"Getting the Future Right – Artificial Intelligence and Fundamental Rights". A view from the European Union Agency for Fundamental Rights

Oreste Pollicino

Full Professor of Constitutional Law, Bocconi University and Member of the Executive Board, European Union Agency for Fundamental Rights, Vienna.

As the section that is special issue of *BioLaw Journal* dedicates to AI & Law since one year shows, the relationship between artificial intelligence along with digital technologies more broadly, and (the protection of) fundamental rights is one of the most interesting and complex fields for legal investigation and legal imagination in current research.

Aside from this legal investigation (and imagination), there is also, at the same time, an increasing need to address the "law in action" aspect to that relationship. More specifically, in my opinion, we have to reflect from an institutional viewpoint on what the right balance is between different (political, judicial or technical) stakeholders and authorities when engaging with the issue of the protection of fundamental rights in the digital age.

More specifically, it is necessary to counterbalance the increasing role (and power) of the courts within our increasingly digitalised and interconnected societies.

This increase – one might view it as a kind of "digital judicial globalisation" – can be explained in at least as follows.

The main (substantive) reason focuses on the traditional gap between law and technology, where law lags behind technological advances. The burden of making up for this inevitable legislative inertia – at national and supranational level – falls heavily on the shoulders of the courts. In order to explain the link between the two, it could be argued that the large-scale implementation of automated technologies has the potential to cause a further transmutation in fundamental rights protection – and consequently the role of courts in overcoming legislative inertia. In addition to the changes already caused by the shift from the world of atoms to the world of bits, where constitutionalism becomes «digital constitutionalism»[1] and power is relocated among different actors in the information society, this relationship between algorithms and the courts leads judicial activism to play a predominant role in the information society.

However, such a role must not be exclusive and must be complemented, if not by political authorities (which may be gripped by inertia), then at least by technical bodies, such as the various EU agencies and, more specifically, the European Union Agency for Fundamental Rights (FRA), on the Executive Board of which I have the privilege to serve. The non-judicial nature and functions of the FRA mirror its essential role in promoting and protecting human rights (also) in the Internet era. The FRA promotes digital awareness campaigns and can play an important role in preventing violations of rights in digital settings. At the same time, it is able to complement the remedies available through the courts, as the courts are not necessarily able to restore the *status quo ante* in every case, and in some cases are not easily accessible for minorities or the vulnerable. In

[1] I take the liberty of citing my forthcoming book: O.POLLICINO, *Protection of Fundamental Rights on the Internet. A Road Towards Digital Constitutionalism?*, *forthcoming*, Oxford, 2021.

ISSN 2284-4503

other words, this an emerging model for the comprehensive technical promotion of fundamental rights in the light of technological challenges, which represents an alternative to the judicial, often fragmented, reaction to violations.
This is the theoretical and institutional backdrop against which the recent (December 2020) FRA report "Getting the Future Right Artificial Intelligence and Fundamental Rights" must be contextualised.
The key findings of the report are organised into two parts. The first general part deals with the protection of fundamental rights in relation to artificial intelligence. The second part considers three horizontal sectors: non-discrimination, data protection and access to justice. It is worth observing that the report only considers the challenges for human dignity, the right to social security and social assistance, the right to good administration (mostly relevant for the public sector) along with consumer protection (particularly important for businesses). However, there is no mention to the challenges for freedom of expression or political rights.

A) Safeguarding fundamental rights: scope, impact assessments and accountability

Considering the full scope of fundamental rights with respect to AI

The first aspect that any regulations adopted in the area of AI must take account of is their impact on fundamental rights as enshrined in the Charter and the EU Treaties. This general statement obliges the EU and Member States to rely on "robust evidence concerning AI impact on fundamental rights" so as to ensure that any restrictions of certain fundamental rights respect the principles of necessity and proportionality. This is a primary safeguard in the field of law and technology where the novel nature of technology means that there may not necessarily be any evidence or case studies for assessing the necessity or proportionality of any limitation on fundamental rights and freedoms. Another challenge is how to define AI, which needs to be updated in line with technological developments.
Within this framework, relevant safeguards must be put in place by law in order to ensure effective protection against arbitrary interference with fundamental rights, as well as ensuring legal certainty for both AI developers and users. Voluntary schemes for observing and safeguarding fundamental rights in relation to the development and use of AI can help to mitigate further any violations of rights. Therefore, the notion of legality is also firmly linked to a precise yet, at the same time, not too rigid definition of AI technologies. In line with the minimum requirements of legal clarity – as a basic principle of the rule of law and a prerequisite for securing fundamental rights – lawmakers must exercise due care when defining the scope of any such AI law.
Outside the field of AI, the challenges of contact tracing have laid bare this dilemma. The dominant narrative, according to which a trade-off must be made between the degree of precision of virus-mapping and the need to respect the quite demanding European data protection rules, is entirely misleading. This essentially focuses only on the proportionality test and not, as Article 52 of the Charter, in addition to Article 23 GDPR and Article 15 of the e-Privacy Directive show, on the necessity of the limitations on privacy. In other words, it is taken for granted that contact tracing applications will be effective in combatting the virus, and it is consequently taken for granted that it will be necessary. This is a typical expression of, in Morozovian terms, technology solutionism according to which every

problem must find an almost immediate technological solution. In this case, the solution should be the digital contact tracing system.

Using effective impact assessments to prevent adverse effects

The protection of fundamental rights in the age of AI also requires a focus on which public and private instruments actors can use to assess the impact of these technologies. In recent years, increasing attention has been placed on the development and implementation of AI technologies by the private sector. In particular, in line with existing international standards – notably the United National Guiding Principles on Business and Human Rights (UNGPs) – businesses should put in place «a human rights due diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights» (Principles 15 and 17). This applies irrespective of their size and sector, and encompasses all businesses that work with AI.

Impact assessments are an important tool not only for businesses but also for public administrations alike in mitigating the potential adverse impacts of their activities on fundamental rights. EU law requires particular forms of impact assessments in specific sectors, such as data protection impact assessments under the General Data Protection Regulation (GDPR). However, although a large number of DPIA have already been carried out, they have taken on different forms. Moreover, prior assessments, when conducted, focus mainly on technical aspects. They rarely address potential impacts on fundamental rights. According to some interviewees, fundamental rights impact assessments do not have to be carried out when an AI system does not, or does not appear to, have any negative effect on fundamental rights.

This last aspect requires an emphasis to be placed on the role of effective impact assessments in preventing adverse effects. This tool is critical in ensuring that EU lawmakers give consideration to all fundamental rights. Impact assessments should cover both private and public sectors, and should be carried out before any AI-system is deployed. They should take various characteristics into account, such as the level of automation and complexity, as well as any potential harm.

Ensuring effective oversight and overall accountability

Assessing any adverse impacts is not enough. In order to guarantee accountability, it is also necessary to ensure effective oversight and enforcement mechanisms. This is necessary in the light of the evolution of AI technologies, combined with their increasing rollout across various sectors.

A variety of bodies are potential candidates for providing AI oversight from a fundamental rights perspective. These also include specialist bodies established in specific sectors, for example banking and data protection supervisory authorities. However, many of those interviewed from the private and public sectors are uncertain about what the responsibilities of AI oversight bodies should be.

One option could be for national human rights institutions to play a primary role. These bodies are increasingly playing a leading role in monitoring and ensuring the effective implementation of international human rights standards at national level. The non-judicial status and functions of NHRIs mirror their essential role in promoting and protecting human rights. These institutions promote awareness campaigns and can play an important role in preventing violations of rights.

At the same time, they are also able to complement the remedies available through the courts, as the courts are not necessarily able to restore the *status quo ante* in every case, and in some cases are not easily accessible for minorities or the vulnerable.

B) Non-discrimination, data protection and access to justice: three horizontal themes

Specific safeguards for ensuring non-discrimination when using AI

The wide-scale implementation of AI technologies has increased the number of discriminatory outcomes. The use of such systems by law enforcement authorities or social media has revealed how this technology can seriously interfere with equality and dignity. AI technologies are used to profile individuals, and to create clusters for classifying behaviours, relationships or other characteristics. This process is heavily influenced by biases introduced by AI developers and also data sources. Moreover, big data analysis also raises the problem of inference between new data and information analysis. It is not always possible to verify the quality of these data, having been generated through correlations hidden within the logic of the AI system.

The obligation to respect the principle of non-discrimination reflects the broader need to ensure protection for human dignity as enshrined in Article 2 TEU, Article 10 TFEU (which requires the Union to combat discrimination on a number of grounds), and Articles 1, 20 and 21 of the Charter (equality before the law and non-discrimination on a range of grounds). In addition, this principle of also enshrined in various specific and detailed provisions contained in a number of EU directives.

Nonetheless, AI can also be used as tool for remedying discrimination. In some cases, AI systems can also be used to test for and detect discriminatory behaviour, which can be encoded within datasets. However, according to the FRA report, very few interviewees mentioned the possibility of collecting such information concerning disadvantaged groups in order to detect potential discrimination. Given the lack of any in-depth analysis of potential discrimination within the actual use of AI systems, there has been also almost no discussion and analysis of the potential positive effects of using algorithms to make decisions fairer.

More guidance on data protection

In the information society, information and data are primary assets. As raw materials, the processing of this information has the potential to create value, while also challenging the protection afforded to privacy and personal data. It is no coincidence that debate has extensively focused on the relationship between AI and data protection. The right to an explanation is just one example of how AI technologies challenge the entire system of data protection, which is based on transparency and accountability. There is a high level of uncertainty concerning the meaning of automated decision making and the right to human review in relation to the use of AI and automated decision making.

Data protection is critical in the development and use of AI. It can act both as a catalyst and a hindrance in this field. Article 8 (1) of the Charter and Article 16 (1) TFEU provide that everyone has the right to the protection of their personal data. The GDPR and the Law Enforcement Directive (Directive (EU) 2018/680) further elaborate on this right, and also incorporate many provisions that are relevant for the use of AI.

ISSN 2284-4503

Within this framework, there is a need for guidance on how data protection rules should apply. European and national authorities have been playing an increasingly important role in providing guidelines about how personal data should be used in various contexts. Contact tracing is a primary example. Some initial responses concerning a common European framework for contact tracing have referred not only to the GDPR but also to: the European Data Protection Board's statement on the processing of personal data in the context of the COVID-19 outbreak; the joint statement of the European Commission and the President of the European Council proposing a European Roadmap for lifting COVID-19 containment measures; the Commission guidance paper on COVID-19 applications; the release of a Common EU Toolbox for Member States by the EU's eHealth Network, a Commission-established body composed of Member State authorities responsible for eHealth matters' as well as a letter written by the EDPB in response to the guidance. In addition, the EDPB has also developed guidelines on the use of location data in contact-tracing applications.

Effective access to justice in cases involving AI-based decisions

Justice was one of the first sectors in which the application of AI raised constitutional issues. In the USA, the initial application of this system in order to calculate the likelihood of reoffending provides just one example of how these technologies are increasingly being incorporated into legal systems. The debate has considered the implementation of AI as a mere support for the activities of judges, or even as a system for replacing judges altogether. Nonetheless, one of the primary issues is ensuring access to justice and remedies against potential violations of fundamental rights and freedoms.

Access to justice is both a process and a goal and is crucial for individuals seeking to benefit from other procedural and substantive rights. It encompasses a number of core human rights. These include the right to a fair trial and to an effective remedy under Article 6 and 13 ECHR and Article 47 of the EU Charter of Fundamental Rights. Accordingly, the notion of access to justice obliges states to guarantee each individual's right to refer a dispute to the courts – or, in some circumstances, an alternative dispute resolution body – to obtain a remedy if it is found that the individual's rights have been violated.

In order to ensure that available remedies are accessible in practice, EU lawmakers and Member States might consider establishing a legal duty for public administrations and private companies that use AI systems to provide those seeking redress with information about how their AI systems work. This could include details on how these AI systems arrive at automated decisions. This obligation would help to achieve equality of arms in cases in which individuals seek judicial redress. It could also support the effectiveness of external monitoring and human rights oversight of AI systems.

A rigorous debate is currently underway concerning remedies in the field of social media content moderation. The lack of remedies for users allowing access to justice against discretionary content removal is one of the primary reasons why the EU has been considering different systems based on procedural safeguards for content moderation, as is shown by the proposal of the Digital Services Act.