

Contini F.♦

DOI: 10.25108/2304-1730-1749.iolr.2022.68.103-105

UDC: 343.1

Artificial intelligence: A new Trojan horse for undue influence on judiciaries?

Abstract: The mass of case-related data is made available by court technology to increase transparency, but how systems internally analyze it is hard to access and difficult to make accountable. Hence, a general question is what the possibilities are to deploy effective controls over ICTs' inner-workings and the algorithms that process the data. A further question is, indeed, how to guarantee proper oversight and accountability over the functioning of technology and whether AI (and more precisely machine learning) is a peculiar case in this accountability exercise.

Technology - whether it is used for case management, simple web forms, or more complicated AI-based tasks - should be introduced into the judicial process if - and only if - proper accountability mechanisms are in place.

Keywords: artificial intelligence; information and communications technology (ICT); court decision; justice; algorithm; Ethical Charter.

For more than three decades, information and communications technology (ICT) advancements have burst into the operations of courts and prosecutors' offices promising transparency, efficiency and radical changes to working practices, such as paperless courts. Even if in most jurisdictions such promises have yet to be fulfilled, software programmes and algorithms are already executing growing chunks of judicial procedures. The impacts such technologies have on the functioning of justice systems and the values endorsed by the Bangalore Principles of Judicial Conduct are mostly positive.

The most recent wave of technology is based upon artificial intelligence (AI) and promises to change the ways in which judicial decisions are taken. This goal is mostly pursued through a specific technology called 'machine learning' that makes predictions by evaluating case files, both procedural documents and the associated judicial decisions. This data set, known as 'training data', is analyzed to build statistical correlations between the cases and the associated judicial decisions. The more the algorithm processes data, the more accurate it becomes in predicting decisions in new cases. For this reason, such systems 'learn' (even if just in terms of improved statistical accuracy) to replicate the outcomes that judges made in similar cases. Unlike the technological implements already in place that digitize the exchange of data and documents, this 'predictive justice' technology (as it is commonly mislabelled) aims to affect judicial decision-making¹. It is not clear if this trend is leading to better decisions or if it is undermining the proper functioning of the system.

The potential impact of such technology on the administration of justice can be explored by considering the challenges posed by information technology already in place, such as case management and e-filing. In England and Wales, a simple calculation error embedded in the official

♦ **Contini Francesco** - Senior Researcher at the Research Institute on Judicial Systems of the National Research Council in Italy (IRSiG-CNR) (Italy). E-mail: francesco.contini@irsig.cnr.it

¹ The label 'predictive justice' is dangerously misleading, as such systems make predictions, but not judicial decisions. Judicial decisions require, as a minimum standard, justifications based on an assessment of the relevant facts and applicable regulations. AI systems make statistical correlations and their forecasts are just the result of those correlations. Hence, it would only be proper to speak of actual predictive justice if the systems were to provide justifications in terms of facts and laws.

form used in divorce cases led to the wrong calculation of alimonies in 3,600 cases over a period of 19 months. The problem is not the error *per se*, but the reasons why the Ministry of Justice and the form users did not detect the error for such a long time. Technology users tend to focus on the interfaces and on the tools that enable the use of technological systems and not on their internal functioning.

The mass of case-related data is made available by court technology to increase transparency, but how systems internally analyze it is hard to access and difficult to make accountable. Hence, a general question is what the possibilities are to deploy effective controls over ICTs' inner-workings and the algorithms that process the data. A further question is, indeed, how to guarantee proper oversight and accountability over the functioning of technology and whether AI (and more precisely machine learning) is a peculiar case in this accountability exercise.

Several jurisdictions, particularly in the United States, have adopted technology that recommends how to make pre-trial detention decisions. Such applications use algorithms that calculate recidivism risks, and 'score' the defendant based on the probability they will commit a crime if released.

This kind of scoring places the judge in an awkward position. Suppose there is a case in which the judge is inclined to release the defendant into pre-trial custody, but the score identifies a high risk of recidivism. Should the judge be ready to go against the risk assessment calculation made by the machine? And what if the defendant is released and then commits a crime? The standard rebuttal to this argument emphasizes that the systems are just taking advantage of the data available. It is argued that the scientific methods employed calculate recidivism risks in a way that is more powerful and reliable than those used by individual judges. This argument is relevant, but how can we ensure that the data has no biases? Guaranteeing accountability is much more complicated than with simpler technology.

ProPublica, an American non-profit organization that conducts investigations in the public interest, compared actual recidivism with predicted recidivism. The analysis of 10,000 criminal defendants' cases found that "black defendants were far more likely than white defendants to be incorrectly judged to be at a higher risk of recidivism, while white defendants were more likely than black defendants to be incorrectly flagged as low risk" [3]. This instance shows that accountability is hard to achieve and that such systems can inject biases into judicial processes.

Technology - whether it is used for case management, simple web forms, or more complicated AI-based tasks - should be introduced into the judicial process if - and only if - proper accountability mechanisms are in place.

The problem with accountability is even more severe with AI systems that are based on machine learning. In this case, the prediction is based on algorithms that change over time. With machine learning, algorithms 'learn' (change) based on their own experience. As the algorithms change, we do not know how they work and why they do things in certain ways. If we cannot adopt effective control mechanisms, how can we guarantee proper accountability? The debate is ongoing and the precautionary principle should be adopted until such questions have been solved from a technical and institutional perspective.

The cautions and the precautionary principle mentioned in this article follow the same direction as the Council of Europe's Ethical Charter on the Use of Artificial Intelligence in Judicial Systems [2], particularly the principles of respecting fundamental rights and of keeping users in check. However, how to implement such guidelines is not yet clear. Lawyers, case parties and individual judges can definitely not be charged with such a task. It is a challenge that should be faced by pooling multifaceted competences, monitoring the functioning of the systems and benchmarking AI against the core values of the Bangalore Principles of Judicial Conduct. It is a challenge that Global Judicial Integrity Network participants are well-positioned to face [1].

References

1. The Bangalore Principles of Judicial Conduct. The Hague, November 25-26, 2002. Available at: https://www.unodc.org/pdf/crime/corruption/judicial_group/Bangalore_principles.pdf
2. CEPEJ European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment. Available at: <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>
3. Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin. How we analyzed the COMPAS Recidivism Algorithm. Available at: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>
Original source – Francesco Contini - the United Nations Office on Drugs and Crime - https://www.unodc.org/dohadeclaration/en/news/2019/06/artificial-intelligence_-a-new-trojan-horse-for-undue-influence-on-judiciaries.html

Контини Ф.*

DOI: 10.25108/2304-1730-1749.iolr.2022.68.103-105
УДК: 343.1

Искусственный интеллект: Новый способ ненадлежащего воздействия на судебные органы?

Аннотация: Судебные технологии обеспечивают доступ к массиву данных о судебных делах для повышения прозрачности, но то, как проходит внутренняя оценка системами этих данных, оценить и проконтролировать сложно. Поэтому основной вопрос состоит в том, возможно ли создать эффективные механизмы контроля над внутренней работой ИКТ и алгоритмов, обрабатывающих данные. Еще один вопрос - как гарантировать надлежащий контроль над работой технологий и их подотчетность, в частности на примере технологий ИИ (а точнее машинного обучения).

Технологии, будь то системы делопроизводства, простые онлайн-формы или более сложные программы, использующие ИИ, должны использоваться в судебных процессах только при наличии надлежащих механизмов контроля.

Ключевые слова: искусственный интеллект; информационно-коммуникационные технологии (ИКТ); судебное решение; правосудие; алгоритмы; Этическая хартия.

* Контини Франческо - старший научный сотрудник Института изучения судебных систем Национального исследовательского совета Италии (IRSiG-CNR) (Италия). E-mail: francesco.contini@irsig.cnr.it