Artificial Intelligence and the Transformation of Humans, Law and Technology Interactions in Judicial Proceedings

Francesco Contini
Legal Informatics and Justice Systems Research Institute, National Research Council of Italy, Italy

Abstract

The paper connects the potentially disruptive effects of Artificial Intelligence (AI) deployment in the administration of justice to the pre-existing trajectories and consequences of court technology development. The theoretical framework combines Luhmann’s theory of technology with actor–network theory to analyse how the new digital environment affects judicial agency. Then, it explores law and technology dynamics to map out the conditions that make legal the use of technologies in judicial proceedings. The framework is applied to analyse ‘traditional’ digital technologies (simple online forms and large-scale e-justice platforms) and AI-based systems (speech-to-text and recidivism assessment). The case comparison shows similarities and dynamics triggered by AI and traditional technologies, as well as a radical difference. While system developers and owners remain accountable before the law for the functioning of traditional systems, with AI, such accountability is transferred to users. Judges—users in general—remain accountable for the consequences of their actions supported or suggested by systems that are opaque and autonomous. This contingency, if not adequately faced with new forms of accountability, restricts the areas in which AI can be used without hampering judicial integrity.

Keywords: ICT; Artificial Intelligence; e-justice; e-government; judicial independence; courts.

Introduction

In the last few years, scholars, practitioners and policymakers have realised that Artificial Intelligence (AI) might bring disruptive changes to the functioning of courts and particularly judicial decision-making. With few courts using AI-based systems, the debate has been mostly conjectural and futuristic, while scholars elaborate their analysis on the few available systems that support judges’ decision-making. The disruption is yet to come, and software developers are inflating expectations. However, the very notion that judges deciding cases will have to consider suggestions made by algorithms, or that algorithms will decide cases without the involvement of judges catches the attention of all those interested in the administration of justice.

1 The paper deals with the use of digital technologies within the courts’ systems, and not with the broad field of AI and the law.
2 As noticed by Surden, “Artificial Intelligence and Law,” 1306.
3 Since the end of the 1980s, a first wave of AI technologies has promised to revolutionise the legal field and the adjudication of cases, with negligible practical results. There is a growing suspicion that the current AI cycle has also reached its peak of inflated expectations. Laurence Goasduff, “Gartner Hype Cycle,” Reiling, “De Rechtspraktijk, Toepassing Van Ai in De Rechtspraak.”
4 Niiler, “Can AI Be a Fair Judge in Court?”
The debate considers AI as another powerful tool to be used to improve the functioning of justice and is decoupled from the past and current trajectories of information communication technology (ICT) innovation in justice systems. AI is presented as a radically new technology that has nothing in common with the digital technologies used by courts for many years, such as case management applications or e-filing. However, this paper approaches the issue of digital innovation in the administration of justice from a different perspective. It argues that the similarities and differences between AI and the other digital technologies are relevant and help to identify the consequences of AI’s introduction into the judicial domain. To emphasise the distinction between AI and the other digital systems used by courts, the paper refers to this second group as ‘traditional’ digital technologies. Traditional technologies encompass a variety of applications, including those for case management, e-filing, integrated justice chains, e-justice platforms, video technologies, legal databases, human resources and accounting systems. The trajectory of these ‘traditional’ applications and assessment of some of the dynamics observed when law and technology work together in judicial proceedings help to anticipate some of the potential impacts of AI. The impact assessment aims to avoid the new wave of technological innovation hampering the fundamental values supporting judicial decision-making and the administration of justice—particularly independence, integrity and fairness—and hence undermining the fundamental rights protected by courts. The risk is not hypothetical. The Council of Europe’s Ethical Charter on the Use of Artificial Intelligence in Judicial Systems states that the use of AI should be done ‘responsibly, with due regard for the fundamental rights of individuals as set forth in the European Convention on Human Rights and the Convention on the Protection of Personal Data’. The Charter spells out several principles that should be fulfilled to achieve this goal, three of which are entangled with the focus of this paper. The fundamental rights principle underlines that ‘the design and implementation of AI must be compatible with fundamental rights’; the user control principle aims to ‘ensure that users are informed actors and in control of the choices made’; and the transparency, impartiality and fairness principle highlights that ‘data processing methods must be accessible and understandable, external audits must be authorised’. Transparency is one of the requirements for the accountability of technological systems and of the agency they enable or mediate—issues that, even if not mentioned by the Charter, are pivotal in AI ethics and crucial for this paper.

In the judicial domain, technology is often approached as a tool and consequently analysed within the functional paradigm, focusing on its instrumental nature. However, as emphasised by Lanzara, technology is not just a tool and more nuanced approaches are needed to analyse the consequences of ICT deployment in the judiciaries. Going beyond the dominant functional paradigm, the paper analyses the interactions between law, technology and humans, starting from a theoretical framework based on the concepts of functional simplification and closure coupled with those of inscription and delegation developed by actor–network theory. The effects of the mixed normativity derived from the entanglement between law and technology are then explored.

The case study analysis highlights similarities and differences between the new AI-based systems and the ‘traditional’ ICT applications and discusses some of the dynamics that occur when the law, technology and humans interact: the lack of transparency, the change of ownership of the procedure (co-ownership, or the transfer of ownership from the previous owner—typically judges or clerks—to the owner of the software code), and the often-underestimated issues of human oversight and the accountability of the new forms of agency. The concluding section of the paper, in consideration of the institutional changes triggered by the digitisation of judicial procedures, identifies the conditions that must be fulfilled to use digital technologies to support, automate or guide judicial procedures and judicial decisions.

---

5 Han, “How to Upgrade Judges with Machine Learning.”
6 Traditional technologies rely on databases, workflow and document management, communication systems, encryption and digital identity. AI systems rely essentially on natural language processing, machine learning and neural networks.
7 Council of Europe, European Ethical Charter.
9 See for instance Cerrillo, E-justice, see Steelman, Caseflow Management, chapter VII.
11 Luhmann, Risk; Kallinikos, “Order of Technology.”
Law, Technology and Humans in the Administration of Justice

Since the 1980s, ICTs have proliferated in the operations of courts, promising transparency, efficiency and radical changes to working practices. The systems already in place (e.g., case management systems, e-filing and the digital exchange of data and documents) execute and automate procedures well-regulated by the relevant codes of procedure. The impacts of such “traditional” digital technologies on the functioning of the judiciaries are mostly positive: they have helped to improve access and equal treatment, and the efficiency and effectiveness of judicial procedures. More recently, the development of AI promises a new wave of changes affecting not just procedures, but also decisions: legal analysis and advice performed by autonomous devices (legal analytics), the prediction of judicial decisions based on jurisprudence and other criteria (predictive systems), and even the capacity for autonomous decision-making delegated to ‘robot judges’. For judiciaries under pressure from high caseloads, backlogs and a lack of resources, AI brings the attractive promise of inexpensive, consistent and fast decisions.

Within the functional approach, digital technologies are considered tools to support the delivery of justice. The ministries of justice develop case management systems, e-filing or make interoperable court and prosecutor systems to automate procedures and improve efficiency, effectiveness, compliance with procedural law and access to justice. These efforts lead to e-justice, understood as the ICT-enabled and -empowered administration of justice. However, the effects of the digital transformation on the administration of justice are felt not just at the functional level, but also in institutional settings, challenging values such as independence, impartiality, fairness and accountability. The instrumental approach does not grasp the full range of consequences triggered by the introduction of ICT, particularly the new dynamics between law, technology and human actors. Actor–network theory and Luhmann’s work on technology offer more fine-grained perspectives to understand how law, technology and humans interact, and how the new systems affect the administration of justice. Actor–network theory identifies the crucial dynamics and consequences of technological innovation and offers a focus on the interplay of humans and technology. Luhmann’s concepts of functional simplification and closure give a robust view of technology at the design and operational levels, and of its interplay with formal rules.

Automation, Inscription and Delegation

The digitisation of judicial proceedings entails the migration of operations and tools from paper-based to digital media. The migration requires changes at various levels: the actors that undertake action, the steps to be taken by each of them, the data to be recorded in the official registries, the procedural documents to be completed and exchanged. All these operations combined in various workflows are pre-established by a thick normative setting that regulates how case-related information must be exchanged to have fair and accountable procedures and decisions. The normative layer is designed to drive human agency in a way that is coherent with fundamental judicial values.

Therefore, from an information systems perspective, judicial proceedings are the regulated exchanges of information required to deliver the information the judge needs to make the decision and to make accountable the entire process. Automation is one of the most visible consequences of the digital transformation. Computer programs, with their interfaces and data processing modules, automate some activities previously performed by humans. The automatic drafting of summons based on data collected in registries is an example of this dynamic. In this case, the agency shifts from humans—clerks—to the machine performing the task.

Actor–network theory, and more precisely the concepts of inscription and delegation, allows a more nuanced account of the phenomena. In addition to automating tasks, the system drives the actions of clerks, lawyers and judges: it tells humans what to do, and when and how to do it. This guidance is much more compelling than that provided by the law. While law requires human interpretation, leaving the interpreter the possibility to choose from different options, an ICT system provides users with

---

13 Fabri, Justice and Technology in Europe; Reiling, “Technology in Courts in Europe”; Nihan, “Using Technology Improve Administration.”
14 Lupo, “Designing and Implementing e-Justice Systems,” 373.
15 Sourdin, “Judge v Robot?”
16 The EU e-justice action plans are typical examples of this approach. See 2019–2023 Action Plan European e-Justice.
18 Latour, Reassembling the Social.
a ‘pre-packed’ interpretation of the law. The commands of formal regulations are ‘inscribed’ into software codes. Human actors can only do what the machine, its interfaces, data sets and data processing allow (see the PDF Form and e-Justice Platform cases below). Systems users cannot change the way the system elaborates the data entered into an electronic form. Working against the constraints and affordances offered by the machine can be impossible or demanding in terms of costs, time or personal engagement.

Through this process, the execution and enforcement of the commands of the law are delegated to software codes that become self-enforcing artefacts. In other words, while the law is a set of commands (often abstract) that must be interpreted and enforced by supervision, hierarchy and controls, the commands provided by software applications are self-enforcing. The subject who encodes the legal rules into the software code is making the actual interpretation of the law. The system will enforce that interpretation any time the human actor interacts with the machine to execute that procedure.

With this process, components of the procedures as established by the law are inscribed into the technology. The delegation of agency to machines is then twofold: machines perform actions previously done by humans (automation), while also guiding humans in performing their actions. The more e-justice is widespread, the more the authority of the law (and of formal regulations in general) is delegated to software codes. As stated by Lessig, the software becomes the code.19 Compliance obtained by technology ‘appears to be more objective or at least less questionable than formal authority, legal rule or human supervision’.20 Despite the frequent claim that ICT improves transparency, its apparent objectivity makes it difficult to question actions requested by software codes21 (see the PDF Form case below). Hence, transparency is more of an aspiration than an actual consequence of digitisation. The process of inscription and delegation, through which rule-based action is delegated or enforced by software codes, requires the encoding of formal rules and actual practices in digital media (i.e., the design, development and deployment of the system). Whether the system automates some tasks previously performed by judges or clerks or provides guidance to their actions, two key questions arise: what should the role of legal professionals be in making the inscription, principally when it entails the core of judicial functions; and how can it be checked that the inscription—and the consequent pre-packed interpretation of the law—is coherent with the text of the law and the interpretation made by individual judges or superior courts. This coherence cannot be taken for granted and there are already cases in which superior courts have ruled against the interpretation of the law inscribed in specific systems.22 The concepts of functional simplification and closure provide a second perspective that is relevant to the discussion of this issue.

**Functional Simplification and Closure**

It can be difficult to transfer into digital media practices that are smoothly running in paper-based mode, such as signature or identification of lawyers and case parties.23 Indeed, the sets of opportunities and constraints offered by conventional, paper-based methods and digital media are different. Despite the pervasiveness of the legal framework regulating the functioning of courts and judicial proceedings, each court and judge defines and establishes specific and personal working practices within the boundaries of the legal rules.24 These are working practices that shape the way the law on the book is applied, and the steps to be followed in its application. Sometimes the differences are minimal; other times they are quite relevant and based on a different interpretation of the law. Judges and clerks working in a court may have tools and interests to agree on common working practices. If we then compare the relative standardisation reached at court level among the different courts, we will probably discover that the working practices adopted by the various courts are different. The magnitude of such differences increases further when the same law has to be applied by judges operating in different judicial systems, as with several European regulations such as the European order for payment.25 This heterogeneity has to be reconciled with the features of digital technology.

---

19 Lessig, Code and Other Laws Cyberspace.
21 Contini, “Elusive Mediation Law and Technology,” section V.
22 Consiglio di Stato, “Sentenza N. 8472.”
25 Ontanu, “Adapting Justice to Technology,” sections 5.3 and 5.4.
Making Court Technology Legal

Justice systems operate using a broad range of technologies. The dynamics described in the previous section focus on how technology absorbs and changes the pre-existing forms of agency and interactions between law and humans in judicial proceedings. Some of the technologies are just input or output interfaces, while others execute or support judicial procedures. The first group of technologies includes printers, keyboards and computers performing generic functions. Such devices do not require a specific legal endorsement. The sentence can be handwritten, drafted with a word processor or with an old typewriter. What matters is to respect the formal and substantive features of the sentence, such as the structure, the signature of the judge and clerk, the registration number and the stamp.

26 Cordella, “E-Government and Organizational Change”, Section 3.
28 Kallinikos, Governing through Technology, 77
31 Ministero della Giustizia, Specifiche Tecniche, 13.
32 Cordella, “E-government and Organizational Change.”
33 In civil law countries, case assignment and the identification of priority cases are judicial prerogatives. Even if this may not be the case in common law jurisdictions, it is undeniable that they influence judicial actions, and are not simple administrative tasks.
The second group of technologies includes electronic forms, case management systems, e-justice platforms and any other technological component that registers, processes, guides and executes procedures. These applications must be approved by some regulation for use in judicial proceedings. This occurs, even if with some differences, in both civil and common law countries. In the old times, paper registries were approved by ministerial decree, establishing the data set and physical feature of the books. Codes of procedure establish if and how to take hearings records, and if transcripts must be abridged or unabridged. Some governance authorities (ministries, judicial councils or court presidents) have legalised the use of case management systems to replace paper registries, or approved the use of e-filing through some specific technological platform.

Technologies affecting proceedings must be made legal to become performative from a legal perspective. If a regulation establishes that procedural documents can be exchanged only via the court portal, procedural documents sent to the court registry via e-mail cannot be used in the proceeding. Document received in this way cannot be admitted in the proceeding; they are not considered to exist. Once the technology is made legal, it produces not just functional results, such as the transmission of data or documents, but achieves legal effects: data and documents transmitted using the authorised technology can be used in the proceeding. Thus, the filing of a document via the approved platform changes the status of the procedure. At this level, ICT applications execute and enforce legal rules; hence, technology integrates its regulative power with that provided by the law to deliver the legal effects sought.

Legalisation ensures that technology can be used to produce the expected legal results. It also certifies that the commands of the law have been properly inscribed into the technological system and that the software codes execute or guide human action respecting in full the legal code. Hence, technology is closed from a technological perspective (black-boxed) and certified for its legal compliance. The technical closure and legal certification guarantee the stability and predictability of its functioning. All this makes systems accountability feasible: the technology-based actions can be checked against the relevant normative provisions.

The discussion so far has focused on ‘traditional’ ICTs, systems variously used by courts for 25 years. Does this theoretical framework of technology in judicial proceedings fit with AI? From this perspective, the main difference between traditional digital technologies and AI is the stability of the software codes. Once closed, the software codes of the ‘traditional’ digital technologies are designed to remain stable. AI encompasses a set of distinct technologies such as machine learning and the neural network that are understood as ‘a growing resource of interactive, autonomous, self-learning agency, which enables computational artefacts to perform tasks that otherwise would require human intelligence to be executed successfully’. The self-learning and autonomous nature of AI means that the software code changes over time autonomously through learning processes. Therefore, when an authority authorises the use of AI in court proceedings, the software code that was made legal is likely to be different from the software code that is actually used. The question of how to keep legal self-learning and autonomous systems remains open and will be further explored by the paper.

Case Studies

The theoretical framework just outlined is now used to analyse four case studies. Two of these cases concern applications based on traditional ICT (PDF electronic forms and an e-justice platform), while the other two pertain to AI-based systems (speech-to-text and predictive systems). Analysing the similarities and differences among the four cases allows the identification of features and dynamics generally applicable to digital technologies, and those specific to AI.

The PDF Form

Electronic forms are one of the simplest technologies used in e-justice. As paper forms, they enable users to provide requested information in a simple and straightforward way. Due to their simplicity, they provide an excellent starting point for the analysis.

34 Vismann, Files.
On 17 December 2015, UK newspapers published the news that a divorce software error had affected thousands of family law settlements, miscalculating spouses’ financial worth.\(^{36}\) The software error had the potential to inflate the financial worth of a wife or husband, and the wrong calculation may have affected the decision regarding financial rights and duties to the parties, particularly as regards spousal maintenance. The error went unnoticed for a total of 18 months between April 2011 and January 2012 and between April 2014 and mid-December 2015. In the two periods, the software calculated incorrect figures for the net assets in 3,638 proceedings.\(^{37}\) For the cases open at the time of the discovery (a total of 1,403), the courts corrected the errors automatically. In the other 2,235 cases, the divorce procedure had to be reopened and information resubmitted.\(^{38}\) The question is not the error per se, but the reasons why the Ministry of Justice and the users did not detect the error for such a long time.

The error was nested in one of the PDF forms used by case parties and lawyers involved in civil proceedings. The faulty formula affected Form E, which aims to offer court users a manageable way to present the pieces of evidence required to decide the financial side of a divorce. It involves a pre-selection of the relevant information to be entered out of the myriad of entries that parties could consider crucial for the decision. The form is functionally simplified (it pre-selects the data set relevant for the procedure) and closed (it does not allow additional information or changes to the background calculation formula). It also shows the inscription of the rules of practice that regulate the procedures and the delegation of performative power. The form is not just the result of the work of practitioners; it was officially endorsed by Her Majesty’s Courts Service (now the Ministry of Justice) to act in that specific proceeding.

The case shows that even a minimal inscription of legal rules and procedures into a software program can have profound and extensive consequences for judicial affairs. The malfunctioning of a simple, trivial PDF form helps us to disclose some of the hidden entanglements between law and technology in court operations. Therefore, this simple case is a good starting point to explore the intricacies of law and technology in judicial proceedings and to frame our argument about the elusive nature of digital mediation.

Technology users tend to focus on the interfaces and tools that enable the technology’s use, rather than on its internal functioning. Just as car drivers tend to take for granted the internal processes of their car’s engine as they drive, case parties using an official form or web interface made available by the competent authority focus on the use of the tool itself and not on its internal functioning, technical and legal performativity, or effectiveness.

More generally, the masses of case-related data made available by court technologies increase transparency. At the same time, ICTs’ internal functioning is hard to access and difficult to make accountable even in cases as simple as the one discussed. Its correct functioning is taken for granted by users and certified by the sealing of Her Majesty’s Courts Service. Hence, a general question concerning court technologies is whether it is possible to deploy adequate controls on their inner workings and the algorithms that process the data. The actual question is how to guarantee proper oversight and the accountability of the technology in its day-to-day use. Is AI a peculiar case in this accountability exercise? This question is considered below.

**The E-Justice Platform**

Unlike PDF forms, e-justice platforms are complex systems in terms of their functionalities, components, types of users, software complexity and agency delegated to the technology. They integrate case management systems, e-filing, e-summoning and other functions. In a nutshell, they are digital platforms supporting or automating the tasks required to conduct judicial proceedings. The analysis of one of these platforms allows the discussion of the dynamics of law, technology and humans that occur when relevant aspects of legal agency are delegated to digital platforms.

Due to constitutional and historical reasons, the Italian Ministry of Justice is in charge of ICT development for courts and prosecutors’ offices. Therefore, this institution owns and controls the software, including the platform (suites of applications), used by clerks and judges to handle civil proceedings. Clerks and judges contributed to the design of the system, sharing their input and suggestions with software developers. The inscription of formal rules and working practices in the various

---

\(^{36}\) Bowcott, “Revealed.” The discussion of the case is based on a previous analysis made by the author. See Contini, “Elusive Mediation between Law and Technology.”

\(^{37}\) “Update on Investigation.”

\(^{38}\) “Update on Investigation.”
components of the platform took more than 15 years of hard work,\textsuperscript{39} and its use is now mandatory as established by the law and by-laws. From filing to disposition, technology enables any procedural step established by the code of procedure, automating tasks and procedural steps, and provides systematic guidance to clerks and judges. The transformation of the code of civil procedure into a software code is almost complete.

In practical terms, the human actors execute their tasks by interacting with system interfaces and background routines. On the one hand, the system executes several functions by itself (e.g., the preparation and service of documents, data exchange, events recording, data checks and statistical reporting); on the other, it provides a workflow that identifies the task and options available (e.g., action and tasks lists, pre-established procedural tracks, deadlines and priorities). Since the self-enforcing features of ICT are much more compelling than those of formal regulations, the large variety of court and individual practices found at the national level before the introduction of the platform have been standardised. As a result, a function or a procedural decision envisaged by the code of procedure can be easily accomplished only if inscribed in the platform. Decisions previously made on a case-by-case basis by clerks and judges are now made at the level of software development. Here, the software code establishes and controls the execution of the procedure and becomes visibly the pre-packed interpretation of the legal code.

In March 2018, a breakdown made clear this new state of affairs. Without any consultation, the Ministry of Justice (the owner of the platform) stopped the service for uploading judicial decisions into the national jurisprudential database of first instance and appeal courts decisions, and shut down access to the decisions already uploaded.\textsuperscript{40} The law is very clear about the publication of judicial decisions. The Privacy Law\textsuperscript{\textsuperscript{41}} states ‘judgments and other decisions of the judicial authority […] are also made accessible […] on the Internet, observing the precautions provided for in the following chapter’. Thus, the decision to publish rests with the judges, while the Ministry has a duty to make decisions available on the Internet. Consistent with existing regulations, the interface used by judges had various options available, including to upload the decision in full text or to ask for anonymisation.

The Ministry did not explain the decision to stop the service, but information leaks point to a data protection breach: several cases were uploaded without having anonymised the names of the persons involved. The state of affairs remained unchanged for several months; judges and bar associations harshly criticised the decision, and after some changes, the jurisprudential database went back online in July 2018.\textsuperscript{42} Regardless of any evaluation of the decision made by the Ministry, the case clarifies the unprecedented control of the day-to-day handling of judicial procedures exerted by the owner of the software code. In the Italian constitutional framework, this means a switch from the individual judge to the executive power. If technology affects the constitutional separation of powers, it is difficult to argue that it is a neutral tool. ‘Traditional’ digital technologies already change institutional settings and procedural roles. Software codes execute tasks established by the law and guide the actions of judges, lawyers and clerks. At this level, new governance mechanisms should be identified to keep technology in line with the values supporting fair trial.

\textit{Speech-to-Text Applications}

AI brings in further changes and challenges to those already discussed. The case of speech-to-text applications helps to keep separate the issues related to the specific features of AI (such as complexity, autonomy and issues of accountability) and those related to the areas in which it is applied and agency is delegated: data entry, data analysis and decision-making.

Embedded into the operating systems of smartphones, computers and other ‘smart devices’, speech-to-text is a quintessential example of AI applications already in use in everyday life, as well as in many courts and prosecutors’ offices. Speech-to-text implies the inscription of the full set of equivalences between phonemes and written terms, and hence of the rules regulating oral and written communications. This occurs through complex operations beyond the scope of this paper, except for the fact

\textsuperscript{39} Carnevali, “Pushing at the Edge,” 167–176.

\textsuperscript{40} Ministero della Giustizia, “Uffici Giudiziari,” March 14, 2018: “We inform you that the consultation of the jurisprudential archive is temporarily suspended due to systems updates. We will notify of the reactivation of this function, which will take place quickly” [translation of the author].

\textsuperscript{41} Legislative Decree 196/2003, article 51 (“Privacy Code”).

\textsuperscript{42} On 26 June 2018, the Ministry of Justice issued this message: “We inform you that the consultation of the Jurisprudential Archive is temporarily suspended due to a system update. The consultation of the Archive will be online again on 9 July 2018” [translation of the author]. See Ministero della Giustizia, “Uffici Giudiziari,” June 26, 2018.
that they work thanks to inscrutable ‘deep-learning and neural network algorithms that constantly improve the accuracy of their performance’.43

The user of the system delegates to algorithms the conversion of a spoken act into a written text. As in the cases discussed above, the process involves functional simplification and closure. The analogic flow of acoustic waves is deconstructed and digitally reconstructed in a way that is processable by machines. The technology samples and filters the acoustic waves and makes them digital. Then, the system chunks the digital flow into short signals that, through stochastic techniques (deep learning and neural network), are associated with phonemes and words. The application is closed. Users can interact using only a few commands, such as the selection of the language and the jargon (e.g., medical or juridical), and the start and end of the dictation. At the same time, the application is not frozen. The inaccessible background algorithms work not just to transform speech into text but to improve the accuracy of the transcripts.

In the judicial field, speech-to-text is used to obtain swift and accurate court hearing records, and more generally for document drafting. In Italian courts, judges commonly use such applications for hearing records in civil proceedings. They ‘dictate’ to the machine, check the output on the computer screen and, if needed, do some editing. Once ready, they process the document following the code of procedure as enabled by the platform described in the previous section: most often, they digitally sign the decision or hearing record and pass it to the clerk for the routine service of documents. Speech-to-text is a significant improvement over the previous procedure in which judges had to type documents with a keyboard, or handwrite minutes and pass them to a typist. The continuing and voluntary uptake of speech-to-text in the legal field by judges pressed by caseload also shows that the AI-based procedure is quicker and more effective than the previous one.

Further, when used for writing records, speech-to-text can increase operational transparency, with many judges having adopted a double screen system in court hearings. One of the screens is for the judge, while the other is for case parties, who can control in real-time the hearing report and ask for amendments. This crosschecking by the judge, lawyers and case parties zeros the risk of having the content of the record disputed by lawyers. In this case, the use of AI increases transparency and efficiency, and does not create problems of accountability or interfere with the due process of law. Indeed, the use of such systems does not represent a problem if the judge (or an interested subject) can swiftly control the equivalence between the inputs and outputs. Often, the use of AI is associated with critiques and ethical concerns, usually focusing on the opaqueness of the algorithm and the lack of accountability.44 However, in the case of speech-to-text, as applied in Italian courts, such applications show how even inscrutable algorithms can increase transparency. In this case, the use of AI is not a problem per se. It only becomes a problem if the equivalence between input and output cannot be checked or, if this first check cannot be performed, the procedure that transforms the input to output remains obscure.

In addition, the use of speech-to-text does not require any legal approval. Here technology is just an input device and as with any other input device, such as a keyboard or pencil, it is the user’s responsibility to check that the output is correct. The output (e.g., a hearing record or a judicial decision) becomes performative only when the draft has been processed according to the procedural rules (i.e., signed, stamped, recorded in the system and served). If speech-to-text makes an error in transforming the spoken act into a written text and the judge does not correct it, the judge is by default accountable for the error. It would be difficult to demonstrate the fault of the software developer for a wrong expression found in the text of a sentence. While with traditional digital technologies (e.g., e-forms and e-justice platforms), software developers and the ministries are accountable for the functioning of the system, with AI, such as speech-to-text, users are accountable. The silent transfer of accountability from developers to users is a deep-seated change in the human, law and technology interplay when AI applications are adopted.

Predictive Systems: The Case of Risk Assessment

Not all AI-based applications allow such simple forms of control as those discussed above. Predictive systems are the quintessential example of a more complex state of affairs. When applied in the judicial domain, such systems are designed to predict, support (or replace) traditional human decision-making, often under the label of ‘predictive justice’. The term is highly evocative: judicial decisions suggested or taken by machines, leading to automatic, inexpensive and objective decisions.

---


The expression ‘predictive justice’ is dangerously misleading. Such systems make predictions or forecasts, but such predictions are far from being similar to judicial decisions. Judicial decisions require, as a minimal standard, justifications based on an assessment of the relevant facts and the applicable regulations. In theory, what is inscribed into such systems is the fact-checking and legal assessment that informs judicial decisions. In practice, AI systems make statistical correlations between data. Then, such information is packed, summarised and delivered to the judge to ‘support’ his or her decisional duties. The decision remains with the judge, but it can be difficult for the judge to resist these ‘disinterested’ and ‘science-based’ suggestions. Therefore, the risk is that while systems developers delegate a suggestion to the system, they end up achieving a delegation of the decision they are supposed to support.

COMPAS is the best-known example of this class of technology. It is thus informative to highlight some of the implications of its use. The system has been adopted by several US jurisdictions to support judicial decision-making in pre-trial detention cases. The software provider describes the system as:

one of the most scientifically advanced assessments available—allows you to easily select any combination of its 22 risk and needs scales to effectively and efficiently inform decisions. After selecting your scale combinations COMPAS saves them as custom ‘scale sets’ for repeated use in the assessment wizard. The main COMPAS Bar Chart and accompanying Narrative Report provide for easy case interpretation.\(^{45}\)

Despite this emphasis on scientifically informed judicial decisions, COMPAS is well known for its ‘bias against blacks’ (defendants), as discovered by ProPublica.\(^{46}\)

Like similar applications, COMPAS uses algorithms (most often machine learning) that predict recidivism risks and ‘score’ defendants based on the probability that they will commit a new crime if released. Since the algorithms have not been disclosed, the methods they use to estimate recidivism risks remain unknown. The complexity of the factors that may affect recidivism is simplified in a dataset of 137 records that describe the accused and the criminal records.\(^{47}\) A questionnaire completed by probation officers provides most of the data the system elaborates. Hence, the system works based on a pre-established set of data established by software developers and algorithms never made public because they are considered industrial secrets. The problem of technological accountability is in this case harsher and almost intractable.

A non-judicial authority (a private company) has inscribed in the system the rules to be adopted to calculate recidivism risks. Such rules are closed, black-boxed, justified in general terms, and non-accessible to judges or those involved in the proceedings. Once the system is in place, it is likely the user will delegate to the machine the decision about pre-trial custody. The scoring puts the judge in an awkward position. Consider a case in which the judge is inclined to grant bail, but the scoring of the system identifies a high risk of recidivism. Is the judge ready to go against a risk assessment calculation made by a machine? What if, once released, the defendant commits a crime? In a different scenario, the judge may decide to keep a defendant in custody due to her assessment of the recidivism risk, but the AI system could predict a low chance of recidivism. Is the judge ready to ignore the suggestion of the machine and keep the defendant in custody? The risk is high that the system will increase the number of false-positives; that is, persons kept in custody who could have been released. The use of such systems provides a new and undue influence on judicial decision-making, endangering judicial independence and fair trial.\(^{48}\)

The case of recidivism assessment and predictive justice (by COMPAS and similar systems) reveals another dynamic between law and technology. In the US, court administrations make such systems available to judges without specific legal endorsement. Various reports endorsed by Federal agencies and the National Center for State Courts support the adoption of such systems.\(^{49}\) This endorsement provides a kind of scientific support for the use of these systems, but it is not a legal authorisation. The court administration invites judges to use these systems to support their decision-making. Hence, while the judge motivates the decision, they also consider the technological suggestion provided by the recidivism assessment algorithm, which is neither transparent nor accountable. This positions judges and these technologies in a loop; however, it is the judges, not the technology developers, who are accountable for the decision and its consequences. Indeed, just as with speech-to-text applications, there

---

\(^{45}\) Equivant, “Northpointe Specialty Courts.”

\(^{46}\) ProPublica defines itself as an independent, non-profit newsroom that produces investigative journalism. Angwin, “Machine Bias.”

\(^{47}\) Dressel, “Accuracy, Fairness, and Limits,” 1.

\(^{48}\) Garapon and Lassègue also envisage the risk of effet moutonnier or herd behaviour, with judges following the jurisprudential line identified by the algorithm. Garapon, Justice Digitale, 279.

is an implicit, silent accountability transfer from the developers and owners of the software to its users. The difference is that while the speech-to-text user can easily check the input–output coherence, the recidivism assessment user cannot perform the same check.

Would the same problems be present in public algorithms based on machine learning or other AI technologies? Would it be different if the decision support systems were based on a more straightforward and stable algorithm, such as one calculating a regression coefficient between two variables? To what extent do the problems relate to the peculiar features of AI? Further, to what extent do they relate to the kind of agency delegated to the technology? Comparing the four case studies helps to consider these issues.

Concluding Remarks

Most of the dynamics observed with AI also occur with ‘traditional’ digital technologies. However, unlike those technologies, AI seems to put at risk fundamental judicial values. Some questions emerge due to the features of the technology (e.g., accountability issues and system autonomy). Other issues emerge when AI applications aim to support judicial decision-making.

The four cases illustrate how functional simplification and closure capture essential dynamics of software development, and how—through this process—agency is inscribed and delegated to the ICT system. In the four cases, we find a new form of agency in which different tasks are automated, enabled by the system or executed by humans with the guidance of the system itself. The overarching effect of this transformation is a progressive relocation of the administrative and judicial agency to the digital systems, whether traditional or AI-based. This change entails a notable institutional transformation since clerks and judges increasingly work with digital systems that provide a pre-packed interpretation of the law made at the design stage. In the judicial sector, software development is not just a technical endeavour, but also a matter of interpretation of the law made outside the judicial process. The implications are twofold.

First, some form of legal authorisation to use the system is required. As with the old paper-based systems, any technology that has an official function in judicial procedures has to be approved by some formal rule or by some authority. As noticed, legalisation is not required with speech-to-text since it is just an input device and does not have a direct effect on proceedings. This difference between speech-to-text and the other case studies is not related to the features of the technology (traditional ICT or AI), but rather to the specific functions of the technology in the proceedings (input device v. official standing).

Accountability is the second issue that affects both traditional and AI-based systems, albeit in radically different ways. Even the most straightforward ICT system, such as a PDF form, can reveal limited transparency and create problems of accountability. The system is functionally simplified and closed so that users can focus on the request of the interface. The system closure contributes to reducing users’ awareness about the features and functioning of the form, including the simple background calculations. The user of a digital artefact that is approved, official and made legal does not even think there can be an error nested in its background functioning. This is a form of alienation, since human-made artefacts (the assemblage of law and technology embedded in the form) become de facto unquestionable by single users (see Mohr in this issue).

Accountability is also an issue in the e-justice platform. The case shows how the owner of the code (the Italian Ministry of Justice) takes control of the tasks the law assigns to the judges. Once the system is deployed, the dynamic becomes ubiquitous and usually invisible to users. The pre-packed interpretation of the code of procedure made available by the e-justice platform becomes taken for granted. Judges or clerks remain free to criticise the system if they think it is not coherent with legal requirements, but in practice this seldom occurs. A system breakdown—the lack of access to the jurisprudential archives—was needed to trigger critiques.

Legal authorisation is pivotal in the accountability exercise. In both the PDF form and e-justice platform cases, the legal endorsement made by the ministries (the software owners) certifies that the systems work respecting formal legal requirements.

---

51 The dynamic described occurs just when the system is deployed and becomes the taken-for-granted artefact to be used in day-to-day court operation. Conversely, during the design and implementation stage, the different interpretations of the law made by the various subjects affected by the innovation can create problems and hinder the deployment of the system.
The ministries can be called to explain why the PDF form was faulty, and why the jurisprudential archive was not available. The systems can be ‘opened’, the software codes analysed, and their functioning checked against specifications and requirements established by the law. The discovery of a difference between the specification of the legal code and what is implemented in the software code is technically possible: the systems built with ‘traditional’ ICT are stable and not designed to improve and learn over time like AI. Hence, the functional closure guarantees the software codes remain unchanged and the systems work as planned. The ministries (the owners of the systems) authorise and take responsibility for the changes and functioning of the software codes. The criterion against which the system has to be made accountable is the respect of relevant legal provisions. Stability over time is the condition that makes accountability enduring rather than ephemeral, and that makes software developers and owners accountable for the functioning of the system. The principle of transparency, spelled out by the Council of Europe’s Ethical Charter on the Use of Artificial Intelligence in Judicial Systems, is relevant, and may be difficult to implement with ‘traditional’ digital systems, as seen even with PDF forms. However, in this case, there are clear means and criteria to make the functioning of the system transparent, check it against legal standards and make it accountable and respectful of legal provisions and fundamental rights.

This classical accountability exercise for the ‘traditional’ digital technologies—that is, checking the software codes against the legal code and making the software developers and owners accountable for their functioning—cannot work in the case of AI. As seen, AI systems do not have stable software codes. Machine learning, deep learning and neural networks imply autonomous and iterative changes of the software. As a consequence, the accountability issue is more elusive than for the traditional technologies. First, the software codes that transform speech-to-text or assess recidivism risk are not disclosed since they are protected by industrial secrecy. This protection makes opening the black box to check the algorithms extremely difficult, if not impossible. However, even if AI were developed by public institutions and thus not protected by industrial secrecy, the full disclosure of the software code (i.e., full transparency) would not be advisable: vulnerabilities could be identified and exploited, and users may identify tricks to fool the algorithm. An alternative option is to assign accountability functions to an independent body regulated by state laws as suggested by the Office of the Human Rights Commissioner of the Council of Europe. In this case, the disclosure of the software codes would be limited to an expert group that would assess the system.

Prima facie, this option may solve the legal and technical concerns just outlined. To be used by courts, the system has to be certified. An independent body could check the functioning of the system against relevant legal standards and, if found to be coherent, approve its use. Through this process, the independent body would become accountable for the functioning of the system. The approach may allow the identification of gross errors and biases like those discovered in COMPAS by ProPublica. However, due to the autonomy and ongoing evolution of AI-based systems, the certification of the independent body, or at least its effectiveness, would be ephemeral: it would be dealing with a system older than that currently used. Thus, even in the case of independent oversight, the users of a system are still accountable for the consequences of its use. Further, the continuous changes of the software code would require an ongoing accountability exercise to ensure that the functioning of the AI-based system remains aligned with legal requirements. Therefore, even full disclosure, or the certification of an independent body while making the system temporarily transparent, would not solve the accountability issue.

Consideration is now given to the differences between the two AI applications covered in the study. Courts use AI applications such as speech-to-text without concern because, in this case, checking accountability against the legal code is neither needed nor relevant. The judge is accountable for the outcome of the system, but he or she can check the equivalence between the input and output. Further, speech-to-text, being an input device, does not require any form of legal endorsement; it does not have to be proved that it acts according to legal requirements. Even if the functioning of the system is not transparent, speech-to-text can be used without hampering basic judicial values. The control of the judge and the procedure (i.e., the controls of the lawyers) are the ultimate guarantors of fundamental values. As mentioned above, accountability shifts from system designers to system users.

Systems that provide judges with relevant inputs and suggestions about case decisions require a different form of accountability. Here, judges receive suggestions from the system (which can be biased, as in the case of COMPAS); however, they do not have the capacity to check how the system works and calculates recidivism risk. There is a serious lack of transparency. At the same time, the judge is influenced by the system and accountable for the decision taken. Thus, in this case, the AI interferes with fundamental rights and basic principles of judicial independence and fairness. Such problems would be present even if a stable

---

52 Heaven, “Why Deep-Learning AIs Easy to Fool.”
53 Council of Europe Commissioner for Human Rights, Unboxing Artificial Intelligence.
algorithm using a regression coefficient were to establish the risk assessment. Is this ‘suggestion’ acceptable or can it provide an undue influence in judicial decision-making? The answer is not simple, but the question makes clear that the more a system (be it digital, statistical or of any other nature) interacts with judicial decision-making, the more its introduction requires a careful assessment of its consequences, including those that may not be expected (see Mohr in this issue). As seen for both trivial PDF forms and advanced e-justice platforms, once a system affects or plays a role in judicial proceedings, it must be legally endorsed, and the endorsement can occur only when the system respects the requirements of the relevant legal framework. Accountability is part of this exercise.

Technologies—whether electronic forms, e-justice platforms, speech-to-text or predictive systems—can be introduced into the judicial process if, and only if, proper accountability mechanisms are in place. The complexity and autonomous features of AI make this accountability exercise extremely difficult. AI can be unquestionably adopted when the users can guarantee the effective oversight of the system and reasonably be held accountable for its use, as with speech-to-text. However, in the case of recidivism risk assessment,54 users cannot guarantee effective oversight of the functioning of the system and hence cannot be accountable for its use. Here, the ‘suggestions’ from inscrutable systems become undue influences on judicial decision-making, affecting judicial independence and fairness. In this case, it is advisable to adopt a precautionary principle towards such technologies until the questions about their use have been resolved from a technical and institutional perspective. This call for caution is in line with the Council of Europe’s Ethical Charter on the Use of Artificial Intelligence in Judicial Systems,55 and particularly the principles establishing the respect of fundamental rights and of maintaining user control. It is clear that transparency does not guaranty accountability. It is evident that AI can challenge fundamental rights, and it is not clear how to keep users under control. At the same time, the case of speech-to-text makes clear that AI is not a problem per se. It becomes a problem only when it influences judicial decision-making or judicial procedures without there being proper accountability mechanisms in place.

Bibliography


54 Between these two areas of application there are other areas of application—often called legal analytics—that are not discussed in this paper but that are certainly less critical than those assessing recidivism risks.

55 Council of Europe, European Ethical Charter.


http://pst.giustizia.it/PST/it/pst_3_1_wp?previousPage=homepage&contentId=NEW4540

http://pst.giustizia.it/PST/it/pst_3_1_wp?previousPage=pst_3&contentId=NEW5027


https://www.wired.com/story/can-ai-be-fair-judge-court-estonia-thinks-so/


https://quantumvaleat.wordpress.com/2016/01/22/update-on-investigation-into-faulty-online-form-used-in-divorce-proceedings/


