ISSN: 2358-1824

Cadernos Ibero-Americanos de Direito Sanitário
Cuadernos Iberoamericanos de Derecho Sanitario

Article

The European Union legal framework for using artificial intelligence and imaging databases and imaging biobanks for research purposes: applying the notion of fairness

O quadro jurídico da União Europeia para o uso de inteligência artificial e bases de dados de imagens e biobancos de imagens para fins de pesquisa: aplicando a noção de justiça

El marco legal de la Unión Europea para el uso de inteligencia artificial y bases de datos de imágenes y biobancos de imágenes con fines de investigación: aplicando la noción de equidad

Valentina Colcelli[1]
Consiglio Nazionale delle Ricerche, Roma.
https://orcid.org/0000-0001-5488-2991
valentina.colcelli@cnr.it

Submitted on: 10/01/24
Revision on: 10/04/24
Approved on: 10/04/24

Abstract

Objective: To analyze the issue of justice and discrimination in artificial intelligence systems based on medical image databases. **Methodology:** Analysis of documents that constitute the regulatory framework of the European Union for the use of artificial intelligence, compared with the report FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging. **Results:** The study indicates that artificial intelligence trained with unbalanced data tends to generate biased predictions, which can exacerbate health inequalities and affect justice. Discrimination in artificial intelligence systems appears abstract, subtle, and difficult to detect compared to traditional forms of discrimination. **Final Considerations:** Robust regulation is necessary to ensure justice in artificial intelligence systems, considering the need for interdisciplinary collaboration to prepare this new generation of legal professionals with an enhanced perspective on the topic and its various dimensions.

Keywords: Artificial intelligence; Right to health; Legislation as topic.

Resumo

Objetivo: analisar a questão da justiça e da discriminação em sistemas de inteligência artificial, com base em bancos de imagens médicas. **Metodologia:** análise de documentos que compõem o marco normativo da União Europeia para o uso da inteligência artificial cotejados com o relatório FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging. **Resultados:** o estudo indica que a inteligência artificial treinada com dados desbalanceados tende a gerar previsões enviesadas, o que pode exacerbar desigualdades de saúde e afetar a justiça. A discriminação em sistemas de inteligências artificias se mostra abstrata, sutil e de difícil detecção quando comparadas com as formas tradicionais de discriminação. **Considerações finais:** Impõe-se uma regulamentação robusta para garantir justiça nos sistemas de inteligência artificial considerando a necessidade de colaboração interdisciplinar para preparar essa nova geração de juristas com um olhar aprimorado sobre o tema e suas variadas dimensões.

[1] PHD Law Degree, University of Perugia, Perugia, Italy. Senior Researcher, Consiglio Nazionale delle Ricerche, Roma, Italy.

Palavras-chave: Inteligência artificial; Direito à saúde; Legislação como assunto.

Resumen

Objetivo: Analizar la cuestión de la justicia y la discriminación en sistemas de inteligencia artificial, basándose en bancos de imágenes médicas. **Metodología:** Análisis de documentos que constituyen el marco normativo de la Unión Europea para el uso de la inteligencia artificial, cotejados con el informe FUTURE-AI: Principios Rectores y Recomendaciones de Consenso para una Inteligencia Artificial Confiable en Imágenes Médicas. **Resultados:** El estudio indica que la inteligencia artificial entrenada con datos desbalanceados tiende a generar predicciones sesgadas, lo que puede exacerbar las desigualdades en salud y afectar la justicia. La discriminación en los sistemas de inteligencia artificial se muestra abstracta, sutil y de difícil detección en comparación con las formas tradicionales de discriminación. **Consideraciones finales:** Es necesaria una regulación robusta para garantizar la justicia en los sistemas de inteligencia artificial, considerando la necesidad de colaboración interdisciplinaria para preparar a esta nueva generación de juristas con una perspectiva mejorada sobre el tema y sus diversas dimensiones.

Palabras clave: Inteligencia artificial; Derecho a la salud; Legislación como asunto.

Introduction and paper's scope

The new interest in the legal and ethical features of Artificial Intelligence (AI) is under development all over the world, and in the European Union (EU). In the EU legal system, the new interest in juridical and ethical aspects of artificial intelligence is under development due to the publication of important documents at the international level, such as the European Commission's High-Level Expert Group on Artificial Intelligence (AI HLEG); the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems; and particularly the Ethically Aligned Design IEEE document, and especially due to the publication of the 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 so-called AI Act[(1)].

In April 2021, the European Commission proposed the first EU regulatory framework for AI. It says that AI systems that can be used in different applications are analysed and classified according to the risk they pose to users. The different risk levels will mean more or less regulation.

This one rules for the first time in the EU legal system how to produce the products based on the use of AI algorithms based on a risk approach. The regulation summarises a quite long path realised by the EU through various soft law documents that tried to realise an approach to AI based on fundamental rights in the EU legal system.

The main idea underpinning the EU approach to AI is that artificial intelligence algorithms should incorporate the principles of democracy, the rule of law and fundamental rights from the design stage.

The paper underlines how the idea of creating a new culture of AI through design incorporating the principles of democracy, the rule of law, and fundamental rights asks—first of all—to the jurists, judges, and lawyers—a deep analysis of how to apply traditional juridical principles, such as, for instance, the principle of no discrimination because of the algorithm's use.

We know that one of the principles underpinning the application of AI in the light of fundamental rights means that the AI must be trustworthy.

The EU guideline on AI Trustworthy realised by a group of experts for the EU Commission, and now recalled by the Regulation Act, defines Trustworthy by four ethical issues: 1. the respect for human autonomy; 2. the preservation by damage; 3. the explainability 4. and the fairness.

Fairness and the principle of nondiscrimination are realisable by the correct selection of the data and personal data used for training and testing AI algorithms. A very well-known, scholastic, and classic example is the AI algorithm used by a bank that realised discrimination because it was not able to lend mortgages to black women because it was trained just by personal data associated with a man with white skin.

In real life, the point is that usually, violation of the principle of non-discrimination in the case of the application of AI is not so intuitive.

To reach the above-mentioned goal, we analysed the Report titled "FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging. The report reproduces principles and related recommendations defined by accumulated experiences from five EU projects on AI in Health Imaging (AI4HI Network). Report reproduces principles and the related recommendations and best practices were defined by building on accumulated experiences and results from five large European projects on AI in Health Imaging (the AI4HI Network, comprising the EuCanImage, PRIMAGE, CHAIMELEON, INCISIVE and ProCancer-I projects). We will focus on 'Fairness' designed by FUTURE-AI 'For Equitable AI in Medical Imaging'. In the lack of a specific regulation of the biobank phenomenon, the experience of others becomes precious. In this framework, the experience of important research projects could be analysed and taken into consideration.

The paper is organized as follows to show how the principle of fairness in the field of AI is difficult to apply following the traditional juridical approach based on direct and indirect non-discrimination.

The EU legal framework for Artificial Intelligence

A large number of documents are provided by the EU about AI. The issue of the use of algorithms and artificial intelligence, in general, is the subject of documents such as the Resolution of the European Parliament of February 16, 2017, with recommendations for the Commission on "civil law rules on robotics"; the Communication of the European Commission of April 25, 2018, on "Artificial Intelligence for Europe" (COM(2018) 237 final); the Brussels Commission Communication "Building trust in human-centric artificial intelligence" of April 8, 2019, COM(2019) 168 final and, more recently, in the White Paper on Artificial Intelligence of February 19, 2020 (COM(2020) 65 final). Within the framework of the Council of Europe, the Committee of Experts on Internet Intermediaries published the study Algorithms and Human Rights in March 2018 and shortly after, on December 3, 2018, the European Commission for the Efficiency of Justice Systems (CEPEJ) approved the "Ethical Charter on the use of artificial intelligence in justice systems and their environment".

The European Commission published on 19 February 2020 a Communication called "A European strategy for data" that is part of a wider package of strategic documents, including also a "Communication on Shaping Europe's digital future" and a "White Paper on Artificial Intelligence" as the European approach to excellence and trust. Because data is now at the centre of this transformation, and more is to come, the EU decided to delaine a path for creating a Single European data space in strategic economic sectors and domains of public interest – such as f.i the common

European health data space – to enable the EU to become the most attractive, most secure and most dynamic data economy in the world to face China and the USA could emerge as two data and AI superpowers, and so to avoid that data sources could be limited through to concentrations in a few places as we have with an oil-driven economy — EU wants to draw from many, diverse sources and future applications that will emerge from new and unexpected players around data. This is the framework because the EU will create a single market to make it easier for businesses, public authorities and researchers to access high-quality data to boost growth and create value for ensuring the proper functioning of internal markets, as characterised by the free movement of goods, capital, services and persons(2). White Paper on artificial intelligence indicates how the Commission will support and promote the development and uptake of artificial intelligence across the EU. Commission puts forward the proposed regulatory framework on Artificial Intelligence with the following specific objectives:

- ensure that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values;
- ensure legal certainty to facilitate investment and innovation in AI;
- enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems;
- facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation.

Within the scope of the above-mentioned legal and cultural approach, in Europe, especially in the European Union, legislation is being developed to address the issues arising from using artificial intelligence in legally relevant decision-making. In the case of AI, the EU's objective is to build a framework of clear, shared rules based on credibility and excellence to promote an "ecosystem of trust" that constitutes a global reference. This "ecosystem" is based on collaboration with all stakeholders involved in constructing and using IA systems.

After identifying the main principles that are translated into requirements in the HLEGAI 'Ethics Guidelines for Trustworthy AI, ' it is also important to define how to proceed in the implementation phase. According to AI HLEG, fairness, robustness, accountability/traceability/reproducibility, and explainability are the core requirements of Trustworthy AI.

In addition, the Commission has recently approved a Regulation (of the European Parliament and of the Council) establishing "harmonised rules on artificial intelligence (artificial intelligence law) and amending certain legislative acts of the Union" (COM (2021) 206 end of April 21, 2021). Before its adoption and entry into force, EU documents strongly encouraged the use of the Trusted Artificial Intelligence Assessment List (ALTAI) to develop procedures to detect, rate and address potential risks.

This institutional debate, which is constantly evolving, is the basis for the adoption of specific legal instruments or the application of general rules and principles.

From this point of view, other legislative sources of the European Union can be applied to AI systems, such as those that refer to the protection of personal data (see Article 22 of Regulation (EU) no. 2016/679); to the free circulation of non-personal data; the directives against discrimination, the so-called machinery directive, the directive on producer responsibility, the directive on consumer rights and the directives on health and safety at work.

The EU legal framework for Artificial Intelligence and imaging databases and imaging biobanks for research

Originally and usually is the collection of biological material, tissues the core of the "infrastructure" for scientific research qualified as a biobank (see the OECD notion of a biological resources centre[3] and the European Union law idea of research infrastructure[4]. If a biobank may be defined as "any collection of biological materials, whether the source be a human, plant, or animal, fungi, bacteria, microorganisms or other living families, as well as bioinformatics data on such organic materials"[5], in recent years imaging data were included in the notion of biobanks; thus it is possible to talk about imaging biobanks. In this framework, the Imaging Biobanks Working Group (WG) of the Research Committee was established by the European Society of Radiology (ESR) in 2014[6].

Artificial intelligence systems in such biobanks are mainly used to transform the old qualitative or semi-quantitative clinical imaging into the full use of data-driven, quantitative perspective thinking in radiomics. The last one represented the first real effort to achieve this goal, trying to maintain the evolving paradigm of Big Data Science and Artificial Intelligence.

The use of AI for understanding medical imaging (and especially in radiology) is still very baffling in several respects: “(i) data acquisition and preprocessing (uniform protocols across clinical sites, and dependence on the device and operators); (ii) data quality (curation, annotation, segmentation); (iii) data protection for both image data/metadata and health records associated; (iv) known and unknown biases management; (v) continuous monitoring of model/concept/performance drift; (vi) reproducibility, accountability, explainability, and interpretability of model predictions. In other words, even if there exist strategies for balancing and augmenting the dataset or detecting and managing biases, not all of these strategies could be applied in the medical/clinical setting; also, well-known phenomena, such as the model overfitting as well as the underestimation of the assumption of statistical hypotheses (rarely verified), could occur and badly affect the knowledge extraction from radiological images"[7].

The EU Commission Report of the Expert Group on Dealing with Ethical and Regulatory Challenges of International Biobank Research 2012 has underlined the necessity of harmonising the rules and the legal and ethical practices governing research management in the European Union. European Union law provides a very developed regulatory framework for biomedical activities; for instance, see Regulation EU 536/2014 on Clinical Trials on Medicinal Products for Human Use; Regulation EU 2016/679 on the protection of personal data (GDPR). However, these legal sources, as with other EU laws, do not get consideration for the research activities of biobanks into consideration and instead leave regulation to the Member States.

Despite this lack of a specific legal instrument, the principles contained in the clinical trials and data protection laws provide the main regulatory basis to guarantee research participants. The EU data protection regulation factually provides a novel legal framework for the use of biobank data, which has to be adopted by the Member States. This leads to challenges in worldwide biobank cooperation and sample exchange; anyway, implementing the GDPR rules at the national level poses problems due to a lack of harmonisation.

The increasing number of research biobanks and the importance of their role in supporting medical and biological research have resulted in Actions aimed at the development and sharing of biobanking best practices and benchmarking standards, which guarantee the fundamental rights of the donors/patients and increase trust in the system while not unnecessarily aggravating biomedical

research. In fact, the creation and management of Biobank collections cause several ethical and juridical issues and may provoke societal concerns[8].

The main legal challenges in managing biobanks are finding ways to protect the interests of individuals while making essential information available for medical research. In addition to the tremendous technical and organisational challenges, the fundamental principle underpinning the governance framework for medical research is that individual research participants must be respected, while at the same time, essential information must be made available for medical research. Consequently, the data systems accompanying biobanks must be technically implemented and legally managed in an efficient way that allows the generation of social welfare benefits for the international community through collaborative research efforts while respecting the informational privacy of individuals. A known challenge in this field is that the implementation of relevant ethical guidelines and legal instruments impede collaboration and the exchange of information; for example, the EU Data Protection rules, which significantly differ in some respects from the Member State to the Member State, impede collaboration and the exchange of information.

At the same time, only a few legal systems have adopted a special law on research biobanks among EU members. In many legal systems, sometimes characterised by a hybrid system, the spaces left empty by the law, the regulation of the establishment of biobanks, the organisational needs of the infrastructure and the management of biological samples or images, as well as the protection of the rights of the stakeholders, which have to be reconciled with the needs of science, are often the result of interpretations of individuals and soft law instruments that operate according to principles and are often formed in the practice of practitioners. Without a specific binding regulatory framework, the approach becomes how to properly use existing legislation concerning other sectors. In particular, we appeal to data protection rules and clinical trial provisions, always using the classic categories of national civil codes. Although these schemes have enabled practitioners to map out a legal horizon within which to move, they do not always provide an adequate regime. Each borrowed model may be deficient and inappropriate, albeit for different reasons.

The GDPR foresees a specific derogation to the prohibition of the processing of certain special categories of personal data, such as health data, when it is necessary for scientific research (Article 9 (2) (j) GDPR and Article 89 (2) GDPR.) While it is true that Regulation (EU) 2016/679 is a wide-ranging piece of legislation that includes several provisions related to scientific research that favour it – or rather, favour an understanding of its specific needs – its application is not always easy in the context of research, or at least not well understood by researchers themselves, and because the Member States may maintain or introduce more specific provisions adapting the application of the rules of the Regulation about scientific research. Even though the availability of the derogation is not phrased clearly, the GDPR has still brought a big advantage since it has offered the national Legislature the opportunity to reconsider the entire system and design a more complete protection framework for scientific research that is also applicable to biobanking[9]. Even the Clinical Trials Regulation is not a model that adheres to the specific challenges and needs of research biobanks, primarily because the establishment of biobanks does not imply any clinical trials; moreover, the provisions on experimentation do not contemplate the hypothesis of a ‘sharing of the material’ differently from what happens in research[10]. The legislator will soon no longer be able to exempt itself from dealing with research biobanks as a research model in which different skills - ethics, information technology, medicine and law and, in this sector, aspects of the administration, private law and fundamental rights

- are called into question. Since numerous knowledge domains and different operators are involved and the scientific investigation often takes place between different centres, it would be essential to ensure a binding legal framework of reference and to create a harmony in terms that today is already lacking.

Indeed, the national and international proliferation of biobanks requires a common language, considering that only the words chosen to describe the structures can influence the patients' decisions to participate.

In the above context, we have to put Artificial Intelligence (AI) techniques in support of Imaging Biobanks. We also must consider many documents provided by the EU about AI.

The issue of the use of algorithms and artificial intelligence, in general, is the subject of documents such as the Resolution of the European Parliament of February 16, 2017, with recommendations for the Commission on "civil law rules on robotics"; the Communication of the European Commission of April 25, 2018, on "Artificial Intelligence for Europe"[(11)]; the Brussels Commission Communication "Building trust in human-centric artificial intelligence" of April 8, 2019[(12)], and, more recently, in the White Paper on Artificial Intelligence of February 19, 2020[(13)]. Within the framework of the Council of Europe, the Committee of Experts on Internet Intermediaries published the study Algorithms and Human Rights in March 2018 and shortly after, on December 3, 2018, the European Commission for the Efficiency of Justice Systems (CEPEJ) approved the "Ethical Charter on the use of artificial intelligence in justice systems and their environment".

All this has put in the European Commission's scenario a Communication called "A European strategy for data" 14, published on 19 February 2020. It is part of a wider package of strategic documents, including a "Communication on Shaping Europe's digital future" and, of course, the above-mentioned "White Paper on Artificial Intelligence," which is an EU approach to excellence and trust.

As data is now at the heart of societal and technological transformation and more will follow, the EU has outlined a roadmap for creating a 'Single European Data Space' in strategic economic sectors and domains of public interest, such as the Common European Health Data Space. This framework should enable the EU to become the most attractive, most secure and most dynamic data economy in the world, facing China and the United States, the two data and AI superpowers, and prevent data sources from being limited due to concentrations in a few places, as we did with an oil-based economy. The EU wants to tap into many different future sources and applications emerging from new and unexpected actors around data. To achieve this, the EU seems willing to create a single market by

- facilitate access to high-quality data for businesses, public authorities and researchers

- stimulate growth and create value to ensure the proper functioning of internal markets, characterised by the free movement of goods, capital, services and people[(2)].

As part of the above legal and cultural approach, legislation is being drafted in the EU that seeks to address issues arising from using artificial intelligence in legally relevant decision-making processes. In the case of AI, the EU's objective is to build a framework of clear and shared rules based on credibility and excellence to promote an 'ecosystem of trust' that will be a reference at the global level. This 'ecosystem' is based on collaboration with all stakeholders involved in building and using AI systems.

The White Paper on Artificial Intelligence sets out how the Commission will support and promote the development and take-up of artificial intelligence across the EU. The Commission's proposed regulatory framework on AI has the following specific objectives:

- Ensuring that AI systems placed on the Union market are secure and comply with existing legislation on fundamental rights and Union values;
- Ensure legal certainty to facilitate investment and innovation in AI;
- Improve governance and effective enforcement of existing legislation on fundamental rights and security requirements applicable to AI systems; and
- Facilitate the development of a single market for legitimate, secure and trusted AI applications and prevent market fragmentation.

After identifying the key principles that are translated into requirements in the HLEGAI 'Ethics Guidelines for Trustworthy AI', it was also important to define how to proceed in the implementation phase. According to the AI HLEG, fairness, robustness, accountability/traceability/reproducibility, and explainability of AI are the fundamental requirements of trustworthy AI.

In addition, the European Commission has recently approved a Regulation (of the European Parliament and the Council) establishing "harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain pieces of Union legislation"[(14)]. However, even before its adoption and entry into force, EU documents strongly encourage using the Trusted Artificial Intelligence Assessment List (ALTAI) to develop procedures to detect, assess and address potential risks[(15)].

This evolving institutional debate is the basis for the adoption of specific legal instruments or the application of general rules and principles.

At the same time, other sources of European Union law may apply to AI systems, such as those referred to herein

- the protection of personal data (see Article 22 GDPR);
- the free movement of non-personal data;
- the anti-discrimination directives, the so-called Machinery Directive;
- the producer responsibility directive;
- the Consumer Rights Directive; and
- the directives on health and safety at work, among the others.

Analysing fairness under the guiding principles named FUTURE-AI

According to guiding principles named FUTURE-AI, will try to analyse the approach realised through some elements that could find in the law opportunity for greater reflection.

Fairness, Universality, Traceability, Usability, Robustness and Explainability are the guiding principles because AI technologies can play an important role in future medical imaging.

We summarise the main questions related to Fairness aspects focusing on the elements more relevant from the legal point of view. Universality, Traceability, Usability, Robustness and Explicability will not be addressed at this time, although any elements arising from the Fairness analysis have a direct impact on the other principles, in particular Explicability.

Fairness as a subjective notion by nature? An initial clarification

The first principle of the FUTURE-AI guidelines is Fairness, which states that 'imaging AI algorithms should be impartial and maintain the same performance when applied to similarly situated individuals (individual fairness) or different groups of individuals, including under-represented groups (group fairness)'. We can consider the above-mentioned aspects, especially from a legal point of view.

Anyway, because there are many different interpretations of fairness, we believe that because we work for AI development in the EU legal framework, when talking about fairness, we need to apply the definition used in the 'Ethics guidelines for trustworthy AI' as a common framework[16].

As a matter of fact, FUTURE-AI guidelines notice that the notion of fairness is a subjective and not an objective notion in the contest of imaging AI solutions. For FUTURE-AI guidelines, fairness has a subjective definition because of the specific applications that it requires, taking into account the:

1) The Specific goals and end-users of the AI solution. These are especially true in the case of medical context-specific sources of bias.

2) A general setting that should be satisfied horizontally with respect, e.g., sex/gender, etc.

Anyway, if it was possible to understand the meaning of reasoning and the conclusions of FUTURE-AI guidelines related to Fairness as a subjective notion by nature, it is better not to talk about 'subjective notion' for 'fairness' if we take into consideration the analysis the legal point of view. Something that could be defined as 'subjective', is based on personal opinions and feelings rather than facts.

It is necessary to clarify that in the EU legal framework, any type of AI system has to be grounded in the EU's hierarchy of legal sources.

EU Ethics framework essentially means 'fundamental rights', and not a no well good clarified idea of "ethic", soft law, or encouraging basis for identifying abstract ethical principles. The EU needs to realise an AI made in the EU[17] in the framework of the European Strategy for data.

In this Single European data space, the AI has to be reliable and trustworthy. Three components should always be present during the entire life cycle of the AI system: legality, ethicality, and robustness. For the three components to be made operational in socio-technical systems, the EU uses a fundamental rights-based approach, which means respecting fundamental rights within the framework of democracy and the rule of law and providing the most values that can be made operational in AI systems.

In the above-mentioned framework, Fairness cannot be considered a subjective principle because AI systems in the EU do not operate in a "lawless world."

European Union and the national legal framework as a legally binding body of legislation, have already in place an embryonal legal system to permit the development, deployment and use of AI systems, as well as legislation on data protection, consumer protection, non-discrimination and gender equality, Law Enforcement Directive (Directive (EU) 2016/680) with a set of harmonised rules applicable to the design, development and use of certain high-risk AI systems and restrictions on certain uses of remote biometric identification systems, or the AI systems related to products covered by the so-called New Legislative Framework (NLF) legislation (e.g. machinery, medical devices, toys). Despite this, the fact remains that the existing relevant legislation is called for a review so it can fit for purpose of the new opportunities and challenges raised by AI, especially the EU non-discrimination law.

This recall is to remember how each AI system has to be evaluated in light of its present legal worthiness and that—under EU law—"Fairness" is not a principle "highly subjective," as defined in the paper. Its clear definition—also in the context of the imaging AI solution—is linked to European, national, and international laws.

Also, according to the ethical guidelines for a Trustworthy AI in the EU, or the several documents on this topic, fundamental rights are the basis of AI to realise what means the respect for fundamental rights within the framework of democracy and the rule of law, provides the most and values that can be made operational in AI systems.

It is different, on the other hand, to clarify if the adjective ‘subjective’ referred to fairness is referred to the fact that in the EU law system, there is a lack of clear EU legislative guidance to accommodate contextual equality, to face the correct representation of the patient groups in the training of the AI systems. The substantive dimension of fairness is manly-related to equal treatment under the Article 19 Treaty on the Function of European Union (TFUE), art. 21 of the EU Charter of Fundamental Rights (EUCFR) and Article 14 of the European Convention on Human Rights (ECHR), as well as EU non-discrimination law. However, the legal protection offered by non-discrimination law is called into question when it is the AI, not humans, to discriminate.

Many authors remember us:

> ‘Humans discriminate due to negative attitudes (e.g. stereotypes, prejudice) and unintentional biases (e.g. organisational practices or internalised stereotypes), which can act as a signal to victims that discrimination has occurred. Equivalent signalling mechanisms and agency do not exist in algorithmic systems. (…). The increasing use of algorithms disrupts traditional legal remedies and procedures for detection, investigation, prevention, and correction of discrimination which have predominantly relied upon intuition’[18].

In this light, we can think to clarify the definition of fairness like 'subjective notion' if the last one means that discrimination in AI systems, compared to traditional forms of discrimination, 'is more abstract and unintuitive, subtle, intangible, and difficult to detect'[18].

In any case, the EU legal system is the road map to approaching the abstract and unintuitive discrimination in AI systems, as explained above, through the judicial interpretation and the evidential requirements used by the European Union Court of Justice (ECJ), as suggested by some scholars[18].

Sometimes, a multi-disciplinary approach has to be used to permit reduced subjectivity, because this one can identify as many possible sources of application-specific bias and inequity as possible through the interaction between medical doctors, the technical community and social scientists, especially jurists, but not in general 'ethicists':

> ‘What needs to be avoided is a situation where system developers and controllers alone set regulatory thresholds for discrimination at the local and subjective level without external regulatory or judicial input’[18].

The substantive and procedural dimensions of fairness and the correct representation of the patient groups in the training of the imaging AI systems

‘Ethics guidelines for trustworthy AI’ approach fairness in two dimensions: substantive and procedural (see par.).

The substantive dimension of fairness implies a commitment for:

1) "Ensuring the equal and just distribution of both benefits and costs, and ensuring that individuals and groups are free from unfair bias, discrimination and stigmatisation";

2) Fostering “equal opportunity in terms of access to education, goods, services and technology";

3) "The use of AI systems should never lead to people being deceived or unjustifiably impaired in their freedom of choice";

4) Respecting by AI practitioners the principle of proportionality between means and ends.

5) Consider carefully how to balance competing interests and objectives.

From the substantive point of view, in an Imagine Biobank, the first point to handle is the correct representation of the patient groups in the training of the imaging AI systems. From a substantial point of view, fairness can be negatively affected by both quantitative and qualitative biases. From the juridical point of view, the errors - that could be derived from the training of AI algorithms - become biased towards under-represented groups and hence exacerbate existing health disparities.

The correct representation of the patient groups is to be realised both from the medical point of view and in the light of sex, ethnicity, geography and socioeconomics representation, when and if those elements have an impact on existing health disparities. The "existing imaging databases are often imbalanced according to sex, ethnicity, geography and socioeconomics"[22]. In this case, there is a risk that trained AI algorithms become biased towards under-represented groups and hence exacerbate existing health disparities". Because of that, AI tools can generate undetected errors, with harmful consequences to the patient, when they are applied to imaging conditions that may differ or unexpectedly deviate, even slightly, from those used for training[19].

Several studies illustrate that the main cause of algorithmic bias in AI is training datasets that often lack the quantitative and qualitative diversity and balance needed to achieve AI solutions that maintain the same performance across human groups and subpopulations.

Specifically, if an AI algorithm is trained with imaging data that is imbalanced concerning gender[20] socioeconomic or ethnicity[21] and due to differences in health within and between these groups, the resulting model will likely lead to biased predictions.

Although the problem of bias in AI is common to all medical applications, it is particularly relevant in medical imaging. Here, personal attributes such as gender, age, ethnicity, and socio-economics - which are essential to complying with the principle of no direct or indirect discrimination required by the standard - are not always retained during the data preparation and image anonymisation process. This operation, although functional to the image anonymisation process and capable of minimising the possibility of patient identification, can however open the door to the bias of the AI system for possible differences in health both within and between human groups and sub-populations in which sex, age, ethnicity and socioeconomic status are detected.

To avoid any problem linked to the AI algorithm training concerning sex, socioeconomics or ethnicity, and given the health differences within and across these groups, the paper offers a solution to appeal to Multicentre data collection that can increase the diversity of datasets. It is also possible to suggest the use of distant mode training for the AI imaging solution, so-called synthetic data. Since personal and non-personal data in a dataset are inextricable and linked, synthetic data could support training.

Although Multicentre data collection is a good solution, if the prior selection of represented groups guides it, it is not certain the perfect balance between different groups could be achieved in AI

training and testing datasets the way to ensure compliance with the principle of fairness and non-discrimination as established by EU law and its Court. EU law guides the specific fairness settings and general fairness requirements that should be met in general and in specific clinical objectives, using the risk approach to select the fairness requirements that should be met.

In any case, a simple multicentre data collection might not be enough to avoid bias in assessing potential discrimination; because just statistical elements could not be satisfactory. Defining one or more disadvantaged groups, legitimate comparison groups and evidence of 'particular disadvantage' requires making case-specific normative choices that reflect the local political, social and legal dimensions of the case, as well as arguments made by the possible claimants and alleged offenders.

Although the ECJ jurisprudence offers very few clear examples of static requirements or thresholds that can qualify as a legal standard capable of defining key concepts and groups underlying discrimination, it is necessary to resort to the court itself to establish what is meant by prima facie discrimination.

European non-discrimination law exists in both primary and secondary law. We already recalled the primary law (par. 5.1.2). Talking about the secondary law, four non-discrimination directives apply to both the private and the public sectors: the Racial Equality Directive (2000/43/EC), the Gender Equality Directive (recast) (2006/54/EC), the Gender Access Directive (2004/113/EC), and the Employment Directive (2000/78/EC). The non-discrimination directives only ruled a minimal standard and provide a general framework that needs to be transposed into Member states' legal systems, introducing a fragmented standard for non-discrimination regulation across the Member States. In any case, they are the EU standard for non-discrimination

> ‘European non-discrimination law addresses two general types of discrimination: direct and indirect. Direct discrimination refers to adverse treatment based on a protected attribute such as sexual orientation or gender. Indirect discrimination, on the other h describes a situation where “apparently neutral provision, criterion or practice” disproportionately disadvantages a protected group in comparison with other people’[18].

To bring a case concerning alleged direct or indirect discrimination under EU non-discrimination law, the plaintiff must meet several evidentiary requirements that establish *prima facie* discrimination together. To found a claim on prima facie discrimination, claimants have to prove that (1) a specific injury has happened or is oncoming; (2) the injury is revealed or is oncoming in a significant way within a protected group of persons; and (3) the detriment is disproportionate to others in a similar situation. At the same time, the judge must decide whether the disadvantaged group has been or is likely to be significantly disadvantaged compared to another group in a similar situation and, thus, whether there is prima facie discrimination. For example, under the Racial Equality and Employment Directives, a comparison must always be made between the effects of the contested measure on two groups (see Case C-167/97)[18]. As suggested by the same cited scholars[18], the approach to evaluating an AI system in the EU legal system could be realised by evidential requirements used by the ECJ, also under a recognised critical incompatibility between European notions of discrimination and existing work on algorithmic and automated fairness. A static basic fairness metric, in keeping with the 'gold standard' set by the ECJ for assessing potential discrimination, should be set by technicians for Imaging biobanks to support legal practitioners in cross-disciplinary collaboration. *Prima facie*

'gold standard' for assessing discrimination has been put forward by the European Court of Justice but has not yet been translated into standardised assessment procedures for discrimination arising from AI. The reflection on the metric unit of equity on a static basis might have its peculiarities in Imaging Biobank, given the types of biases deriving from AI algorithm trained with imaging data, due to the above mentioned.

The procedural dimensions of fairness and the complete identification of the people involved in the management of image biobank

Fairness analysed under the procedural dimension by Ethics guidelines for trustworthy AI "entails the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them". This is the reason because "the entity accountable for the decision must be identifiable, and the decision-making processes should be explicable".

To reach this goal, we must consider the specific characteristics of a biobank and human components.

> 'Only the correct and complete identification of the people involved in the process of creating and using biobanks and their specific roles allow to correctly define the legal aspects that must be respected in this field. The proper application of regulation is the fundamental requirement of the whole processes'[22].

This is true, especially in the context of biobanks, often characterised by the absence of a law aimed at governing the phenomenon, as we told (par.2).

'People involved with any stage in an AI product's life cycle must understand it deeply'[23]. Different groups of 'people/stakeholders' cover different roles in the process of verifying the application of the requirements for AI:

> 'Persons who develop the requirements, people who use the systems, and end-users, who, for example in the healthcare, can be represented by patients who use treatment based on the results produced by Artificial Intelligence tools'[22].

Regarding medical images in a biobank and their management through AI systems, the characterization of people during the collection and data structuring have to be taken into consideration for approaching the issues related to the application of AI systems to image biobanks (and so, the point a and b above-mentioned). This approach will permit the verification of the whole cycle of AI systems in the use of images for health purposes in the light of the EU framework on AI.

As matter of fact, qualitative biases affecting fairness may consist of cognitive biases of clinicians generating, interpreting or annotating the imaging data[23], thus understanding the human component in an image biobank, is useful for several points, especially accountability of the decision.

As matter of fact, AI tools affect the decision-making and interpretation skills of experienced and less experienced radiologists. It is well known in the medical literature that the integration of AI in real-life clinical practice can raise issues of fairness, including its use by experienced and less experienced physicians and the effect it has on their decision-making abilities. For example, with the interpretation of mammography images, one study reported how AI automated support positively influenced the decision-making of radiologists with less advanced interpretative skills. The same study

above-mentioned describes how the AI system hurt the decision-making of radiologists with advanced image interpretation skills. This is related to overconfidence in the outcome of automated support and devaluation of physicians' professional experience accumulated over years of medical practice. It should also be taken into account that this may be dictated by misconceptions about the limitations and the strengths of AI.

Additional, in the case of the framework suited to testing for discrimination in AI systems, a recent early-stage research study reports evidence of how artificial intelligence algorithms can identify race from a person's medical scans (e.g. chest and hand X-rays and mammograms,), even when human medical experts would not be able to understand it in images without other references to personal data. In this particular case, the algorithmic and human discrimination problems, although different in nature and character, have an impact on the principle of explicitness and fairness. This situation could generate a decision that is said to be based on AI, but, instead, generated by different and unequal treatment of patients according to their race even in cases where the latter is not a relevant medical treatment criterion. In addition, in this case, the characterization of people, during the collection and data structuring, is relevant.

On the other hand, the “end-user” term in a biobank can then assume even a broader meaning (supplier and user or final user): the researchers and the patients, since patients have both a private dimension and a social dimension.

The patient is often represented only as a single person but also has a social dimension, represented by the community of people who generally use the output of processing data contained in the biobanks. In the specific context of biobanks, a further element is added to the classification expressed above the difference between supplier and user is inserted. A supplier is the one who allows the collection of his information, physical and not material. Each sample, even physical, corresponds to a profile of the person to whom that profile belongs. The user or final user is represented by the person who takes advantage of data processing stored in biobanks[22].

The qualification at a specific stage of the process who is the end-user in a biobank opens the question of how to explain decision-making processes if the end-user is represented by a patient/data subject who uses treatment based on the results produced by Artificial Intelligence tools'[22].

Conclusion

Discrimination in AI systems is more abstract and unintuitive, subtle, intangible, and difficult to detect compared to traditional forms of discrimination.

The paper talks about the case - as an example - of the application of AI algorithms to diagnostic imaging, used to detect cancer and/or develop new biomarkers. The references is the European projects on AI in Health Imaging (the AI4HI Network, comprising the EuCanImage)

AI techniques are used—e.g., in imaging biobanks—to transform old qualitative or semi-quantitative clinical imaging into a fully quantitative and data-driven perspective. This approach is known today as radiomics: starting from clinical images, AI applied to diagnostic imaging develops mathematical models to classify diseases and/or predict their development.

It is very well documented in international scientific papers that an AI algorithm trained with imaging data that are imbalanced with respect to sex, socioeconomics or ethnicity will likely result in a model that will lead to biased predictions given the health differences within and across patient groups.

If the problem of bias in AI is common to all medical applications, it is even more problematic in medical imaging as personal attributes such as sex, age, ethnicity and socioeconomics are not always retained during the data preparation to minimise the possibility of patient identification because of the use of images.

Imaging AI algorithms are impartial and maintain the same performance when applied to similarly situated individuals (individual fairness) or to different groups of individuals, including under-represented groups (group fairness).

The central aspect of a fair approach in the use of ai in the treatment of images lies in the choice of an adequate algorithm and the construction of a training dataset that is capable of a representation of patients that does not hide common biases based on gender, ethnicity, income and geographical origin.

From the juridical point of view, the errors—that could be derived from the training of AI algorithms—become biased towards underrepresented groups or develop some discriminatory characteristics of the patient, exacerbating existing health disparities.

If not correctly represented, elements such as sex and race, understood as geographic origin, can have consequences on medical imaging: think of an AI system trained with the physical characteristics of just asian groups applied in other geographic areas in which the AI tools have to analyse images in which the height of the patient has some medical consequence, as in the case of delineation of body surface area. before the radio treatment.

We can use -as an example- how it is not intuitive how an AI algorithm could realise socioeconomic discrimination in the field of radiomics. The starting point is that during the collection and preparation of imaging databases for the development and testing of new AI algorithms, the health personal data of poor people, and the homeless are missing, and thus some specific illnesses relate to those situations (i.f deriving from the abuse of alcohol). and the data has a bias to exacerbate existing health disparities, just because the health systems don't collect their health personal data, because for several reasons they fall the medical control. To face the problem of fairness in AI, 'what needs to be avoided is a situation where system developers and controllers alone set regulatory thresholds for discrimination at the local and subjective level without external regulatory or judicial input', on the other side, because the application of the traditional principle of direct and indirect discrimination is abstract and not intuitive, a new generation of jurists must be prepared through an interdisciplinary collaboration.

Conflict of interest
The author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Sources of Funding
This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement nº 952159 (ProCancer-I).

References

1. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). European Union [Internet]. Jun. 12, 20124 [cited 2024 Sep 30]. Available from: https://eur-lex.europa.eu/eli/reg/2024/1689/oj

2. Klamert, Marcus, and Alexandre-Xavier-Pierre Lewis. Article 28 TFEU. In: Kellerbauer M, Klamert M, Tomkin J (eds). The EU Treaties and the Charter of Fundamental Rights: A Commentary. New York: Oxford Academic [Internet], 2019. [cited 2024 Sep 30]. Available from: https://doi.org/10.1093/oso/9780198759393.003.104

3. OECD. Best Practice Guidelines for Biological Resource Centers. Paris: OECD Publications, 2007. Available from: https://www.oecd-ilibrary.org/science-and-technology/oecd-best-practice-guidelines-for-biological-resource-centres_9789264128767-en

4. European Strategy Forum for Research Infrastructures. ROADMAP 2011. Strategy Report on Research Infrastructure. Milano: ESFRI, Dipartimento di Fisica – Università degli Studi di Milano, 2021, 243 p. Available from: https://roadmap2021.esfri.eu/media/1295/esfri-roadmap-2021.pdf

5. Perry M. Accessing accessions, biobanks and benefit-sharing. In: Pascuzzi G, Izzo U, Macilotti M. Biobanks. Springer: 2013, p. 267.

6. European Society of Radiology (ESR). ESR Position Paper on Imaging Biobanks. Insights Imaging [Internet]. 2015 Aug. [cited 2024 Sep 30]; 6(4):403-10. DOI: 10.1007/s13244-015-0409-x

7. Borgheresi R, Barucci A, Colantonio S, Aghakhanyan G, Assante M, Bertelli E, et al. NAVIGATOR: an Italian regional imaging biobank to promote precision medicine for oncologic patients. Eur Radiol Exp [Internet]. 2022 Nov. [cited 2024 Sep 30]; 8;6(1):53. DOI:10.1186/s41747-022-00306-9.

8. Gaskell G, Allum N, Bauer M, Durant J, Allansdottir A, Bonfadelli H et al. Biotechnology and the European public. Nat Biotechnol [Internet]. 2000 [cited 2024 Sep 30]; 18(9):935-8. DOI: 10.1038/79403

9. Slokenberga S, Tzortzatou O, Reichel J. GDPR and Biobanking: Individual Rights, Public Interest and Research Regulation across Europe. Cham Springer. 2021, p. 1.

10. Wright J, Ploem C, Śliwkac M, Geversb S. Regulating Tissue Research: Do We Need Additional Rules to Protect research Perticipants? European Journal of Healt Law [Internet]. 2010 [cited 2024 Sep 30]; 17(5):455-469. Available from: https://doi.org/10.1163/157180910X525295

11. European Commission. Communication From The Commission To The European Parliament, The European Council, The Council, The European Economic And Social Committee And The Committee. Of The Regions Artificial Intelligence for Europe COM(2018) 237 final. European Union [Internet]. 2018 [cited 2024 Sep 30]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN

12. European Commission. Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions Empty Building Trust In Human-Centric Artificial Intelligence. Com (2019) 168 Final. European Union [Internet]. 2019 [cited 2024 Sep 30]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52019DC0168

13. European Commission. WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust. Com (2020) 65 final. European Union [Internet]. 2020 [cited 2024 Sep 30]. Available from: https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-artificial-intelligence-feb2020_en.pdf

14. European Commission. 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. European Union [Internet]. Apr. 21, 2021 [cited 2024 Sep 30]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

15. European Commission. ALTAI portal. Futurium [Internet]. 2024 [cited 2024 Sep 30]. Available from: https://futurium.ec.europa.eu/en/european-ai-alliance/pages/altai-assessment-list-trustworthy-artificial-intelligence

16. European Commission. Ethics guidelines for trustworthy AI. Publications Office [Internet]. 2019 [cited 2024 Sep 30]. Available from: https://data.europa.eu/doi/10.2759/346720

17. Permanent Representatives Committee. 5808/19 COMPET 75 MI 70 IND 22 DIGIT 15 JUSTCIV 26 RECH 56 EDUC 33. Artificial intelligence: b) Conclusions on the coordinated plan on artificial intelligence-Adoption. Council of the European Union.

2019 [cited 2024 Sep 30]. Available from: https://data.consilium.europa.eu/doc/document/ST-6177-2019-INIT/en/pdf

18. Sandra Wachter, Brent Mittelstadt, Chris Russell, Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. Computer Law & Security Review [Internet]. 2021 [cited 2024 Sep 30]; 41: 1-72 Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547922

19. Lekadir K, Osuala R, Gallin C, Lazkrak N, Kushibar K, Tsakou G, et al. FUTURE-AI: Guiding Principles and Consensus Recommendations for Trustworthy Artificial Intelligence in Medical Imaging. Arxiv [Internet]. 2024 [cited 2024 Sep 30]; 1: 1-54. Available from: https://arxiv.org/abs/2309.12325

20. Larrazabal AJ, Nicolás Nieto, Peterson V, Milone DH, Ferrante E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proceedings of the National Academy of Sciences [Internet]. 2020 [cited 2024 Sep 30]; 117(23):12592–12594. DOI: 10.1073/pnas.1919012117

21. Puyol-Antón E, Ruijsink B, Piechnik SK, Neubauer S, Petersen SE, Razavi R, et al. Fairness in cardiac mr image analysis: An investigation of bias due to data imbalance in deep learning-based segmentation. In: de Bruijne M, Cattin PC, Cotin S, Padoy N, Speidel S, Zheng Y, et al. Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. MICCAI 2021. Lecture Notes in Computer Science [Internet]: Springer. 2021 [cited 2024 Sep 30]. Available from: https://doi.org/10.1007/978-3-030-87199-4_39

22. L Burzagli, User. In the context of biobanks who are the users and which factors relate to the human component to consider? (in printing). Arnold R, Cippitani R, Colcelli V, Brochhausen C (eds.). Paper accepted for the publication in GDPR Requirements for Biobanking Activities Across Europe Handbook. Cham: Springer International Publishing, 2022.

23. Geis JR, Brady AP, Wu CC, Spencer J, Ranschaert E, Jaremko JL, et al. Ethics of Artificial Intelligence in Radiology: Summary of the Joint European and North American Multisociety Statement. Canadian Association of Radiologists Journal [Internet]. 2019 [cited 2024 Sep 30]; 70(4), 329-334. Available from: https://doi.org/10.1016/j.carj.2019.08.010, p.3-4

How to cite

Colcelli V. The European Union legal framework for using artificial intelligence and imaging databases and imaging biobanks for research purposes: applying the notion of Fairness. Cadernos Ibero-Americanos de Direito Sanitário. 2024 oct./dec.;13(4):120-136
https://doi.org/10.17566/ciads.v13i4.1288