for the detection of correlations and the drawing of inferences, could create distortions in the way legal datasets are employed, that is not for gaining legal knowledge, but to detect patterns between judgments and some accidental factors, as age, sex, civil status or birthplace of decision makers.<sup>162</sup> This information could possibly be used for the purposes of forum shopping and thus for the identification of the most convenient judicial venue. These risks have been highlighted also by the European Court of Human Rights with reference to the

<sup>155</sup> G. Comandè, n 19 above.

<sup>156</sup> G. Malgieri and J. Niklas, ‘Vulnerable Data Subjects’ *Computer Law and Security*, 4 (2020). From a gender perspective, see F. Luna, ‘Elucidating the Concept of Vulnerability: Layers Not Labels’ 1 *International Journal of Feminist Approaches to Bioethics*, 121, 139 (2009).

<sup>157</sup> D.J. Solove, and D.K. Citron, ‘Risk and Anxiety: A Theory of Data Breach Harms’ 96 *Texas Law Review*, 737 (2018).

<sup>158</sup> S. Murray, ‘Algorithms Tame Ambiguities in Use of Legal Data’ *Financial Times*, available at <https://tinyurl.com/y3n7xn3s> (last visited 30 June 2022).

<sup>159</sup> See M. Zhu, *An Algorithmic Jury: Using Artificial Intelligence to Predict Recidivism*, available at <https://tinyurl.com/2p9ce32v> (last visited 30 June 2022).

<sup>160</sup> Commission of the European Communities, *Green Paper on Public Sector Information in the Information Society*, COM(1998)585 final (20 January 1999).

<sup>161</sup> *ibid* 11.

<sup>162</sup> The relevance of accidental judgments in judicial decision making has been already highlighted by some studies well before the artificial intelligence-driven revolution. See C. Guthrie, J.J. Rachlinsky and A.J. Wistrich, ‘Inside the Judicial Mind’ 86 *Cornell Law Review*, 777 (2001); S. Danziger, J. Levav and L. Avnaim-Pesso, ‘Extraneous Factors in Judicial Decisions’ *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 108, 6889 (2011).

shortcomings of publishing decisions with the identification of the judge.<sup>163</sup>

France has been the first European Member State addressing this concern, by establishing an exception from the general right to use court decisions to analyse future judicial decision-making in the 2019 Judicial Reform Act.<sup>164</sup> Art 33 of the French Act bans to use judges' identities in published court decisions for the purposes of evaluating and predicting their behaviour.<sup>165</sup>

Similar risks of profiling through artificial intelligence-based tools regard however not only judges, but also lawyers and counterparties. As anticipated, this kind of profiling is perceived as extremely worrisome for the rule of law, since it can create significant leverage on counterparts and judges and exposes to high risks of stigmatization.<sup>166</sup> These instruments produce risks for end-users as well since the 'accuracy' of the profiling and its consequent predictions are highly dependent on the quality and quantity of data<sup>167</sup> and shifts the attention from the quality of the arguments sustaining a legal position to non-legal issues normally and officially not affecting a legal decision. In other terms, while meta-legal elements such as policy arguments have a recurrent and acknowledged place in decisions on legal matters other factors need to be and to remain irrelevant to avoid determining justice by spurious factors undermining the rule of law.

These considerations hold true also in those legal systems, as the Italian one, that are based on collective decision-making mechanisms, where no dissenting opinion is envisaged and where there are no qualitative or quantitative indicators that reveal the actual weight of the involved judges' opinion in the shaping of the final judgment thus minimizing risks of judges' stigmatization patterns.

It must also be noted that, if we do not consider or effectively remedy the risks related to biased datasets/outcomes, the very same tools, under proper vigilance, might be highly beneficial if and when 1) they actually identify and flag the (not permitted) relevance of spurious factors, that is the suspect that a monocratic judge could be driven by unpermitted biases; or 2) they flag factual metalegal<sup>168</sup> patterns such as the relevance of factual situations in driving the adjudicating process (eg local costs for replacing goods; existence of services

<sup>163</sup> Eur. Court H.R., *Vernes v France*, no 30183/06; Eur. Court H.R., *Pretto v Italy*, no 7984/77.

<sup>164</sup> Loi no 2019- 222 du 23 mars 2019 n 3 above.

<sup>165</sup> Art 33, *ibid*: 'the identity data of magistrates and of members of the judiciary cannot be used with the purpose of the effect of evaluating, analysing, comparing predicting their actual or alleged professional practices'.

<sup>166</sup> This is one of the reasons why Art 33 of French law no 2019/222 prohibits the re-use of the identification data of judges and employees for classification, evaluation, comparison or profiling purposes and accompanies this prohibition with a criminal sanction.

<sup>167</sup> In particular, with reference to the dangers related to the identification of the judge in charge, see Eur. Court H.R., *Vernes v France*, no 30183/06. On the risks related to the publication of the decision see *Pretto v Italy*, no 7984/77; while with reference to the necessary impartiality of the judge and the right to a judge established by law (and not only to an alternative resolution mechanism) see *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.

<sup>168</sup> G. Comandé, n 151 above, 173.

available, etc).

For these reasons, flagging as high risk the use only by the judiciary of these tools results in more than a misleading understanding of the risks they entail and the potential loss of their low-risk usefulness. In this light, the Council's proposal to extend the realm of unacceptable AI systems, forbidden under the AI Act, to private social scoring models does not appear to be an efficient solution, since it would ban from the market profiling methods that could have great benefits, in terms of the speed and the quantity of decision-making, also in the legal sector.

A preferable solution would be that of including private social scoring models – not only those models used to profile judges – within the category of high risk systems under the proposed AI Act, and modelling the applicable requirements in consistency with a risk-based approach.

### 5. The 'Digital Legal Divide' and Competition Concerns

The importance of a certain degree of knowledge in the field of informatics and data science by legal professionals is increasingly acknowledged.<sup>169</sup> This change in competences needed for handling technologies employed for the resolution of controversies, could engender risky disparities between legal practitioners and their assisted parties, and especially between more powerful parties as businesses or institutions with stronger technological means, and parties with less technological facilities and thus with less understanding of rendered results. Weaker parties and weaker law firms (meaning also less financially equipped to cope with the costs) would thus become those who are less empowered on the side of technical expertise, this making it hard to litigate 'on equal munitions' with more powerful technologically endowed experts. Implications in constitutional terms to the right to a fair trial and equality of arms are obvious in many jurisdictions.<sup>170</sup>

The growing relevance of artificial intelligence-based tools in the patterns of legal interpretation, implementation and reform, is thus likely to generate new 'digital' legal divides, affecting citizens' accessibility to legal expertise and services in new ways as well as legal practitioners' abilities to cope with technological advancements. From the first standpoint, these divides are soon destined to result in social imbalances deriving from inequalities in the legal protection available to citizens. From the opposite angle, that of legal services providers, the different resources available to public administrations could create technological gaps in less wealthy geographical areas.<sup>171</sup>

<sup>169</sup> K.D. Ashley, *Artificial Intelligence and Law: New Tools for Law Practice in the Digital Age* (Cambridge: Cambridge University Press, 2017) 4, 11; J. Armour, R. Parnham and M. Sako, n 44 above, 57.

<sup>170</sup> For instance in Italy it would trigger doubts under Art 24 of the Constitution and under Arts 6 and 13 ECFR.

<sup>171</sup> See the evidence in respect to 'digital government services', L. Reggi and J.R. Gil - Garcia, 'Addressing Territorial Digital Divides Through ICT Strategies: Are Investment Decisions Consistent

Additional disparities are likely to be experienced in the private sector, where market imbalances may arise as a direct result of the different availability of the highly sophisticated datasets needed to develop and enacting artificial intelligence applications<sup>172</sup> or even the domain interdisciplinary knowledge to master them. In consideration of the very specificities of those datasets, the market of AI-based applications for the legal sector is at risk of being characterised by high barriers to entries as well as by foreclosure behaviours, giving rise to new competition scenarios in the legal services' sector.

## VII. Mapping Solutions: The Legal Framework for AI-Assisted Decision-Making Tools

The variety of concerns described so far triggers a mixture of possible lines of intervention for both interested businesses or institutions engaging in the development of AI-driven legal technologies, and regulators. For the purposes of achieving a more comprehensive and balanced understanding of these challenging transformations, a third focus of enquiry, in addition to the ones given by the market and ethical perspectives, is given by the analysis of the legal system. This is suggested by the same High-Level Expert Group on Artificial Intelligence, which identifies among the three components of the 'Trustworthy AI', not only i) ethics and ii) technical (and thus market) robustness, but also iii) lawfulness. While the first two components have been given more attention respectively from a theoretical and practical perspective, the definition of the lawfulness component has been – strangely – lagging behind.<sup>173</sup>

The identification of the applicable legal framework is, first of all, relevant for circumscribing the realm of *legitimate* employment of machine learning technologies for legal decision-making purposes. In accordance with the very essence of legal rules, the legal perspective comes along with enforcement mechanisms assuring compliance by economic players. Enforceability of legal requirements assures the existence of quick reaction means regarding technologies that come to violate established legal provisions: this is what occurred in the UK where an algorithm employed by the police for facial recognition purposes, was challenged by a civil rights group and ruled unlawful.<sup>174</sup> A correct implementation of identified legal

with Local Needs?' *Government Information Quarterly*, 101562 (2020).

<sup>172</sup> D. Rubinfeld and M. Gal, 'Access Barriers to Big Data' 59 *Arizona Law Review*, 340, 381 (2017); I. Graef, 'Rethinking the Essential Facilities Doctrine in Digital Markets' 53 *Revue juridique Thémis de l'Université de Montréal*, 33, 72 (2019).

<sup>173</sup> This point is made by G. Comandè and D. Amram, 'Feedback for the EU Commission Inception Impact Assessment Towards a 'Proposal for a Regulation of the European Parliament and the Council Laying Down Requirements for Artificial Intelligence'', available at <https://tinyurl.com/5enbarb7> (10 September 2020) (last visited 30 June 2022).

<sup>174</sup> J. Rees, 'Facial Recognition Use by South Wales Police Ruled Unlawful' *BBC*, available at <https://tinyurl.com/2p9demt6> (11 August 2020).

requirements, as backed-up by ready enforcement, may also contribute to guarantee the employment ‘for good’ of legal technologies and thus define, on the side of their theoretical justifications, the sphere of acceptable uses that legitimises them.<sup>175</sup> In other words, prompt enforcement reactions might trigger a sort of spill-over effect that is ultimately useful to (re)define the list of high-risk AI systems in legal services and administration of justice and to unleash their faster uptake.

To begin with, the first issue to be addressed regards what part of the legal framework is relevant for the development of ‘trustworthy’ legal artificial intelligence-based technologies. A strand of the literature has pointed to the human rights’ framework.<sup>176</sup> In our reading, however, the reference to fundamental rights, although necessary, is misleading if not coupled with operational rules. Indeed, as has been highlighted in the previous paragraphs, the same *Guidelines on a Trustworthy AI* appear to consider the protection of fundamental rights as the *end-goal* of an ethically-sound AI design.<sup>177</sup> Moreover, the same fact that fundamental rights are considered as ‘basic moral entitlements’<sup>178</sup> places these rights at a level of normativity, which, although already attaining to the domain of enforceability,<sup>179</sup> is nonetheless very close to that of ethics. In this respect, it appears that the fundamental rights’ perspective offers little practical guidance to effectively bridge ethics in the market without a clear-cut link with clearly legally enforceable rules.

In light of these shortcomings, in our view, the gap needs to be filled looking at the market that has to internalise ethical precepts and thus at the legal- *per se* enforceable- rules that come to regulate the market of AI-driven applications for legal services. In other words, the applicable legal framework for AI based legal services (in any domain for what matters here) needs to be a layered one moving downwards from general principles (fundamental rights protection) to general non sectoral rules (as will be referred to in the paragraphs below), and hence to sectoral rules (as occurs with the reference to the European Union Harmonisation Legislation in Annex 2 of the AI act).

With reference to general non sectoral rules, in addition to general private law rules (including liability ones),<sup>180</sup> by establishing a basic set of rules to which the fast-developing market of AI tools for legal decision-making needs to comply with, regulations provided by open data laws, data protection laws as well as the proposed Regulation on artificial intelligence itself provide concrete opportunities

<sup>175</sup> See L. Floridi et al, n 94 above.

<sup>176</sup> K. Yeung, A. Howes and G. Pogrebna, n 21 above.

<sup>177</sup> G. Comandè and D. Amram, n 173 above, 6, where the Authors make reference to ‘an overall approach grounded on fundamental right protection’ of the Guidelines on a Trustworthy AI.

<sup>178</sup> This was already acknowledged by R. Dworkin, *Taking Rights Seriously* (London: Duckworth, 1977).

<sup>179</sup> Stressing this point, K. Yeung, A. Howes and G. Pogrebna, n 21 above, 5.

<sup>180</sup> On the problems of defining an accurate liability framework for AI systems, see generally, M. Rabitti, ‘Intelligenza artificiale e finanza. La responsabilità civile tra rischio e colpa’ *Rivista trimestrale di diritto dell’economia*, 295, 319 (2021).

for bridging the market reality with the ethical declarations, with that offering to market players a more clear guidance in the uptake of relevant technologies.

Under these premises, in the effort to define the relevant legal framework, from a methodological standpoint, a complementary approach is employed: this approach moves from the assumption that ethical principles themselves can offer interpretative guidance for an effective interpretation of market-based legal precepts.<sup>181</sup> In this perspective, the illustrated ethical principles can act as relevant interpretative criteria for detecting those provisions of the regulatory framework that are most relevant for an ethically-sound market development. From a further perspective, these same principles could also stir the evolutive interpretation of existing legal rules and encourage, where necessary, legal reforms,<sup>182</sup> which come to vitalise relevant ethical principles within the applicable legal and enforceable framework<sup>183</sup> and directly address emerging social and market imbalances.<sup>184</sup>

### 1. Open Data Regulations

The enactment of open data is likely to minimise the risks of inadequate cross-reference and to strengthen the precision of the results of automated processing, in direct consistency with accuracy objectives.<sup>185</sup> Moreover, explicability of AI-driven systems, which has been considered by the High Level Expert Group on AI as a crucial element for ‘building and maintaining users’ trust in AI systems’, is exactly given by the open communication of their design, capabilities, and purposes. Explicability also involves the accessibility of the decision making process’ model, that is of its organisational structure and the degree in which artificial intelligence tools are integrated with human decision making.

In this respect, the open data policies that are being increasingly considered at EU level, within the former Digital Single Market Strategy and the current EU strategy for data, are of particular importance also for the purposes of AI-driven legal services. In the new EU strategy for data, the Commission has stated the need to inform future regulatory and policy actions regarding data, upon the principle of ‘as open as possible, as closed as necessary’,<sup>186</sup> which promotes data

<sup>181</sup> See already, G. Comandé, n 19 above; Id, ‘Multilayered (Accountable) Liability for Artificial Intelligence’ in S. Lohsse, R. Schulze and D. Staudenmayer, *Liability for Artificial Intelligence and the Internet of Things* (London: Bloomsbury Professional, 2019), 165, 187.

<sup>182</sup> See S. Delacroix and B. Wagner, ‘Constructing A Mutually Supportive Framework Between Ethics and Regulation’ 40 *Computer Law & Security Review*, 10520 (2021), calling for the ‘cross-fertilisation’ between ethics and law. Stressing this point also G. Comandé, n 19 above.

<sup>183</sup> See in these regards the proposal by G. Comandé and D. Amram, n 173 above.

<sup>184</sup> A legislative proposal by the European Commission should be due in early 2021 with the explicit purpose of making artificial intelligence ‘ethical, safe and innovative’. It has been announced by European Parliament, ‘Making Artificial Intelligence Ethical, Safe and Innovative’, available at <https://tinyurl.com/y3szuxjk> (1 October 2020) (last visited 30 June 2022).

<sup>185</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 61.

<sup>186</sup> European Commission, ‘A European Strategy for Data’ n 15 above, 15.

re-usability and analysis across different sectors of the economy, through a constant balancing with intellectual property rights.

Direct reflections of these policy statements at European regulatory level are given in particular by the Open Data Directive,<sup>187</sup> the forthcoming Data Governance Act<sup>188</sup> as well as the recently proposed Data Act.<sup>189</sup> These regulatory frameworks lay down rules ensuring data transferability and re-usage of public sector information for the purpose of stimulating innovation in emerging data-driven markets.

The resulting framework is particularly relevant for the development of AI tools destined to the legal sector. Here, indeed, the design and development of such technologies is mostly fuelled by judicial or otherwise publicly held data. As a result, the emerging businesses are often in private-public partnerships with relevant institutions.<sup>190</sup>

Accordingly, the Open Data Directive requires Member States to enact specific access regimes regarding publicly held data,<sup>191</sup> including *judicial data*.<sup>192</sup> In particular, recital 8 highlights how

‘documents produced by public sector bodies of the executive, legislature or judiciary constitute a vast, diverse and valuable pool that can benefit society. *Providing that information, which includes dynamic data, in a commonly used electronic format allows citizens and legal entities to find new ways to use them and create new, innovative products and services*’.<sup>193</sup>

The recital further recalls the full support provided by European Union funding programs for the analysis of available aggregated and combined datasets and the creation of new services and applications, ultimately stirring technological

<sup>187</sup> Directive EU 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the Re-use of Public Sector Information (Open Data Directive), 26 June 2019, *OJ L 172 (26 June 2019)* 56, 83.

<sup>188</sup> European Commission, ‘Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act)’, 25 November 2020, COM(2020) 767 final.

<sup>189</sup> European Commission, ‘Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)’, COM(2022) 68 final (23 February 2020).

<sup>190</sup> This is the case of the Hart algorithm, developed in partnership with Cambridge University, with data rendered available by Durham police. Council of Europe-European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 51. The ‘Partnership for Open Government’ (OGP) has been established exactly for the purposes of incentivizing the enactment of technology projects based on publicly held data. It is an organization including more than seventy Member States, representatives of civil society and digital companies. *ibid* 19.

<sup>191</sup> See Art 10 Open Data Directive.

<sup>192</sup> See recital 43 Open Data Directive, referring to the availability of documents regarding the ‘legal and administrative process’.

<sup>193</sup> Emphasis added. In this perspective, the recital recalls the full support provided by European Union funding programs for the digitisation of public services through the analysis of available datasets.

evolution.<sup>194</sup>

In light of the above, it must be observed how there are at least two main hurdles to a fruitful enactment of the considered regime. The first one regards the need of a national implementation by Member States, which could lead to a highly fragmented European landscape regarding access to public data, including data regarding judicial decisions. This could possibly obstruct or at least slow down the development of the technologies based on such data if national markets as such are deemed insufficient to promote innovation.

The second one concerns the burden placed onto businesses to anonymise their datasets so as to bypass the application of the stricter regulatory regime regarding personal data. The Open Data Directive indeed admittedly does not interfere with the General Data Protection Regulation.<sup>195</sup> In this respect, recital 52 of the Open Data Directive identifies anonymization as a means for ‘reconciling the interests in making public sector information as re-usable as possible (...)’. The same recital nonetheless acknowledges the costs of anonymisation interventions, to be considered as ‘part of the marginal cost of dissemination’ of relevant information.

These costs could be soon cut down in case of a successful development of the above-mentioned artificial intelligence-based technologies for the anonymisation of court decisions.<sup>196</sup> In this respect, the Commission’s *e-Justice* Action Plan, envisages the training of an automated-driven model for the anonymisation of Court decisions and, as a result, the drafting of best practices as well as of technical guidelines,<sup>197</sup> enabling a more secure and more widely spread use of these applications. This would enable a (possibly) greater reliance on the open data framework and, as a result, a more voluminous sharing of the legal datasets needed for the development of technologies relevant to the field.

Greater accessibility of relevant legal datasets could help creating a level playing field for firms and institutions, thus reducing emerging gaps in market power directly related to an unequal collection of valuable data by producers of AI tools for legal decision-making. To these ends, the proposed Data Governance Act is meant to provide additional rules for the sharing of publicly held data, when these are ‘protected’ by the ‘rights of others’, as data protection rights or intellectual property rights.<sup>198</sup> In this respect, it establishes a prohibition of exclusive arrangements regarding public data.<sup>199</sup>

If enacted, this provision could be extremely important to enable access to public data also to medium and smaller enterprises and thus to keep up with competition dynamics in the emerging market of AI-driven legal services. It

<sup>194</sup> See also recital 10 GDPR.

<sup>195</sup> Recital 52 GDPR.

<sup>196</sup> European Commission, ‘2019-2023 Action Plan European e-Justice’ n 26 above, 15.

<sup>197</sup> *ibid*

<sup>198</sup> See Chapter III Data Governance Act.

<sup>199</sup> So Art 4 Data Governance Act.

would indeed stem the creation of market imbalances given by exclusive collection of relevant data by one powerful party and by the creation of preferential channels of data retrieval by businesses from public administrations and agencies. In this light, the prohibition of exclusive arrangements regarding publicly held data would thus ultimately contribute to the creation of a pro-competitive market setting for the blossoming of research and development endeavours in the sector of AI-assisted legal decision-making tools.

From a further perspective, the proposed Data Governance Act establishes a sophisticated data sharing mechanism for ‘data altruism’ purposes, related to the purposes of general interest, such as scientific research purposes or the *improvement of public services*.<sup>200</sup> This sharing mechanism relies on the intermediation of a ‘data altruistic organization’, which are not-for-profit legal entities recognised in the Union<sup>201</sup> and which share data ‘without seeking a reward’.<sup>202</sup> Defined in these terms, the sharing infrastructure for data altruism purposes could be particularly interesting for the collection of data for the development of AI tools for public legal services, either directly by governments or by established public-private partnerships.

## 2. The Data Protection Framework

The General Data Protection Regulation lays down the fundamental requirements and principles for the design of technologies processing personal data. It sets the basic framework, to which such a fast-developing market needs to adhere. In this perspective, the normative grounds laid down by the GDPR are particularly important for information retrieval and the building of cognitive computational models destined to the legal sector, which strongly rely on personal data. The GDPR sets the prerequisite for the free circulation of relevant personal datasets and thus for the achievement of relevant technologies’ interoperability, both of which lie at the very core of the development of the correspondent market.<sup>203</sup> In this perspective, the General Data Protection Regulation operates as both a *facilitator* and *external limit* to the development of intelligent agents in the market of legal services.

By external limit, we mean a set of constraints aimed at moulding the required characteristics of these AI tools to facilitate their uptake, not merely limiting or banning legal factors. This ‘modelling’ function of the General Data Protection Regulation over technologies’ design is directly rooted in its fundamental rights’

<sup>200</sup> Art 2(10) Data Governance Act. Emphasis added.

<sup>201</sup> See recital 36 Data Governance Act.

<sup>202</sup> Art 2(10) Data Governance Act.

<sup>203</sup> Access to information and interoperability are identified by the Commission as the two key factors to be addressed for the purposes of the development of e-justice applications. This is highlighted at the outset of the e-Justice Plan. European Commission, ‘2019-2023 Action Plan European e-Justice’ n 26 above, 9, 10.

foundations,<sup>204</sup> where the fundamental right to data protection, and the apparatus of substantial and procedural rules it establishes, act as enablers for the protection of other fundamental rights. In the field of digitised legal services, the technical implementation of the data protection framework will heighten the protection of fundamental rights of equal access to justice, to a fair trial and, more in general, of equality before the law clearing out many of the concerns already raised by the recalled ethical documents.

For instance, it helps preventing unacceptable profiling of judges and legal experts without criminalizing the use of such data as occurred in France. It guarantees the possibility of sharing costs for pseudonimizing/anonymizing judicial datasets, along with the possibilities offered by the DGA, among businesses who can afterwards compete for their efficient use.

Along these lines, the GDPR takes expressly into consideration personal data sharing objectives, as declared in the recitals 2, 5, 13 GDPR and as directly substantiated in the special data protection regime granted for processing activities carried out for statistical and research purposes, regarding special categories of data, under the combined reading of Arts 9(2) lett. j); 5(1) lett. b); 6(4); and 89 GDPR.<sup>205</sup> An interpretation of the GDPR as a research-friendly data protection framework in accordance with the mentioned recitals,<sup>206</sup> sustains the facilitating role of the General Data Protection Regulation in the consolidation of a new market as the one related to artificial intelligence-based tools for the provision of legal services.

Read through the lenses of its free flow of personal information objectives, EU data protection law as framed in the GDPR appears to uphold the set European policies of open data, integrating the relevant framework. The recalled Open Data Directive directly acknowledges under recital 52 that ‘the re-use of personal data is permissible only if the principle of purpose limitation as set out in point (b) of Art 5(1) and Art 6 of Regulation (EU) 2016/679 is met’. By stating this, the Directive implicitly recognizes data protection law as a fragment, specifically related to personal data, of the broader European policy regarding data sharing and accessibility, ultimately expressed in the European Strategy for Data.<sup>207</sup>

<sup>204</sup> See recital 1 GDPR. See G. Comandè, n 19 above; G. Comandè and G. Schneider, ‘Differential Data Protection Regimes in Data-driven Research: Why the GDPR is More Research-friendly Than You Think’ *German Law Journal* (2022); D. Amram, ‘Building up the ‘Accountable Ulysses’ Model. The Impact of GDPR and National Implementations, Ethics, and Health-data Research: Comparative Remarks’ 37 *Computer Law & Security Review*, 105413 (2020); D. Amram, ‘The Role of the GDPR in Designing the European Strategy on Artificial Intelligence: Law-Making Potentialities of a Recurrent Synechdoche’ *Opinio Iuris in Comparatione*, 1, 7 (2020).

<sup>205</sup> G. Schneider, ‘Health Data Pools under European Policy and Data Protection Law: Research as a New Efficiency Defense?’ *JIPITEC* 49, para 1(2020).

<sup>206</sup> See G. Comandè and G. Schneider, ‘Can the GDPR Make Data Flow for Research Easier? Yes it Can, By Differentiating!’ *Computer Law & Security Review* 41, 105539 (2021).

<sup>207</sup> European Commission, ‘A European Strategy for Data’ n 15 above, 4. Arguing for an integrated consideration of the EU open data policies and data protection law, I. Graef, R. Gellert and M. Husovec, ‘Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of

Against this backdrop, specific GDPR provisions offer further evidence of their service to these sharing goals.

### **a) Legal Bases Under Arts 6 and 9 GDPR**

An eligible legal basis for the processing of personal data to design of AI-based legal services is offered by the legitimate interest of (private) developers enshrined in article 6(1)(f) GDPR. This basis could however be relied on only by legal professionals, as law firms, for the in-house development of technologies that directly support the conduction of their legal research regarding their clients. In these cases, indeed, a data subject could reasonably expect 'that processing for that purpose may take place', in light of the existence of a

'relevant and appropriate relationship between the data subject and the controller in situations such as where the data subject is a client or in the service of the controller'.<sup>208</sup>

Conversely, the ground regarding the operator's legitimate interest does not seem to be suitable for businesses developing artificial intelligence applications for legal decision-making purposes in respect to which no direct relationship between the business and the client is to be found. Moreover, the same lawful basis cannot be relied on by public authorities. These ones could find an appropriate legal ground for the development of artificial intelligence tools for legal decision-making in the 'performance of a task carried out in the public interest or in the exercise of official authority vested in the controller', in accordance with Art 6(1) lett e) GDPR. Nevertheless, their initiatives might be stifled by the proposed AI Act as we discussed before, at least for those AI system that would directly assist in 'a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts'.

Moreover, the processing of personal data for the development of algorithmic models could be generally linked to research and statistical objectives.<sup>209</sup> When the processing of personal data for these objectives is adequately grounded in one of the mentioned lawful bases under Art 6 GDPR, a special data protection regime is applicable, facilitating data controllers' processing through the possibility to 'derogate' to the principle of purpose limitation under Article 5(1)(b) GDPR and storage limitation under Article 5(1)(e) GDPR and to data subjects' rights as the right to be forgotten under Art 17(3) GDPR and the right to be informed under Art 14(5) GDPR.

Non-personal Data is Counterproductive to Data Innovation' TILEC Discussion Paper, DP 2018-028, available at <https://tinyurl.com/2p9d8krm> (September 2018) (last visited 30 June 2022).

<sup>208</sup> See recital 47 GDPR.

<sup>209</sup> Matching the purposes of research and statistical enquiry under the GDPR and the development of algorithmic models, S. Wachter and B. Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection in the Age of AI' 2 *Columbia Business Law Review*, 494 (2019).

However, to counterbalance these derogations, Art 89(1) GDPR requires data controllers to implement safeguards, assuring the respect of fundamental data protection principles, as the principle of data minimization, and providing ‘appropriate’ technical and organisational measures for the protection of data subjects’ rights and freedoms. These safeguards encompass, first of all, the pseudonymisation of the data. The appropriateness of these safeguards will have to be considered in light of the specificities of the technology to be developed. The technical and organizational requirements mandatory by data protection by design and default rules under Art 24 GDPR provide a fundamental benchmark to these ends.

### **b) Data Segregation Under Art 11 GDPR**

A positive combination of the open data approach under the framework of the GDPR might unfold around an evolutive notion of anonymity, as long as pursuant to the combined reading of the notions described under Art 4 and recital 26 GDPR. In a layered evolution of the available tools, if and when a strong automated pseudonimization can be offered on judicial data, further use of judiciary (open) data could be envisaged also with reference to the underexplored possibilities offered by Art 11 GDPR.

A careful reading of Art 11 GDPR suggests that it provides a pattern to navigate between technical needs to work with personal data and the obvious concerns to their uses/abuses in the legal services domain. In an admittedly cryptical way, Art 11 GDPR suggests a lighter regime for those data controllers whose data processing does not require or does no longer require ‘the identification of a data subject by the controller’. Such data controllers should not be obliged ‘to maintain, acquire or process additional information in order to identify the data subject for the sole purpose of complying with this Regulation’.<sup>210</sup> The provision thus admits the possibility for data controllers to segregate judicial data by way of a strong level of privacy-preserving pseudonimization, freeing them from many compliance burdens.

Yet, such a possibility might be at odds with the relative notion of data anonymity. Recital 26 clearly declares that ‘to determine whether a natural person is identifiable (and his/her data *data are thus not anonymous*), account should be taken of all the means reasonably likely to be used, such as singling out, *either by the controller or by another person* to identify the natural person directly or

<sup>210</sup> See Art 11 GDPR: ‘If the purposes for which a controller processes personal data do not or do no longer require the identification of a data subject by the controller, the controller shall not be obliged to maintain, acquire or process additional information in order to identify the data subject for the sole purpose of complying with this Regulation. 2. Where, in cases referred to in paragraph 1 of this Article, the controller is able to demonstrate that it is not in a position to identify the data subject, the controller shall inform the data subject accordingly, if possible. In such cases, Articles 15 to 20 shall not apply except where the data subject, for the purpose of exercising his or her rights under those articles, provides additional information enabling his or her identification’.

indirectly' (emphasis added). Reference is made expressly not only to the means of data controllers but also of 'another person'. Thus, the mere technical prevention of identifiability by the data controller does not amount to anonymity automatically.

Yet again,

'to ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments'.

While a first literal reading of the latter quoted text seems to refer to abstract availability of technologies and resources, once we concentrate on the purpose of this availability ('to ascertain whether means are reasonably likely to be used to identify the natural person') we realize that Art 11 casts a sort of relative presumption of non-identifiability on data controllers who 'do not need or no longer need' to identify the data subjects whose data they are processing.

Considering that Art 6, para 4, lett e names 'the existence of appropriate safeguards, which may include encryption or pseudonymisation' as an enabler of further processing, we can easily identify a specific case of relatively (limited to specific data processing activities and data processors) anonymous data that are under the special regime of Art 11 GDPR. This can clearly be the case of strongly pseudonymized judicial data open to further public use. Such a reading, still in need of further exploration, would create a fair data pool potentially enabling the competitive development of new AI-based services balancing their use according to the used safeguards. Appropriate safeguards would need to rely also on supervisory authorities' auditing powers. The Data Altruism Organisations under the proposed Data Governance Act, could be directly entrusted with these oversight tasks.

### **c) Human-Centric Technology Design Under Art 22 GDPR**

Art 22 GDPR prohibits the issuing of 'decisions based solely on automated processing, including profiling, which produces legal effects concerning' the addressee of the decision.<sup>211</sup> The provision first addresses the issue of profiling activities, highlighted as a significant risk related to the collection and processing of judicial data.

The prohibition nonetheless appears to have broader implications for the design of technologies employed for automated legal decision-making. It seems to require the developer and the controller of related technologies to establish a

<sup>211</sup> For a more detailed comment on Art 22 GDPR, see S. Wachter and B. Mittelstadt, n 209 above, 494; M. Brkan, 'Do Algorithms Rule the World? Algorithmic Decision-Making in the Framework of the GDPR and Beyond' 27 *International Journal of Law and Information Technology*, 2, 91 (2019).

delimitation between the human and the automation sphere, leaving room for the conduction of self-standing human choices and requiring human oversight of employed technologies.<sup>212</sup>

Moreover, by requiring a human-centric approach, the provision also suggests the need of adopting a scaled approach in the allocation of the tasks between humans and artificial intelligence tools. This scaling could be determined in light of the results drawn from the data protection impact assessments, which data controllers are required to conduct under the conditions provided for by Art 35 GDPR.<sup>213</sup> The mapping of the risks ‘to the rights and freedoms of natural persons’<sup>214</sup> related to the employment of developed technologies, could indeed suggest the opportunity of envisaging a more invasive ‘user control’ when the risks are higher, and conversely of allowing a less rigorous human control, when the risks are deemed lower. In this respect, it would be desirable to define the thresholds regarding the different degree of human supervision required in respect to differently risk-rated artificial intelligence tools for legal decision-making.

For these purposes, a valuable starting point is given by the risk-based approach offered in the proposal for a Regulation on AI, which tailors its rules to the ‘intensity and scope of the risks that AI systems can generate’.<sup>215</sup> Accordingly, it prohibits specific types of artificial intelligence applications under Art 5 and lays down specific rules for high-risk tools identified under Art 6. For this latter category of AI systems, the proposal requires the implementation of a risk management and a data governance system<sup>216</sup> as well as transparency safeguards,<sup>217</sup> which come to support the human oversight requirement under Art 14, requiring the effective oversight ‘by natural persons during the period in which the AI system is in use’. The human oversight requirement is functional to the minimisation of risks to

‘fundamental rights that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse’.<sup>218</sup>

Overall, Art 22 GDPR, as read in combination with Art 14 of the proposed Regulation on AI, establishes a mandatory minimum autonomy requirement,

<sup>212</sup> Stressing this point, G. Comandé and G. Malgieri, ‘Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation’ 7(4) *International Data Privacy Law*, 243, 265 (2017).

<sup>213</sup> M. Kaminsky and G. Malgieri, ‘Algorithmic Impact Assessments under the GDPR: Producing Multi-layered Explanations’ *International Data Privacy Law*, 1 (2020).

<sup>214</sup> So Art 35 GDPR.

<sup>215</sup> Cf Recital 14 of the proposal for a Regulation, affirming the importance of a ‘defined risk-based approach’ for AI. European Commission, ‘Proposal’ n 9 above.

<sup>216</sup> *ibid* Art 9 and Art 11.

<sup>217</sup> *ibid* Arts 12 and 13.

<sup>218</sup> *ibid*, Art 14, comma 2.

demanding the supervision by human subjects regarding the functioning of employed technologies: this supervision relates to the ability of reviewing issued decisions and with that the data that sustain a specific outcome.

For these purposes, legal professionals using artificial intelligence tools for their decision-making will need to be equipped with adequate competences, so as to ward off the perils of the automation bias.<sup>219</sup> Interestingly, in this respect, the considered provisions under the GDPR and the proposed Regulation on AI appear to provide legal grounds to the CEPEJ's principle of 'under user control', which recommends the implementation of 'computer literacy programmes' for users of technologies employed in judicial systems.<sup>220</sup> As some scholars have interestingly observed,<sup>221</sup> however, this change of competence by legal professionals, directly providing AI-driven legal services or using third parties' developed legal technologies, does not only imply a mere acquisition of basic informatics knowledge by them. It triggers instead a much more complex process of combining different skills, involving data science competences for the interpretation of machine-rendered results; traditional legal expertise for the evaluation of the implications of the same results, and ultimately client skills, needed to convey the performed analysis to the party concretely bearing the effects of the decision.

#### **d) Supervisory Authorities' Oversight and Auditing Powers Under Arts 30; 35; 36(1) and 58(1)(b) GDPR**

The relevance of supervisory authorities' oversight and auditing powers in respect to the design of automated legal services is a further aspect to be considered. The information over the nature of employed datasets and the structural features of processing technologies that data protection authorities are entitled to access under Arts 30; 35; 36(1) and 58(1)(b) GDPR could indeed be extremely precious exactly for the conduction of a sound and effective supervision over the technologies eventually employed in the sector of legal services.

In accordance with Art 36(1) GDPR, data controllers have the obligation to consult the supervisory authority prior to the processing, when the data protection impact assessment shows that the processing would result in a high risk, in the absence of measures to mitigate the risk. For the purposes of this prior consultation, the controller shall provide the supervisory authority with the data protection impact assessment performed under Art 35 GDPR, together with any other information requested by the same supervisory authority, as the one contained in businesses' records of processing activities required by Art 30 GDPR.

Moreover, Art 58(1)(e) GDPR establishes data protection authorities' power to carry out investigations in the form of data protection audits, enabling the

<sup>219</sup> K.D. Ashley, n 169 above, 5.

<sup>220</sup> See Art 5, Council of Europe-European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter' n 7 above, 12.

<sup>221</sup> J. Armour, R. Parnham and M. Sako, n 44 above, 27.

same data protection authorities to access ‘all personal data’ and ‘all information necessary for the performance of its tasks’.<sup>222</sup> Through their investigative powers, data protection authorities can enquire which data are the objects of processing activities, and how exactly they are technically processed. Accordingly, they can access the most detailed information regarding the content and structure of employed datasets<sup>223</sup> (eg whether pseudonymisation techniques have been adequately applied), monitoring the actual compliance with the safeguards needed to enjoy the ‘facilitated regime’ under art 11 GDPR.

Also from the private side, the accountability precepts enshrined in Art 5 GDPR are leading some research institutes to develop automated-driven auditing tools for the assessment of businesses’ algorithms.<sup>224</sup> These tools could on the one hand support businesses developing artificial intelligence tools for the legal sector in the performance of their compliance tasks,<sup>225</sup> and on the other assist also supervisory authorities in their auditing powers.

### **3. Standards and Certifications Under the Proposed Regulation on Artificial Intelligence**

The CEPEJ Ethical Charter recommends the enactment of corrective measures that limit or neutralize the risk of harmful effects stemming from the existence of biases in datasets employed for AI-based applications destined to the legal sector.<sup>226</sup> In addition to this, it welcomes initiatives that raise awareness about the presence of biases in datasets among interested stakeholders.

These issues have been partly addressed in the European Commission’s proposed Regulation on artificial intelligence,<sup>227</sup> that has established mandatory requirements on training data, record-keeping about datasets and algorithms, transparency and quality management for high-risk artificial intelligence tools. The proposal indeed sets ‘data and data governance’ requirements regarding the ‘training, validation and testing data sets’, and assuring that in all the mentioned phases data are ‘relevant, representative, free of errors and complete’,<sup>228</sup> with particular attention to ‘the specific geographical, behavioural or functional setting’.<sup>229</sup> In this perspective, data governance practices need to assure that AI-developed tools are bias-free. The relevance to the specific setting in which the AI

<sup>222</sup> Art 58(1)(e) GDPR.

<sup>223</sup> A. Casey, A. Farhangi and R. Vogl, ‘Rethinking Explainable Machines: The gdpr’s ‘Right to Explanation’ Debate and the Rise of Algorithmic Audits in Enterprise’ 34 *Berkeley Technology Law Journal*, 145, 188 (2019).

<sup>224</sup> Ada Lovelace Institute, ‘Examining the Black Box: Tools for Assessing Algorithmic Systems’, available at <https://tinyurl.com/nhj85y8z> (29 April 2020) (last visited 30 June 2022).

<sup>225</sup> M. Kaminsky and G. Malgieri, n 213 above, 7.

<sup>226</sup> European Commission for the Efficiency of Justice (CEPEJ), ‘European Ethical Charter’ n 7 above, 9.

<sup>227</sup> See G. Comandè and D. Amram, n 173 above, 1.

<sup>228</sup> Art 10 especially under para 1 and 2. European Commission, ‘Proposal’ n 9 above.

<sup>229</sup> *ibid*, Art 10, para 4.

tool is intended to be employed, clearly aims to avoid that what has been referred above as translational bias. Nevertheless, the reference to absence of errors might need some qualification since it is technically a difficult, even if possible, requirement to guarantee.

Although the establishment of these requirements is to be certainly welcomed, these still remain quite vague and raise concerns over when a dataset used to train an AI-driven legal tool is qualitatively satisfying. In this respect, the reference to ethical standards may be useful for a more effective concretization of the proposed rules.

As the Guidelines on a Trustworthy AI underline, integrity of the employed datasets is obtained through the identification of the system's vulnerabilities, and the enactment of appropriate safeguards for the prevention of data pollution.<sup>230</sup> These ones are first related to the alignment of employed systems to relevant standards, as the ISO<sup>231</sup> and IEEE.<sup>232</sup> Adherence to these standards would contribute to control the system's accuracy and the eventual presence of biases in accordance with the principle of *data sanitisation*<sup>233</sup> even if not guaranteeing the dataset is free of error. Accordingly, for the purposes of controlling the accuracy and integrity of employed datasets, the CEPEJ Ethical Charter recommends the 'use of certified sources and intangible data with models conceived in a multi-disciplinary manner, in a secure technological environment'.<sup>234</sup> This means that the datasets from judicial decisions that feed the system's algorithms should be employed only if they come from certified sources and should not be modified before they are processed by the system.<sup>235</sup> These examples illustrate how the ethical considerations can be employed in further operational readings of the existing and forthcoming legal rules.

The quality of employed datasets is also given by their inclusiveness, that is the ability of the same datasets to reflect different population groups, without generating any unfairly biased outputs in respect to these same different groups. In particular, designers of artificial intelligence-based tools should collect relevant datasets from the relevant justice professionals, as judges, prosecutors, attorneys, or by researchers in the field, in accordance with a multidisciplinary approach.<sup>236</sup> Mixed project teams for the technology design could be a first important measure

<sup>230</sup> European Commission-High Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' n 13 above, 27.

<sup>231</sup> Iso, *Standards*, available at <https://tinyurl.com/3fjyzaxa> (last visited 30 June 2022).

<sup>232</sup> IEEE, *Security & Privacy*, available at <https://tinyurl.com/3dy3p5dm> (last visited 30 June 2022).

<sup>233</sup> See B.W. Goodman, 'A Step Towards Accountable Algorithms?', *Algorithmic Discrimination and the European Union General Data Protection' 29th Conference on Neural Information Processing Systems*, available at <https://tinyurl.com/mr3zkbkv> (2016) (last visited 30 June 2022).

<sup>234</sup> European Commission for the Efficiency of Justice (CEPEJ), 'European Ethical Charter' n 7 above, 10.

<sup>235</sup> *ibid*

<sup>236</sup> *ibid*

to assure interdisciplinarity.<sup>237</sup>

In respect to the subsequent moment of bias detection, the proposed Regulation on AI comes to support the transparency requirements envisaged under the GDPR,<sup>238</sup> proceduralizing the development of AI-driven tools, by imposing onto developers the obligation to develop a ‘technical documentation’ before the system is placed on the market<sup>239</sup> and to design the system in a manner that enables ‘the automatic recording of events (‘logs’) while the high-risk AI system is operating’.<sup>240</sup> Ultimately, it demands that the design and development of the systems is performed ‘in a way to ensure that their operation is sufficiently transparent to enable users to interpret the system’s output and use it appropriately’,<sup>241</sup> together with the development of clear instructions that should guide the operator.<sup>242</sup>

To correctly address the harms potentially resulting from biased datasets in AI-tools for legal services, systems should be traceable and data subjects’ ability to contest and search for an effective remedy to the unfair treatment received by the employed AI tools should be guaranteed. Transparency of applications employed in the legal sector becomes a crucial matter also for legal professionals to eventually state the reason for deviating from software’s recommendation preventing at the same time their professional autonomy and eventual professional liability. Indeed, an adequate motivation for having followed or not the indications rendered by the IA, could in turn protect him/her from judicial action for professional liability.

Exactly for the purposes of a fair and equitable design of AI-based technologies, the High Level Expert Group on Artificial Intelligence encourages the involvement of the different stakeholders that will be most impacted by their employment. It is assumed that by properly addressing needs and feedbacks of a wider range of users, resulting devices would also better protect data subjects’ freedom of choice.<sup>243</sup> In this respect, collective and diffuse auditing schemes could be a relevant means for assuring a societal oversight of the collection, processing and storage cycle.

Ultimately, the proposed Regulation establishes a general presumption of conformity to the above illustrated obligations, in case of conformity of AI tools to ‘harmonised standards’<sup>244</sup> and when the provider has followed conformity assessment procedures.<sup>245</sup> It also establishes the possibility for AI tools developers to

<sup>237</sup> *ibid*

<sup>238</sup> See Arts 12-15 GDPR. In this respect see G. Comandè and G. Malgieri, n 212 above, 243, 265.

<sup>239</sup> Art. 11, para 1, European Commission, ‘Proposal’ n 9 above.

<sup>240</sup> *ibid*, Art 12, para 1.

<sup>241</sup> *ibid*, Art 13, para 1.

<sup>242</sup> *ibid*, Art 13, para 2.

<sup>243</sup> European Commission-High Level Expert Group on Artificial Intelligence, ‘Ethics Guidelines for Trustworthy AI’ n 13 above, 19.

<sup>244</sup> Art 40 European Commission, ‘Proposal’ n 9 above.

<sup>245</sup> *ibid*, Art 43.

certify their products through certificates issued by notified bodies.<sup>246</sup>

Relevant certification schemes are also envisaged in the Regulation UE 2019/881 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification.<sup>247</sup> The Cybersecurity Act also establishes a framework for the implementation of common European cybersecurity schemes for digital services and products, which could be relevant for enacting certification regimes specifically regarding artificial intelligence tools for the legal sector.

For these purposes, the Agency for Cybersecurity has an important consultancy role in technological matters. The cybersecurity schemes drafted by the ENISA and enacted by the European Commission in accordance with Arts 47-50 of the Regulation, can be relied on by businesses to issue the EU statement of conformity whereby the fulfilment of the requirements set out in the scheme is declared.<sup>248</sup>

The adherence to issued certifications for artificial intelligence-based products employed for legal decision making could be a further means to ensure that offered services follow requirements that assure the integrity and accurateness not only of stored and processed data, but also of given results/outputs generated by the machine.<sup>249</sup>

## **VII. Conclusions: Bridging Market Developments, Ethical Principles and the Legal Framework**

The definition of the relevant ethico-legal framework in the design of the highly specialized computational models destined to AI-drive legal services is a precondition for a legally-sound expansion of the correspondent emerging market. At a policy level, it also poses fruitful grounds for the theoretical debate regarding the ethical legitimacy of the deployment of such tools in the legal practice and judicial decision-making as well as the more balanced relationship between ethical and legal boundaries of their development and use.

Our analysis shows how in front of the emerging phenomenon of machine-driven legal decision making, both businesses and regulators should take concrete actions to more efficiently combine the merits of traditional and alternative legal services delivery models.<sup>250</sup> The study has demonstrated that a correct implementation of the described existing legal framework into the design of employed automated driven tools can minimise the threats to some fundamental

<sup>246</sup> *ibid*, Art 44.

<sup>247</sup> Regulation UE 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act), OJ L 151, 7 June 2019, online available at <https://eur-lex.europa.eu/eli/reg/2019/881/oj>.

<sup>248</sup> *ibid*, Art 53.

<sup>249</sup> *ibid*, Recital 75.

<sup>250</sup> D.B. Wilkins and M.J. Esteban, n 96 above.

values of the legal system, such as the principle of equality and non-discrimination, the principle of transparency of judicial decision-making and the principle of fair treatment/process, while reducing legal constraints (e.g. by a rational use of Art 11 GDPR and the Open Data Directive) and opening legally-relevant data to further use and economic exploitation, benefiting both the aims of the legal system and the development of a suitable market. In this perspective, each of the identified lines of intervention should be considered more in depth.

The employment of artificial intelligence techniques as a tool for conducting legal assessments activities still leaves open additional concerns. Part of them directly stem from the difficulties of subsuming traditional legal notions and concepts into the standardised criteria the machine reads: general legal notions, in particular so-called general clauses, as for example the good faith clause, may be of too vague and ambiguous nature to be incorporated in machine learning processes. Questions thus arise regarding whether these can be transformed into machine-readable parameters that are adequately understood, processed and thus applied by automated-driven systems; or whether automated systems can correctly combine different normative criteria, the association of which can return different legal results on the basis of the weight to be given to each of the considered criteria relevant for the case under scrutiny. Moreover, the same normative parameters could be subject to reforms or changes in the interpretation, not only directly, but also indirectly.<sup>251</sup> A change in interpretation could, eventually, require an update of the criteria governing the machine, if this does not occur automatically due to the open configuration of employed datasets.

Similarly, a second, more general, problem on the background of the present analysis relates to the persisting difference between artificial intelligence-driven decision-making and human legal reasoning. The human mind is indeed capable of ‘creatively’ substantiating legal notions in the unpredictable and unique set of circumstances of a specific case. As Floridi has argued, it is capable of giving meaning to it, in accordance with its experience and its acquired ‘semantic capital’.<sup>252</sup> Could the machine reproduce this creative, and to some extent ‘irrational’ component of legal reasoning? The problem arises especially in respect to general clauses – and in respect to standards as far as the American system is concerned –, which have traditionally opened up the leeway to transformative interpretations by the judicial.<sup>253</sup>

These considerations lead to the much broader and problematic issue regarding the building of a legal system that incorporates algorithm-driven legal services and where the element of the creativity of the human reasoning in legal

<sup>251</sup> This occurs for example in case of cross-referencing between different regulations.

<sup>252</sup> L. Floridi, ‘Semantic Capital: Its Nature, Value and Curation’ 31 *Philosophy & Technology*, 481 (2018).

<sup>253</sup> This is underlined by R. Pardolesi, and A. Davola, ‘Algorithmic Legal Decision Making: la fine del mondo (del diritto) o il paese delle meraviglie?’ *Questione Giustizia*, 1 available at <https://tinyurl.com/uz9scdxy> (2020) (last visited 30 June 2022).

matters is nonetheless preserved and adequately balanced with the opportunities provided by the considered innovations.

Against this backdrop, after having enquired how the market for AI-driven legal tools needs to evolve *vis à vis* the existing ethic-legal framework, the further question needs to be addressed regarding how these technologies are to be built into the legal system. This last question refers to the need to find an adequate point of balance between machine-generated and human-reasoned law. In respect to the specific issue of a digitised legal system, this means that the legal tasks that are easier to be performed – because the legal analysis is more ‘proceduralized’ and requires less interpretative efforts, as occurs with the calculation of family allowances – can more easily be allocated to automation – always with the supervision of the competent judge or lawyer evaluating algorithms and revising eventually the decisions – for cost saving and efficiency reasons. More complex scenarios shall be, conversely, analysed by actual judges and lawyers. On these premises, a question arises: who decides what is sufficiently simple to be fed by the machine and challenging enough to be assessed by a human?

The integration between machine-driven tools and human interpretation endeavours can however be looked at also from the very opposite perspective. The greater information processing capabilities offered by intelligent agents could be exploited for the solution of more complex legal situations, enabling in some cases to overcome the traditionally acknowledged limits of the legal system, regarding the structural incompleteness of legal rules and of the information held by legal practitioners.<sup>254</sup> In this perspective, the new possibilities of personalised law offered by technologies based on algorithmic processing techniques are being currently evaluated by a strand of the scholarship,<sup>255</sup> for their potential to deliver ‘ex ante behavioural prescriptions finely tailored to every possible scenario’<sup>256</sup> and, in particular, to specific categories of addressees, which can be informed and thus oriented in their actions nearly in real time.<sup>257</sup> From this perspective, thus, an additional issue will have to be addressed in a near future, directly regarding how these prescriptive commands stemming from what has been defined as ‘self-driving law’,<sup>258</sup> are to be integrated and combined with the abstract and widely remedial call of black letter law.

<sup>254</sup> *ibid*

<sup>255</sup> A.J. Casey and A. Niblett, ‘Self-driving Laws’ 66(4) *University of Toronto Law Journal*, 429 (2016); *Id.*, ‘Framework for the New Personalization of Law’ 86(2) *University of Chicago Law Review*, 333, 358 (2019). See also, G. Wagner and H. Eidenmueller, ‘Down by Algorithms? Siphoning Rents, Exploiting Biases and Shaping Preferences- The Dark Side of Personalised Transactions’ *University of Chicago Law Review*, 86, 581 (2019).

<sup>256</sup> A.J. Casey and A. Niblett, ‘The Death of Rules and Standards’ 92(4) *Indiana Law Journal*, 1401, 1402 (2017).

<sup>257</sup> Reflecting over the personalisation opportunities for the legal system given by artificial intelligence, C. Busch and A. De Franceschi, *Algorithmic Regulation and Personalized Law* (London: Bloomsbury Publishing, 2021).

<sup>258</sup> The expression is employed by R. Pardolesi and A. Davola, n 253 above.

All these further, challenging, issues related to the ongoing processes of automatization of our legal system can be managed with greater responsiveness by policy makers and regulators if solid patterns of interconnection between the market, the ethical and the legal spheres are established. This analysis has tried to show possible interaction channels between these three different levels. It has demonstrated how the legal framework can act as a catalyst in order to blend evolving markets of AI tools for legal decision-making with established ethical principles.

In this respect, the study has shown how contrary to the dominant view that stresses the importance of a fundamental rights-based approach, the focus should be set onto the market-based regulatory framework, which comes to directly model emerging businesses in the AI-driven legal sector. In this respect, open data regulations, data protection law under the GDPR and the announced proposal of a regulation on artificial intelligence have been assessed as a primary means to implant ethical values in the market. From an opposite perspective, it has been suggested how ethical declarations themselves can be a driving force for the design and implementation of future regulations over artificial intelligence.

As we believe, the proposed hermeneutical model is particularly important for the purposes of an ethically-sound development of the sensitive market of artificial intelligence applications for the legal sector. It could nonetheless be of paradigmatic relevance also for the regulation of other fundamental rights-invasive applications of artificial intelligence, as is the case of AI-driven health applications, or of artificial intelligence tools deployed in the financial sector.